

# Faithful Implementations of Distributed Algorithms and Control Laws

Takashi Tanaka\*, Farhad Farokhi†, Cédric Langbort‡

\*LIDS, Massachusetts Institute of Technology, USA ttanaka@mit.edu

†Dept. of Electrical and Electronic Eng., University of Melbourne, Australia ffarokhi@unimelb.edu.au

‡Coordinated Science Lab., University of Illinois at Urbana-Champaign, USA langbort@illinois.edu

**Abstract**—When a distributed algorithm must be executed by strategic agents with misaligned interests, a social leader needs to introduce an appropriate tax/subsidy mechanism to incentivize agents to faithfully implement the intended algorithm so that a correct outcome is obtained. We discuss the incentive issues of implementing economically efficient distributed algorithms using the framework of indirect mechanism design theory. In particular, we show that indirect Groves mechanisms are not only sufficient but also necessary to achieve incentive compatibility. This result can be viewed as a generalization of the Green-Laffont theorem to indirect mechanisms. Then we introduce the notion of asymptotic incentive compatibility as an appropriate solution concept to faithfully implement distributed and iterative optimization algorithms. We consider two special types of optimization algorithms: dual decomposition algorithms for resource allocation and average consensus algorithms.

## I. INTRODUCTION

In this paper, we consider a society comprised of a single leader and  $N$  followers. The leader makes a social decision  $z \in \mathcal{Z}$ ,<sup>1</sup> which incurs cost  $v_i(z; \theta_i)$  to the  $i$ -th follower. For every  $i = 1, 2, \dots, N$ , assume that cost function  $v_i(z; \theta_i)$  has a known parametric model while parameters  $\theta_i \in \Theta_i$  are private. (For instance,  $v_i(z; \theta_i)$  can be a 10-th order polynomial of a scalar variable  $z$  whose coefficients  $\theta_i \in \mathbb{R}^{10}$  are private.) The leader desires to make a social decision  $z^*$  that minimizes the sum of the followers' individual costs;

$$z^* \in \arg \min_{z \in \mathcal{Z}} \sum_{i=1}^N v_i(z; \theta_i). \quad (1)$$

A social decision satisfying (1) is said to be (*economically efficient*). Efficient decision-making requires distributed algorithms involving leader-follower communication, since the leader has no access to the private parameters. A challenge here is that such decision mechanisms must be designed so that self-interested followers are given no incentive to manipulate the algorithm in an effort to minimize their individual costs. Note that this requirement is different from the fault resilience requirement considered in, for instance, the Byzantine generals problem [2]. We are interested in decision mechanisms in which the followers could manipulate the result, but they choose not to do so.

Roughly speaking, the task of the leader is to design a game that produces an efficient decision  $z^*$  as a consequence of game-theoretic equilibrium strategies of the followers. Designing such games systematically in various multi-agent decision making situations (e.g., auctions, elections, resource allocations), frequently using a carefully designed tax/subsidy rule, is a subject of interest in the Economics literature under the umbrella of *mechanism design* theory. Developments since the 1970s in mechanism design theory have resulted in a rich and established discipline; basic information about mechanism design theory can be found in, e.g., [3]–[8]. One of the best-known positive results is the Groves mechanism, which provides clear guidelines to design tax/subsidy rules incentivizing the followers to be collaborative in the process of computing efficient decisions.

Recently, the theory of mechanism design has been applied to various engineering and computer scientific problems. These applications have raised new challenges to the traditional mechanism design theory, in term of computational difficulties (e.g., combinatorial auctions [9] [10], job scheduling [11]) and communication difficulties (e.g., inter-domain routing [12]). For instance, the standard Groves mechanism becomes computationally intractable if the optimization problem (1) is NP-hard. In such cases, the goal of mechanism design must be set alternatively to incentivize strategic followers to act faithfully in a (computationally feasible) algorithm that only approximates an optimal solution. It turns out that this task is not straightforward, since mechanisms that naively approximate the Groves mechanism are in general *not* “approximately incentive compatible” *at all* [13] [14]. This implies that a fundamental departure from the Groves mechanism is inevitable when the underlying optimization problem is computationally hard. The interplay between incentives, computation, and communication now forms the field of *algorithmic mechanism design (AMD)* [14] [15]. It should be noted that several recent papers discuss similar ideas regarding game designs for distributed optimization/control without referring to the AMD theory explicitly [16]–[18]. However, their intrinsic connections to the mechanism design theory should be clarified to facilitate the further developments beyond their current problem-specific nature.

The simplest approach that the leader can take to obtain a solution in (1) is to incentivize followers to report their private parameters  $\theta_i$  truthfully, so that the leader can solve the optimization problem using the central computer. This

An early version of this article was presented at the 52nd IEEE Conference on Decision and Control [1].

<sup>1</sup>For ease of presentation, we assume  $\mathcal{Z} \subset \mathbb{R}^{nz}$ .

particular type of decision making procedure, called *direct revelation mechanism*, has been the main focus in the mechanism design literature. In many cases, this concentrated interest can be justified by the *revelation principle* [4], which proves the existence of an incentive compatible direct mechanism for every incentive compatible indirect mechanism, and thus guarantees no loss of generality with focusing only on direct mechanisms. However, as recognized in the AMD literature, there are many situations in which the revelation principle should not be naively relied on:

- i) In a direct revelation mechanism, the leader is solely responsible for solving a possibly large-scale optimization problem. Even if there exist more practical distributed algorithms, direct mechanisms do not allow distributed implementations of algorithms by the followers;
- ii) Reporting  $\theta_i$  is not a trivial action: It may be difficult for the followers to identify their own cost function;
- iii) Reporting  $\theta_i$  means a complete loss of privacy;
- iv) Direct and secure communication links between the leader and the followers are not always available<sup>2</sup>;
- v) In homogeneous environments (e.g., internet), it might be difficult to establish a leader<sup>3</sup>.

To resolve these issues, it is invaluable to develop a general guideline to design distributed algorithms that induce *truthful actions* by the followers. This requirement is far more general than the one in the direct mechanism regime, where only *truthful reports* are considered. The possibility of such generalization is foreseen by several encouraging results. In [22], it is shown that a natural generalization of the Groves mechanism to the indirect mechanisms (referred to as indirect Groves mechanisms in this paper) implements a socially optimal set of strategies in *ex-post* Nash equilibria. The idea is employed in [23], where several concrete distributed algorithms are shown to be faithfully implementable by strategic agents. Coordination of strategic agents in a dynamic decision making process is considered in [24].

The approach of [22] is particularly attractive since, unlike many results in AMD that are problem-specific, the result there is applicable to a wide range of distributed computation and communication protocols. In this paper, we pursue the same direction of research and make several additional observations that are essential especially when deploying these results in distributed numerical optimization and control problems. We present the following technical contributions in this paper.

1) *Necessity of indirect Groves mechanisms*: We prove in Theorem 2 that incentive compatible indirect mechanisms must be in the class of indirect Groves mechanisms whenever the space  $\Theta_i$  for some  $i$  is rich enough to parametrize all quadratic cost functions. This is an extension of the Green-Laffont Theorem [25] to the indirect mechanism setting.

<sup>2</sup>This issue is raised in [19]. See also Remark 3.

<sup>3</sup>Since we will always assume that a leader exists in this paper, the item v) is beyond our scope. However, we note that there are a few successful mechanism design examples under such environments, including multi-cast cost sharing [20] and interdomain routing [12]. Some important results in distributed algorithmic mechanism design (DAMD) as of 2002 are summarized in a review paper [21].

2) *Asymptotic incentive compatibility*: We introduce this solution concept to justify the use of approximated Groves taxes to incentivize followers to “act right” in a wide class of optimization algorithms, including continuous optimization algorithms. Due to the nature of the continuous optimization, the exact solution cannot be obtained in finite time and hence, Groves taxes must be inevitably approximated as well. However, it has not been fully discussed in the literature whether the use of approximated Groves mechanisms in this context is justifiable or not. We argue that the use of approximated Groves mechanisms is justifiable whenever we have an iterative distributed algorithm, which can be iterated as many times as we wish, and we can compute approximated Groves taxes from its output that diminish the followers’ incentives for cheating to an arbitrary small  $\epsilon$ . We believe such a situation is satisfactory to convince followers to “act right” in the algorithm, and hence is a practical solution concept. We name this solution concept “asymptotic incentive compatibility.”

As mentioned earlier, the issue of approximating Groves taxes is well studied in the AMD literature. However, our focus in item 2) above is different. In the AMD literature, the research focus is almost exclusively on discrete optimization with approximation threshold strictly greater than zero (in the language of [26]). In such cases, the research focus must be on non-Groves mechanisms, since Groves mechanisms are computationally impractical. On the other hand, our focus is still on the (indirect) Groves mechanisms. We consider their applications to continuous optimization problems and clarify in what sense a mechanism can be “incentive compatible” in those cases.

This paper is organized as follows. We start with a motivating example in Section II. Section III formally introduces the framework of indirect mechanism design. Section IV develops the notion of asymptotic incentive compatibility. In Sections V and VI, we discuss faithful implementations of dual decomposition algorithms and average consensus algorithms. Section VII contains some additional discussion and conclusions.

## II. MOTIVATING EXAMPLES

Consider a resource allocation problem of the form

$$\min \sum_{i=1}^N v_i(z_i; \theta_i), \quad (2a)$$

$$\text{s.t. } Rz = c. \quad (2b)$$

A vector  $z = (z_1, \dots, z_N) \in \mathbb{R}^{n_{z1}} \times \dots \times \mathbb{R}^{n_{zN}}$  is a concatenation of the social decision variables, and  $R = [R_1 \ \dots \ R_N]$ . Define  $\mathcal{Z}$  to be the set of all  $z$  such that  $Rz = c$ .

Let  $L(z, \lambda) = \sum_{i=1}^N v_i(z_i; \theta_i) + \lambda^\top (Rz - c)$  be the Lagrangian with a Lagrange multiplier  $\lambda$ . The primal-dual optimal solution  $(z^*, \lambda^*)$  is a saddle point of  $L(z, \lambda)$ , and assuming that cost functions are strictly convex, the saddle point value  $L^*$  is equal to the optimal value of (2); see [27]. The following iterations are guaranteed to converge to  $(z^*, \lambda^*)$  if the step size  $\gamma$  is chosen to be sufficiently small [28]:

$$\hat{z}_i^k = \arg \min_{\hat{z}_i} (v_i(\hat{z}_i, \theta_i) + (\lambda^{k-1})^\top R_i \hat{z}_i) \quad (3a)$$

$$\lambda^k = \lambda^{k-1} + \gamma(R\hat{z}^k - c). \quad (3b)$$

Notice that the above algorithm has an attractive form for a distributed implementation since (3a) can be executed by the followers. This type of parallelization is known as *dual decomposition*. By increasing the number of iterations, the optimal social decision can be approximated with an arbitrary accuracy, provided that the followers faithfully implement (3a).

What kind of side payment (tax/subsidy) mechanism do we need to incentivize the followers to execute (3a)? Let us make a first attempt. Suppose that the leader introduces the following format of auction mechanism:

**Step 1:** Each follower has some initial value  $z_i^0$ , and the leader has some initial value  $\lambda^0$ .

**Step 2:** Run iteration (3) until it reaches a convergence to  $(z_1^*, \dots, z_N^*, \lambda^*)$ .

**Step 3:** The leader determines the allocation according to  $(z_1^*, \dots, z_N^*)$ , and each player makes a payment  $p_i^0 = \lambda^* z_i^*$  to the leader.

In microeconomic theory,  $\lambda^*$  is known as the “market-clearing price” under which demand  $\sum_{i=1}^N R_i z_i$  and supply  $c$  are balanced. The above mechanism employs a particular type of tax rule  $p_i^0$  which is very natural: the tax imposed on the  $i$ -th player is calculated by the share he has won times the market-clearing price.

Unfortunately, the above auction mechanism is *not* incentive compatible. It is easy to demonstrate that it is vulnerable to strategic manipulations. Suppose  $N = 2$ ,  $c = 1$ ,  $R_i = 1$ , and  $v_i(z_i; \theta_i) = (z_i - 1)^2$  for  $i = 1, 2$ . If both players follow the suggested algorithm, the iteration reaches the optimal solution  $(z_1^*, z_2^*, \lambda^*) = (1/2, 1/2, 1)$ , which brings a net cost of  $v_i(z_i^*; \theta_i) + \lambda^* z_i^* = 3/4$  to each player. Now, suppose that player 1 bids  $\hat{z}_1^k = 1/3$  at every iteration (instead of executing (3a) faithfully). Then it can be shown that the iteration arrives at a different fixed point  $(z_1', z_2', \lambda') = (1/3, 2/3, 2/3)$ . This result brings net cost of  $v_1(z_1'; \theta_1) + \lambda' z_1' = 2/3$  to player 1, which is less than  $3/4$ . Hence, player 1 is indeed better off by deviating from (3a).

We also note that, when followers are “price-takers” (which is the case, for instance, when every follower has sufficiently small market power and the price cannot be affected by his sole action), it makes sense to assume that each follower executes (3a) in an effort to minimize his own net cost. However, many realistic markets are oligopolistic, in which a stakeholder agent knows that his sole action has a certain effect on the market-clearing price [29]. In this case, he might be better off by “exercising market power” rather than following (3a) as shown in the above example. Analyzing strategic bidding in a given auction mechanism (as in [30]–[32]) is an important topic. In this paper, however, we are more interested in designing mechanisms in which strategic manipulation by a follower brings no benefit to him.

The primal-dual algorithm considered in this section is just a motivating example. The result of the next section is applicable to a much more general class of distributed algorithms.

### III. INDIRECT MECHANISM DESIGN

#### A. Framework

The diagram in Fig. 1 shows an abstract framework for indirect mechanisms. Let  $\Theta = \Theta_1 \times \dots \times \Theta_N$  be the space

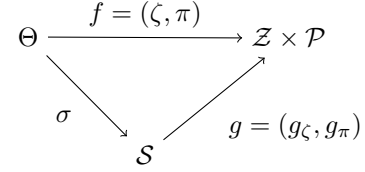


Fig. 1: Framework of indirect mechanisms.

$\Theta = \Theta_1 \times \dots \times \Theta_N$ : The space of private parameters.

$\mathcal{P} \subseteq \mathbb{R}^N$ : The space of tax values  $p = (p_1, \dots, p_N)$ .

$\mathcal{Z} \subseteq \mathbb{R}^{n_z}$ : The space of social decisions  $z$ .

$\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_N$ : The space of actions.

$\zeta : \Theta \rightarrow \mathbb{R}^{n_z}$ : Decision rule.

$\pi : \Theta \rightarrow \mathcal{P}$ : Tax rule.

$\sigma_i : \Theta_i \rightarrow \mathcal{S}_i$ : Strategy function.

$f : \Theta \rightarrow \mathcal{Z} \times \mathcal{P}$ : Social choice function.

$g : \mathcal{S} \rightarrow \mathcal{Z} \times \mathcal{P}$ : Outcome function.

of private parameters. A function  $\zeta : \Theta \rightarrow \mathbb{R}^{n_z}$  is called a *decision rule*. For fixed sets  $\Theta$  and the space  $\mathcal{Z} \subseteq \mathbb{R}^{n_z}$  of social decisions, a decision rule is said to be *efficient* if  $\zeta(\Theta) \subseteq \mathcal{Z}$  and  $\sum_{i=1}^N v_i(\zeta(\theta); \theta_i) \leq \sum_{i=1}^N v_i(z; \theta_i)$  for all  $\theta \in \Theta$  and for all  $z \in \mathcal{Z}$ . Let us also introduce a *tax rule*  $\pi : \Theta \rightarrow \mathcal{P}$  where  $\mathcal{P} \subseteq \mathbb{R}^N$  is the space of tax values assigned to the followers. The pair  $f = (\zeta, \pi)$ ,  $f : \Theta \rightarrow \mathbb{R}^{n_z} \times \mathbb{R}^N$  is called a *social choice function*.

When the leader designs a decision making mechanism, the *action space*  $\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_N$  must be specified, where  $\mathcal{S}_i$  can be thought of as the space of all possible programming codes that the  $i$ -th follower can potentially execute in the distributed computation (for instance, executing (3a) is a valid action in the algorithm considered in Section II, while bidding  $\hat{z}_1^k = 1/3$  at every step is another). The  $i$ -th follower with private parameter  $\theta_i$  determines his actions according to the *strategy function*  $\sigma_i : \Theta_i \rightarrow \mathcal{S}_i$ . Outputs of the followers’ algorithms are processed by the leader’s algorithm called the *outcome function*  $g : \mathcal{S} \rightarrow \mathcal{Z} \times \mathcal{P}$ , which determines a social decision and tax values.

**Remark 1.** For example, the above formulation can express the following abstract model of multi-stage interactions between the leader and the followers. At each stage (indexed by  $k = 1, 2, \dots, n$ ), the leader broadcasts his current computational output  $\eta_L^k$  to the followers. Each follower transmits his current computational output  $\eta_i^k$  to the leader and other followers. Assume that the leader and followers can be modeled as a state-based computer with the internal state  $\xi_L^k$  and  $\xi_i^k, i = 1, \dots, N$  respectively. Given initial states  $\xi_L^0, \xi_i^0$  and  $\eta_L^0, \eta_i^0, i = 1, \dots, N$ , the state evolves according to:

$$\xi_L^k = G_L^k(\xi_L^{k-1}, \eta_1^k, \dots, \eta_N^k); \quad \eta_L^k = H_L^k(\xi_L^k), \quad (4a)$$

$$\xi_i^k = G_{i,\theta_i}^k(\xi_i^{k-1}, \eta_L^{k-1}, \eta_{-i}^{k-1}); \quad \eta_i^k = H_{i,\theta_i}^k(\xi_i^k), \quad (4b)$$

for  $k = 1, 2, \dots, n$ . Finally, we require that  $\eta_L^n = H_L^n(\xi_L^n) = (\zeta(\theta), \pi(\theta))$ , which will be the value of the social choice. In this communication model, the action of the  $i$ -th follower is the sequence of functions in (4b), i.e.,  $\sigma_i(\theta_i) = \{(G_{i,\theta_i}^k, H_{i,\theta_i}^k) : k = 1, 2, \dots, n\}$  parametrized by his type  $\theta_i$ . The outcome function is defined by the sequence of functions in (4a), i.e.,  $g = \{(G_L^k, H_L^k) : k = 1, 2, \dots, n\}$ . For a practical implementation,  $n$  must be a finite number.  $\square$

Formally, a *mechanism* is a triplet  $M = (\sigma, \mathcal{S}, g)$  of a strategy function  $\sigma$ , the space of collective actions  $\mathcal{S}$ , and an outcome function  $g$ . Notice that a mechanism  $M$  suggests followers to employ a particular strategy function specified by  $\sigma$ , but it is followers' choice to be faithful or not; followers are allowed to take any actions in  $\mathcal{S}$ .

A particular case with  $\mathcal{S} = \Theta$ ,  $g = f$ , and  $\sigma = Id$  (i.e., identity map) is called a *direct mechanism*, in which followers are asked to report their types  $\theta_i$  to the leader directly. A mechanism  $M = (\sigma, \mathcal{S}, g)$  is said to be *dominant strategy incentive compatible* if implementing the suggested action  $\sigma(\theta) = (\sigma_1(\theta_1), \dots, \sigma_N(\theta_N)) \in \mathcal{S}$  constitutes dominant strategies among the followers. However, this requirement turns out to be often difficult to attain in indirect mechanism design settings. Hence, we employ a weaker notion of incentive compatibility.

**Definition 1.** A mechanism  $M = (\sigma, \mathcal{S}, g)$  is said to be single fault tolerant if for any  $i$ ,  $s_i \in \mathcal{S}_i$ , and  $\theta \in \Theta$ , the mechanism produces a feasible outcome  $g(s_i, \sigma_{-i}(\theta_{-i})) \in \mathcal{Z} \times \mathcal{P}$ .

**Definition 2.** A mechanism  $M = (\sigma, \mathcal{S}, g)$  is said to be incentive compatible if  $\forall i, \forall s_i \in \mathcal{S}_i, \forall \theta \in \Theta$ ,

$$\begin{aligned} & v_i(g_\zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) \\ & \leq v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})). \end{aligned}$$

In this case, the mechanism is said to implement a social choice function  $f = g \circ \sigma$  in ex-post Nash equilibria<sup>4</sup>.

Intuitively, single fault tolerance requires the mechanism to make a valid social decision (if not optimal) even if at most one follower did not implement suggested strategies faithfully. Incentive compatibility requires that no follower is incentivized to deviate from the suggested strategy if all other followers faithfully implement suggested strategies.

## B. Indirect Groves mechanism

In this paper, we focus on *efficient* distributed algorithms (i.e., those that minimize social cost), and designing a tax rule that induces the followers' faithful actions in such algorithms. The question is rephrased as follows: Given a pair  $(g_\zeta, \sigma)$  such that  $g_\zeta \circ \sigma$  is efficient, how can we design  $g_\pi$  so that  $M = (\sigma, \mathcal{S}, g)$  is incentive compatible?

**Definition 3.** A mechanism  $M = (\sigma, \mathcal{S}, g)$  is said to be in the class of indirect Groves mechanisms if, for every  $i \in \{1, \dots, N\}$ , there exists a function  $k_i : \mathcal{S}_{-i} \rightarrow \mathbb{R}$  satisfying:

- For every  $\theta_{-i} \in \Theta_{-i}$  and  $s_i \in \sigma_i(\Theta_i)$ , the tax rule is

$$\begin{aligned} & g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) \\ & = \sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(\sigma_{-i}(\theta_{-i})) \end{aligned} \quad (5)$$

<sup>4</sup>The term *ex-post* is commonly used to mean that  $\sigma_i(\theta_i)$  is the best strategy even without knowing  $\theta_{-i}$ . See [15], Section 9.

- For every  $\theta_{-i} \in \Theta_{-i}$  and  $s_i \in \mathcal{S}_i \setminus \sigma_i(\Theta_i)$ , the tax rule satisfies

$$\begin{aligned} & g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) \\ & \geq \sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(\sigma_{-i}(\theta_{-i})). \end{aligned} \quad (6)$$

**Theorem 1.** (Sufficiency) A single fault tolerant mechanism  $M = (\sigma, \mathcal{S}, g)$  with an efficient decision rule  $g_\zeta \circ \sigma$  is incentive compatible if it is in the class of indirect Groves mechanisms.

*Proof:* Suppose on contrary that

$$\begin{aligned} & v_i(g_\zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) \\ & > v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})). \end{aligned}$$

for some  $i, s_i \in \mathcal{S}_i, \theta \in \Theta$ . The equality (5) is applicable on the left hand side, while (5) or (6) is used on the right hand side depending on  $s_i$ . In both cases, the above inequality implies

$$\sum_{i=1}^N v_i(g_\zeta \circ \sigma(\theta); \theta_i) > \sum_{i=1}^N v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_i).$$

Since  $M$  is single fault tolerant,  $g_\zeta(s_i, \sigma_{-i}(\theta_{-i})) \in \mathcal{Z}$ . Hence, this is a contradiction to the efficiency of  $g_\zeta \circ \sigma$ . ■

Theorem 1 is due to [22]. A less trivial fact is that the converse of Theorem 1 also holds when each agent's private parameter space  $\Theta_i$  is rich enough.

**Assumption 1.** For every  $i$  and every quadratic function  $q(\cdot) : \mathcal{Z} \rightarrow \mathbb{R}$ , there exists  $\theta_i \in \Theta_i$  such that  $q(\cdot) = v_i(\cdot; \theta_i)$ .

**Theorem 2.** (Necessity) Suppose Assumption 1 holds. A single fault tolerant mechanism  $M = (\sigma, \mathcal{S}, g)$  with an efficient decision rule  $g_\zeta \circ \sigma$  is incentive compatible only if it is in the class of indirect Groves mechanisms.

*Proof:* Complete proof can be found in Appendix A. The basic idea of the proof is attributed to the celebrated result by Green and Laffont [25]. The proof for Case 1 is a modification of Theorem 10.4.3 in [7]. ■

**Remark 2.** Unlike the Groves mechanism in the direct mechanism design, the indirect Groves mechanism does not generally implement the desired algorithm in a dominant strategy. Indeed, it was shown by Proposition 9.23 in [15] that an indirect mechanism is dominant strategy incentive compatible only if every map  $\sigma_i : \Theta_i \rightarrow \mathcal{S}_i$  is surjective. We will see a concrete example of this fact in Example 1. □

**Remark 3.** The set  $\mathcal{S}$  can be extremely rich, since it is the space of all programs that can be executed by the followers during the course of the algorithm. For instance, followers are allowed to write a code to learn about other followers during the algorithm to make future decisions. However, as an implicit premise for Theorem 1 and 2, we must preclude the followers' ability to "hack the rule of the game." For instance, the intended algorithm must be securely announced to every follower without strategic interventions by other followers. Similarly, the tax value must be securely computed based on the formula (5) and (6) without a danger of manipulation. Such interventions are actually possible if a follower has an opportunity to modify other players' messages [19]. For the

same reason, followers are not allowed to drop out of the game in midstream to escape from a punitive tax.  $\square$

**Remark 4.** Direct communication links between the leader and the followers are not constantly required during the course of the algorithm. For instance, the main body of the average consensus algorithm in Section VI requires only peer-to-peer communications among neighboring followers, but followers still cannot be better off by cheating. However, in order to satisfy the requirement of the previous remark, secure communication links between the leader and followers are assumed at the initial phase (to announce the algorithm) and at the final phase (to calculate taxes securely).  $\square$

### C. Individual rationality, Budget balance

A mechanism is said to be *individually rational* [7] if the net cost  $v_i(g_\zeta(s); \theta_i) + g_{\pi_i}(s)$  is non-positive for every  $i$ . This is a basic requirement for a mechanism that does not incentivize the followers to quit the mechanism, when quitting the mechanism is cost-free for them.

A mechanism is said to be *budget balanced* (resp. *weakly budget balanced*) [7] if the tax income  $\sum_{i=1}^N p_i$  is zero (resp. non-negative).

Unfortunately, there may not exist a mechanism that simultaneously satisfies (1) efficiency, (2) incentive compatibility, (3) individual rationality, and (4) budget balance. The aforementioned result by Green and Laffont [25] shows that the only efficient direct mechanisms that are dominant strategy incentive compatible are Groves mechanisms. This observation allows us to construct a simple example in which no efficient mechanism simultaneously achieves dominant strategy incentive compatibility, weak budget balance, and individual rationality. In a Bayesian setting, [33] demonstrated that there exists a simple exchange environment in which no (*ex post*) efficient mechanism is simultaneously (Bayes-Nash) incentive compatible, weakly budget balanced, and (*ex interim*) individually rational. This result was generalized in [34] using the revenue maximization principle. We also note that a recent study [35] of a particular indirect mechanism shows that no budget balanced mechanism implements efficient decisions in Nash equilibrium.

## IV. ASYMPTOTIC INCENTIVE COMPATIBILITY

In this section, we generalize Theorem 1 so that it is applicable to *approximately* efficient decision rules. This is an important generalization, since in many realistic cases, social decision must be made upon the result of iterative numerical optimizations over continuous decision variables that, if terminated at some finite step, only returns an approximate solution. In such cases, incentives may even be needed to guarantee not only that the suggested algorithm is implemented, but also that the actions taken by the followers lead it to converge. Define  $\text{dist}(z', \mathcal{Z}) = \inf_{z \in \mathcal{Z}} \|z' - z\|_2$ , and let  $\text{Proj}(z'; \mathcal{Z})$  be the projection of  $z'$  onto  $\mathcal{Z}$ .

**Definition 4.** For every  $n \in \mathbb{N}$ , let  $\zeta^n : \Theta \rightarrow \mathbb{R}^{nz}$  be a decision rule. A sequence of decision rules  $\{\zeta^n\}_{n \in \mathbb{N}}$  is said to be asymptotically efficient if  $\text{dist}(\zeta^n(\theta), \mathcal{Z}) \rightarrow 0$  as  $n \rightarrow \infty$

and for every  $\epsilon > 0$ , there exists  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\sum_{i=1}^N v_i(\zeta^n(\theta); \theta_i) \leq \sum_{i=1}^N v_i(z; \theta_i) + \epsilon, \quad \forall \theta \in \Theta, \forall z \in \mathcal{Z}.$$

**Definition 5.** A sequence of mechanisms  $\{M^n\}_{n \in \mathbb{N}}$ ,  $M^n = (\sigma^n, \mathcal{S}^n, g^n)$ , is said to be asymptotically incentive compatible if for every  $\epsilon > 0$ , there exists  $n_0 \in \mathbb{N}$  such that for  $n \geq n_0$ ,

$$\begin{aligned} & v_i(g_\zeta^n(\sigma_i^n(\theta_i), \sigma_{-i}^n(\theta_{-i})); \theta_i) + g_{\pi_i}^n(\sigma_i^n(\theta_i), \sigma_{-i}^n(\theta_{-i})) \\ & \leq v_i(g_\zeta^n(s_i^n, \sigma_{-i}^n(\theta_{-i})); \theta_i) + g_{\pi_i}^n(s_i^n, \sigma_{-i}^n(\theta_{-i})) + \epsilon \end{aligned}$$

$\forall i, \forall s_i^n \in \mathcal{S}_i^n, \forall \theta \in \Theta$ . In this case, the mechanism is said to asymptotically implement a social choice function  $f := \lim_{n \rightarrow \infty} g^n \circ \sigma^n$  in *ex-post* Nash equilibria, if the limit exists.

**Remark 5.** For iterative algorithms,  $n$  can be understood as the number of iterations before termination. Since this only gives an approximation of efficient social decisions,  $M^n$  may not be incentive compatible for a fixed  $n \in \mathbb{N}$ . However, as  $n \rightarrow \infty$ ,  $M^n$  provides every follower a diminishing incentive to deviate from the suggested algorithm. Without loss of generality, we assume that taxes are paid after the algorithm has terminated.  $\square$

The next proposition presents a sequence of mechanisms  $\{M^n\}_{n \in \mathbb{N}}$  motivated by the Groves mechanism by which the followers' incentive to deviate from the intended algorithm can be made arbitrary small.

**Proposition 1.** Let  $\{M^n\}_{n \in \mathbb{N}}$ ,  $M^n = (\sigma^n, \mathcal{S}^n, g^n)$ , be a sequence of single fault tolerant mechanisms such that  $\{g_\zeta^n \circ \sigma^n\}_{n \in \mathbb{N}}$  is asymptotically efficient. If the payment rule is

$$g_{\pi_i}^n(s^n) = \sum_{j \neq i} v_j(g_\zeta^n(s^n); \theta_j) + k_i(s_{-i}^n) \quad \forall s^n \in \mathcal{S}^n,$$

then  $\{M^n\}_{n \in \mathbb{N}}$  is asymptotically incentive compatible.

*Proof:* Suppose there exist  $\epsilon > 0$ ,  $i, \theta \in \Theta$ , a sequence of strategies  $s_i^n \in \mathcal{S}_i^n$ , and a subsequence  $\{n_\ell\}_{\ell \in \mathbb{N}}$  such that

$$\begin{aligned} & v_i(g_\zeta^{n_\ell}(\sigma_i^{n_\ell}(\theta_i), \sigma_{-i}^{n_\ell}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_\ell}(\sigma_i^{n_\ell}(\theta_i), \sigma_{-i}^{n_\ell}(\theta_{-i})) \\ & > v_i(g_\zeta^{n_\ell}(s_i^{n_\ell}, \sigma_{-i}^{n_\ell}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_\ell}(s_i^{n_\ell}, \sigma_{-i}^{n_\ell}(\theta_{-i})) + \epsilon \end{aligned}$$

for all  $\ell \in \mathbb{N}$ . This implies that for all  $\ell \in \mathbb{N}$ ,

$$\begin{aligned} & \sum_{i=1}^N v_i(g_\zeta^{n_\ell}(\sigma_i^{n_\ell}(\theta_i), \sigma_{-i}^{n_\ell}(\theta_{-i})); \theta_i) \\ & > \sum_{i=1}^N v_i(g_\zeta^{n_\ell}(s_i^{n_\ell}, \sigma_{-i}^{n_\ell}(\theta_{-i})); \theta_i) + \epsilon. \end{aligned}$$

Since  $g_\zeta^{n_\ell}(s_i^{n_\ell}, \sigma_{-i}^{n_\ell}(\theta_{-i})) \in \mathcal{Z}$  due to the single fault tolerance, this contradicts the asymptotic efficiency of  $g_\zeta^n \circ \sigma^n$ .  $\blacksquare$

In practice, the result of Proposition 1 is used as follows. First, the leader chooses  $\epsilon > 0$  to which he wishes to diminish the followers' incentives to misbehave. Second, the leader identifies  $n_0 \in \mathbb{N}$  satisfying the condition in Definition 4 by analyzing the asymptotically efficient sequence of decision rules to be implemented. Finally, the leader announces a mechanism  $M^n$  as defined in Proposition 1 with some  $n \geq n_0$ .

As a result, no follower has more than  $\epsilon$  incentive to deviate from the suggested algorithm.

Proposition 1 requires  $M^n$  to be single fault tolerant, i.e., that the social decision made by  $M^n$  be feasible even if one of the agents is misbehaving. In some applications (such as *dual decomposition* algorithms to be considered in Section V), this requirement can be met by simply projecting the intermediate result onto the feasible set. However, this is not always possible (as we will see in Section VI where the *average consensus* algorithm over a cyber-physical system is considered). To circumvent this difficulty, we need a sequence of tax rules  $\{g_{\pi_i}^n\}$  that leads the algorithm to converge to the optimal solution even under the lack of single fault tolerance. Several ideas can be exploited. The following result is useful when it is easy for the leader to observe  $\text{dist}(\cdot, \mathcal{Z})$  and the knowledge of the upper bound  $\beta(n)$  on  $\text{dist}(g_{\zeta}^n(\sigma^n(\theta)), \mathcal{Z})$  is available *a priori*. Roughly speaking, it penalizes all followers if the expected convergence rate (to the feasible set) is not observed.

**Proposition 2.** *Assume that  $\mathcal{Z}$  is a closed set and  $v_i(z, \theta_i)$ ,  $\forall i$ , is a continuous function of  $z$  for all  $\theta_i \in \Theta_i$ . Let  $\{M^n\}_{n \in \mathbb{N}}$ ,  $M^n = (\sigma^n, \mathcal{S}^n, g^n)$ , be a sequence of mechanisms such that (i)  $\{g_{\zeta}^n \circ \sigma^n\}_{n \in \mathbb{N}}$  is asymptotically efficient and (ii)  $\sup_{\theta \in \Theta} \text{dist}(g_{\zeta}^n(\sigma^n(\theta)), \mathcal{Z}) \leq \beta(n)$  for some sequence  $\{\beta(n)\}_{n \in \mathbb{N}}$  with  $\lim_{n \rightarrow \infty} \beta(n) = 0$ . If for any  $s^n \in \mathcal{S}^n$ , a payment rule is chosen as*

$$g_{\pi_i}^n(s^n) = \begin{cases} \sum_{j \neq i} v_j(g_{\zeta}^n(s^n); \theta_j) & \text{if } \text{dist}(g_{\zeta}^n(s^n), \mathcal{Z}) \leq \beta(n) \\ C_i & \text{otherwise} \end{cases}$$

then, for sufficiently large  $C_i$ ,  $\{M^n\}_{n \in \mathbb{N}}$  is asymptotically incentive compatible. In particular, the following choice suffices

$$C_i = - \inf_{\theta_i \in \Theta_i} \inf_{z \in \mathcal{Z}'} v_i(z; \theta_i) + \sup_{\theta \in \Theta} \sup_{z \in \mathcal{Z}'} \sum_{j=1}^N v_j(z; \theta_j), \forall i$$

where  $\mathcal{Z}' = \{g_{\zeta}^n(\sigma^n(\theta)) : \forall \theta \in \Theta\}$ .

*Proof:* The tax rule belongs to the indirect Groves class (Definition 3). Proof can be found in Appendix B. ■

**Remark 6.** Practical usefulness of the notion of asymptotic incentive compatibility heavily depends on the computational complexity of the algorithm. For computationally hard problems, realistically there is no mechanism that reduces undesirable incentive to  $\epsilon$  in polynomial time. In such cases, asymptotic incentive compatibility may not be a convincing reasoning to induce faithful behaviors of followers. However, the issue of computational complexity requires more problem-specific discussions, which is not our focus in this paper. □

## V. FAITHFUL IMPLEMENTATION OF DUAL DECOMPOSITION

### A. Algorithm

Recall the dual decomposition algorithm considered in Section II. Based on the developments so far, we are now going to design a tax rule that incentivize followers to execute (3a) faithfully. The idea is to design an asymptotically incentive compatible sequence of mechanisms  $M^n = (\sigma^n, \mathcal{S}^n, g^n)$ , parameterized by the number of iterations  $n$ . Proposition 1

**Algorithm 1** Distributed VCG mechanism  $M^n = (\sigma^n, \mathcal{S}^n, g^n)$  for dual decomposition algorithms.

**Output:** Allocation decision  $\zeta^n(\theta)$  and tax assignments  $\pi^n(\theta)$

- 1: (L) The leader announces the following algorithm;
- 2: // Solve optimization problem (2) by dual decomposition;
- 3: (L) Initialize and broadcast  $\lambda^0$ ;
- 4: **for**  $k = 1, \dots, n$  **do**
- 5: (F) Find  $\hat{z}_i^k = \arg \min_{z_i} (v_i(z_i; \theta_i) + (\lambda^{k-1})^\top R_i z_i)$  and report the result to the leader;
- 6: (L) Update and broadcast  $\lambda^k = \lambda^{k-1} + \gamma(R \hat{z}^k - c)$ ;
- 7: **end for**
- 8: // Solve marginal problems  $P_{-j}(0)$  for every  $j$ ;
- 9: **for**  $j = 1, \dots, N$  **do**
- 10: (L) Initialize and broadcast  $\lambda^{-j,0}$ ;
- 11: **for**  $k = 1, \dots, n$  **do**
- 12: (F) Every follower  $i (\neq j)$  finds and reports  $\hat{z}_i^{-j,k} = \arg \min_{z_i} (v_i(z_i; \theta_i) + (\lambda^{-j,k-1})^\top R_i z_i)$ ;
- 13: (L) Update and broadcast  $\lambda^{-j,k} = \lambda^{-j,k-1} + \gamma^{-j} (\sum_{i \neq j} R_i \hat{z}_i^{-j,k} - c)$ ;
- 14: **end for**
- 15: **end for**
- 16: // Compute social outcomes;
- 17: (L) Compute and broadcast  $z = \text{Proj}(\hat{z}^n, \mathcal{Z})$  and  $z^{-j} = \text{Proj}(\hat{z}^{-j,n}, \mathcal{Z}^{-j})$  for every  $j = 1, \dots, N$ ;
- 18: (F) Every follower  $i$  computes  $\hat{v}_i = v_i(z_i; \theta_i)$  and  $\hat{v}_i^{-j} = v_i(z_i^{-j}; \theta_i)$  for every  $j \neq i$  and report them to the leader;
- 19: (L) Determine taxes  $\pi_i^n(\theta) = \sum_{j \neq i} \hat{v}_j - \sum_{j \neq i} \hat{v}_j^{-i}$  and allocations  $\zeta^n(\theta) = z$ ;

shows that tax rules attaining our goal are not unique, since the choice of  $k_i$  is arbitrary. In this section, we employ a particular tax rule among them inspired by the VCG mechanism. As we will see in the sequel, this tax rule turns out to be a natural choice since it is intimately related to the notion of “market-clearing prices.”

The VCG mechanism is also called the *pivot mechanism*, since the tax for the follower  $i$  is calculated based on the degree to which his presence/absence changes the social cost [5]. For every  $i \in \{1, \dots, N\}$ , consider the “marginal optimization problem”  $P_{-i}(x)$  defined by

$$P_{-i}(x) : \min \sum_{j \neq i} v_j(z_j; \theta_j) \quad (7)$$

$$\text{s.t. } \sum_{j \neq i} R_j z_j = c - R_i x.$$

Notice that  $P_{-i}(x)$  is the optimization problem (2) in which follower  $i$ 's allocation is fixed at  $z_i = x$ . If his absence means  $z_i = 0$ , the optimal social cost in his absence is obtained by solving  $P_{-i}(0)$ . Let  $(z^*, \lambda^*)$  and  $(z^{-i*}(x), \lambda^{-i*}(x))$  be the primal-dual optimal solution to (2) and (7) respectively. The VCG-like tax for the  $i$ -th follower is defined by

$$p_i^{VCG} = \sum_{j \neq i} v_j(z_j^*; \theta_j) - \sum_{j \neq i} v_j(z_j^{-i*}(0); \theta_j). \quad (8)$$

In Algorithm 1, we propose a VCG-like mechanism for the faithful implementation of the dual decomposition algorithm applied to the resource allocation problem (2). In the iteration, variables  $(\hat{z}^k, \lambda^k)$  and  $(\hat{z}^{-i,k}, \lambda^{-i,k})$  are intended to approximate  $(z^*, \lambda^*)$  and  $(z^{-i*}(0), \lambda^{-i*}(0))$  respectively.

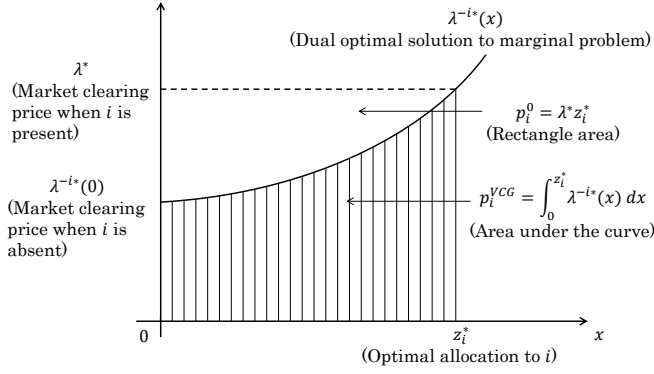


Fig. 2: Graphical interpretation of the VCG taxes.

**Proposition 3.** Assume that  $v_i(\cdot; \theta_i)$ ,  $i = 1, 2, \dots, N$ , are strictly convex for every  $\theta_i \in \Theta_i$ . Then the sequence of mechanisms  $\{M^n\}_{n \in \mathbb{N}}$  provided in Algorithm 1 is asymptotically incentive compatible.

*Proof:* See Appendix C. ■

Notice that lines 8–15 in Algorithm 1 are devoted to calculating  $k_i(s_{-i}^n)$  in Proposition 1. This is just one example of such function (which is motivated by taxes in VCG mechanisms). Removing the above mentioned lines from Algorithm 1 and setting some other values to  $k_i(s_{-i}^n)$  (e.g.,  $k_i(s_{-i}^n) = 0$ ) still attains asymptotic incentive compatibility. To relax strict convexity assumption, one may alternatively consider the ADMM algorithm [36] (distributed implementation of the augmented Lagrangian algorithm [37]) in Algorithm 1.

**Example 1.** As mentioned earlier, indirect Groves mechanisms in general implement efficient decision rules in *ex-post* Nash equilibria but not in dominant strategies. As an example, consider  $\min_{z_1, z_2} \sum_{i=1,2} v_i(z_i)$  s.t.  $z_1 = z_2$  where  $v_1(z_1) = (z_1 - 1)^2$  and  $v_2(z_2) = (z_2 - 2)^2$ , and the dual decomposition algorithm (Algorithm 1) is used. Suppose that the second player chooses to act as  $\hat{z}_2^k = \arg \min_{z_2} (\hat{v}_2(z_2) + \lambda^{k-1} z_2)$  where  $\hat{v}_2(z_2) = (z_2 - 3)^2$  instead of line 5 in the algorithm. If the first player executes the algorithm faithfully,  $\hat{z}_1^k$  and  $\hat{z}_2^k$  converge to 2, and the first player’s net cost converges to  $v_1(2) + p_1 = v_1(2) + v_2(2) = 1$ . However, if the first player deviates from the suggested algorithm and executes  $\hat{z}_1^k = \arg \min_{z_1} (\hat{v}_1(z_1) + \lambda^{k-1} z_1)$  where  $\hat{v}_1(z_1) = z_1^2$  instead of line 5 in the algorithm,  $\hat{z}_1^k$  and  $\hat{z}_2^k$  converge to 1.5, and the first player’s net cost converges to  $v_1(1.5) + p_1 = v_1(1.5) + v_2(1.5) = 0.5$ . Hence, the first player is better off by not following the intended algorithm. This means that faithful execution is not a dominant strategy. □

### B. Connection between VCG and clearing prices

Invoking Theorem 2, it is now clear why the “clearing price” mechanism considered in Section II fails to be incentive compatible. This is simply because the “clearing price” mechanism is not in the class of indirect Groves mechanisms. However, it can be shown that the “clearing price” mechanism is *approximately* incentive compatible under the pure competition

[29] (i.e., when individual followers have negligible market power to control market-clearing prices). We show this fact by pointing out an intimate connection between the tax rule  $p_i^0 = \lambda^* z_i^*$  considered in Section II and the VCG-like tax in (8). For simplicity, we assume that cost functions  $v_i$  are strongly convex and continuously differentiable.

To see a connection, for each  $i$ , consider a smooth path  $x_i : [0, 1] \rightarrow \mathbb{R}^{n_{z_i}}$  defined by  $x_i(t) = t z_i^*$ . Intuitively, the path  $x_i$  continuously connects the follower  $i$ ’s allocations in two distinct situations:  $x_i(0) = 0$  corresponds to the case where follower  $i$  is absent (allocation is zero), and  $x_i(1) = z_i^*$  is the optimal allocation in (2). For every point on the path  $x_i(t)$ , it is possible to consider the primal-dual optimal solution  $(z^{-i^*}(x_i(t)), \lambda^{-i^*}(x_i(t)))$  to the marginal optimization problem (7). In particular,  $\lambda^{-i^*}(x_i(t))$  shows how the market-clearing price determined by the rest of society (excluding  $i$ ) changes if follower  $i$ ’s allocation is fixed at different values  $x_i(t)$  between 0 and  $z_i^*$ .

**Proposition 4.** The VCG payment for the  $i$ -follower is obtained by integrating the market-clearing price  $\lambda^{-i^*}$  along the path  $x_i$ , i.e.,  $p_i^{VCG} = \int_0^{z_i^*} \lambda^{-i^* \top}(x_i) R_i dx_i$ .

*Proof:* For notational ease, dependency of  $v_i$  on  $\theta_i$  is suppressed. Notice that

$$\begin{aligned} \int_0^{z_i^*} \lambda^{-i^* \top}(x_i) R_i dx_i &= \int_0^1 \lambda^{-i^* \top}(x_i(t)) R_i \frac{dx_i(t)}{dt} dt \\ &= - \sum_{j \neq i} \int_0^1 \lambda^{-i^* \top}(x_i(t)) R_j \frac{dz_j^{-i^*}(x_i(t))}{dt} dt \quad (9) \\ &= \sum_{j \neq i} \int_0^1 \nabla_{z_j} v_j(z_j^{-i^*}(x_i(t))) \frac{dz_j^{-i^*}(x_i(t))}{dt} dt \quad (10) \\ &= \sum_{j \neq i} [v_j(z_j^{-i^*}(z_i^*)) - v_j(z_j^{-i^*}(0))] \quad (11) \\ &= \sum_{j \neq i} v_j(z_j^*) - \sum_{j \neq i} v_j(z_j^{-i^*}) = p_i^{VCG} \end{aligned}$$

The identity  $R_i \frac{dx_i(t)}{dt} + \sum_{j \neq i} R_j \frac{dz_j^{-i^*}(x_i(t))}{dt} = 0$  is used in (9) and (10) is from the optimality condition  $\nabla_{z_j} v(z_j^{-i^*}(x)) + \lambda^{-i^* \top}(x) R_j = 0, \forall j \neq i$ . The fundamental theorem of calculus is used in (11). ■

A pictorial interpretation of  $p_i^{VCG}$  is shown in Fig. 2 (for simplicity, assume  $R_i = 1$  for every  $i$  in the resource allocation problem (2) in Section II). This figure shows that  $p_i^{VCG} \approx p_i^0$  when the function  $\lambda^{-i^*}(x_i(\cdot))$  is nearly constant. In other words, when the individual followers have negligible market power, the “clearing price” mechanism can be identified with the VCG mechanism.

### VI. FAITHFUL IMPLEMENTATION OF AVERAGE CONSENSUS

Consider a multi-robot rendezvous problem in which  $N$  robots want to meet in a single position (i.e., achieve a consensus in space). Considering that each robot utilizes fuel/battery to reach the rendezvous point, the social planner may want them to end up at a point that minimizes the sum of their distances from starting points, i.e., the average of the initial

positions denoted by  $\theta_i$ . Here, we design a mechanism that can be used by the social planner (leader) to coordinate robots (followers) to faithfully implement a distributed algorithm, which lead them to the rendezvous point. In particular, we consider an iterative *average consensus* algorithm [38] in which robots are required to communicate with neighboring robots in each iteration. Notice that an incentive design is needed in such situations, since otherwise a particular robot may choose to stand stationary, hoping that all other robots will move towards it thereby, not using any fuel. Due to the nature of the average consensus algorithms, the robots do not achieve exact consensus in their position in finite time. Moreover, if we consider the position  $\zeta^n(\theta)$  of robots after  $n$  iterations of communication as the social decision, the decision rule  $\zeta^n(\theta)$  with finite  $n$  may not yield a feasible solution. Hence the mechanism cannot be made single fault tolerant, and Proposition 1 is not applicable. Instead, we use the result of Proposition 2. Note that the existence of a leader does not mean that communications are required between the leader and the followers at every iteration, and hence does not ruin the advantage of distributed consensus algorithms. Indeed, in our design, a consensus is formed solely by local communications among robots, and the leader plays its role only at the beginning (to announce the tax rule and the algorithm to be implemented) and at the end (to compute taxes based on the information about the final positions of robots).

In what follows, we formally introduce a mechanism that asymptotically implements the average consensus algorithm. Let an undirected graph  $\mathcal{G} = (\{1, \dots, N\}, \mathcal{E})$ , with vertex set  $\{1, \dots, N\}$  and edge set  $\mathcal{E}$ , be given to illustrate the communication links between the agents. Following [39], we can achieve average consensus by solving

$$\min_{z \in \mathbb{R}^N} \sum_{i=1}^N (z_i - \theta_i)^2, \quad (12a)$$

$$\text{s.t. } z_i = z_j, \forall (i, j) \in \mathcal{E}, \quad (12b)$$

where  $z_i \in \mathbb{R}$  is the decision variable of follower  $i$  and  $\theta_i \in \Theta_i \subseteq \mathbb{R}$  is its type. Let us define the incidence matrix  $B \in \{-1, 0, +1\}^{N \times |\mathcal{E}|}$  so that  $b_{ij} = 1$  if  $e_j \in \mathcal{E}$  leaves vertex  $i$ ,  $b_{ij} = -1$  if  $e_j \in \mathcal{E}$  enters vertex  $i$ , and  $b_{ij} = 0$  otherwise (assignments of directions to edges are arbitrary).

**Proposition 5.** *Let  $\mathcal{G}$  be a tree. The sequence of mechanisms  $\{M^n\}_{n \in \mathbb{N}}$  provided in Algorithm 2 is asymptotically incentive compatible.*

*Proof:* Proof can be found in Appendix D. ■

## VII. DISCUSSION AND CONCLUSIONS

We have discussed a general indirect mechanism design framework for faithful distributed algorithm implementations. As examples, we have considered dual decomposition and average consensus algorithms.

The framework of this paper is directly applicable to many *distributed control problems*. Although the issue of incentive is usually neglected in the control theory literature, it is an important challenge that always arises when distributed agents

**Algorithm 2** Distributed mechanism  $M^n = (\sigma^n, \mathcal{S}^n, g^n)$  for asymptotically implementing the average consensus.

**Output:** Consensus decision  $\zeta^n(\theta)$  and tax assignment  $\pi^n(\theta)$

- 1: (L) The leader announces the following algorithm;
- 2: // Solve optimization problem (12);
- 3: (L) Set  $\alpha \in (0, 1/\lambda_{\max}(B^T B))$  and broadcast it;
- 4: (L) Set  $\rho = \lambda_{\min}(B^T B)/\lambda_{\max}(B^T B)$ ;
- 5: (F) Initialize  $z_i^0 = \theta_i$  for each  $1 \leq i \leq N$ ;
- 6: **for**  $k = 1, \dots, n$  **do**
- 7: (F) Calculate  $z_i^k = z_i^{k-1} + \alpha \sum_{j \in \mathcal{N}_i} (z_j^{k-1} - z_i^{k-1})$ , where  $\mathcal{N}_i$  is the neighbors of agent  $i$  in  $\mathcal{G}$ , and transmit to neighbors;
- 8: **end for**
- 9: (F) Transmit  $z_i^n$  and  $\hat{v}_i = (z_i^n - \theta_i)^2$  to the leader;
- 10: // Compute social outcomes;
- 11: (L) Determine and broadcast  $\zeta^n(\theta) = z^n$ ;
- 12: **if**  $\text{dist}(z^n, \mathcal{Z}) \leq (1-\rho)^n \|B(B^T B)^{-1}\|_2 \sup_{q \in \Theta} \|B^T q\|_2$  **then**
- 13: (L) Determine  $\pi_i^n(\theta) = \sum_{j \neq i} \hat{v}_j$ ;
- 14: **else**
- 15: (L) Determine  $\pi_i^n(\theta) = \sup_{\theta \in \Theta} \sup_{z \in \mathcal{Z}'} \sum_j (z_j - \theta_j)^2$ , where  $\mathcal{Z}' = \mathcal{Y}^N$  with  $\mathcal{Y} = \{\sum_j \alpha_j \theta_j : \sum_j \alpha_j = 1, \alpha_j \geq 0, \theta_j \in \Theta_j, \forall j\}$ .
- 16: **end if**

are strategic. For example, a distributed control algorithm proposed in [40] assumes that agents are faithful to the algorithm. However, this requirement can be removed by introducing the tax mechanism in Algorithm 1.

In the future, we may also consider the faithful implementations of *distributed model predictive control* (MPC) algorithms [41]. Distributed MPC is expected to be a powerful tool in large-scale social engineering problems (e.g., operations of power systems [42] and transportation systems [43]). Faithful implementations of distributed MPC requires online (real-time) mechanisms. Several appropriate modifications need to be made to the current framework (e.g., replacement of the solution concept from Nash equilibrium to Markov perfect equilibrium). Further study will be required in this research direction. We believe this is a great opportunity for economic theory (i.e., mechanism design) and engineering (i.e., control) to merge in order to tackle challenging problems in the society.

## APPENDIX

### A. Proof of Theorem 2

**Case 1:** We will first show that the tax rule must be in the form of (5) when  $s_i \in \sigma_i(\Theta_i)$ . Proof is by contradiction. Suppose that  $g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) = \sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(s_i, \sigma_{-i}(\theta_{-i}))$  where there exist  $s_i, s'_i \in \sigma_i(\Theta_i)$  such that

$$k_i(s_i, \sigma_{-i}(\theta_{-i})) \neq k_i(s'_i, \sigma_{-i}(\theta_{-i})). \quad (13)$$

(Step 1): Suppose  $g_\zeta(s_i, \sigma_{-i}(\theta_{-i})) = g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))$ . Let  $\theta_i, \theta'_i \in \Theta_i$  satisfy  $s_i = \sigma_i(\theta_i)$  and  $s'_i = \sigma_i(\theta'_i)$ . If player  $i$ 's true parameter is  $\theta_i$ , acting  $s_i$  minimizes his net cost since  $M$  is incentive compatible. Thus,

$$\begin{aligned} v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) \\ \leq v_i(g_\zeta(s'_i, \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(s'_i, \sigma_{-i}(\theta_{-i})) \end{aligned}$$

or  $g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) \leq g_{\pi_i}(s'_i, \sigma_{-i}(\theta_{-i}))$ . On the other hand, if player  $i$ 's true parameter is  $\theta'_i$ , it must be that



$g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) \geq g_{\pi_i}(s'_i, \sigma_{-i}(\theta_{-i}))$ . Since both cases could occur, the only possibility is

$$g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) = g_{\pi_i}(s'_i, \sigma_{-i}(\theta_{-i})). \quad (14)$$

The left hand side of equation (14) can be written as  $\sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(s_i, \sigma_{-i}(\theta_{-i}))$ , while the right hand side is  $\sum_{j \neq i} v_j(g_\zeta(s'_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(s'_i, \sigma_{-i}(\theta_{-i}))$ . Since we are assuming  $g_\zeta(s_i, \sigma_{-i}(\theta_{-i})) = g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))$ , equation (14) implies

$$k_i(s_i, \sigma_{-i}(\theta_{-i})) = k_i(s'_i, \sigma_{-i}(\theta_{-i})).$$

This is a contradiction to (13).

(Step 2): Now we can assume  $g_\zeta(s_i, \sigma_{-i}(\theta_{-i})) \neq g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))$ . Without loss of generality, we can also assume  $k_i(s_i, \sigma_{-i}(\theta_{-i})) < k_i(s'_i, \sigma_{-i}(\theta_{-i}))$ . Then,

$$k_i(s_i, \sigma_{-i}(\theta_{-i})) < k_i(s'_i, \sigma_{-i}(\theta_{-i})) - \epsilon \quad (15)$$

for some  $\epsilon > 0$ . For every  $j \neq i, 1 \leq j \leq N$ , assume  $v_j(\cdot; \theta_j)$  are quadratic functions. Since  $\Theta_i$  exhausts the space of all quadratic functions, there exists  $\theta''_i \in \Theta_i$  such that

$$v_i(z; \theta''_i) = - \sum_{j \neq i} v_j(z; \theta_j) + \epsilon \left( \frac{\|z - g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))\|^2}{\|g_\zeta(s_i, \sigma_{-i}(\theta_{-i})) - g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))\|^2} - 1 \right). \quad (16)$$

By incentive compatibility, player  $i$  whose true type is  $\theta''_i$  attains smaller net cost by acting  $\sigma_i(\theta''_i)$  than acting  $s_i$ . (Here,  $\sigma_i(\theta''_i)$  and  $s_i$  may or may not be equal.)

$$v_i(g_\zeta(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})); \theta''_i) + g_{\pi_i}(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})) \leq v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta''_i) + g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})). \quad (17)$$

By efficiency, the outcome  $z$  of the mechanism minimizes

$$v_i(z; \theta''_i) + \sum_{j \neq i} v_j(z; \theta_j) = \epsilon \left( \frac{\|z - g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))\|^2}{\|g_\zeta(s_i, \sigma_{-i}(\theta_{-i})) - g_\zeta(s'_i, \sigma_{-i}(\theta_{-i}))\|^2} - 1 \right).$$

Thus, the decision by the mechanism is

$$g_\zeta(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})) = g_\zeta(s'_i, \sigma_{-i}(\theta_{-i})). \quad (18)$$

Substituting (18) into (17) gives

$$v_i(g_\zeta(s'_i, \sigma_{-i}(\theta_{-i})); \theta''_i) + g_{\pi_i}(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})) \leq v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta''_i) + g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})).$$

Rewriting the above relation using (16),

$$\begin{aligned} & - \sum_{j \neq i} v_j(g_\zeta(s'_i, \sigma_{-i}(\theta_{-i})); \theta_j) - \epsilon \\ & + \sum_{j \neq i} v_j(g_\zeta(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})) \\ & \leq - \sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) \\ & + \sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(s_i, \sigma_{-i}(\theta_{-i})). \end{aligned}$$

Using (18) again, this can be simplified to

$$k_i(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})) - \epsilon \leq k_i(s_i, \sigma_{-i}(\theta_{-i})). \quad (19)$$

Since (18), by applying the argument in Case 1, it must follow that

$$k_i(\sigma_i(\theta''_i), \sigma_{-i}(\theta_{-i})) = k_i(s'_i, \sigma_{-i}(\theta_{-i})). \quad (20)$$

Substituting (20) into (19), we have  $k_i(s'_i, \sigma_{-i}(\theta_{-i})) - \epsilon \leq k_i(s_i, \sigma_{-i}(\theta_{-i}))$ . This is a contradiction to (15). Hence, we have shown that the tax rule must be in the form of (5) when  $s_i \in \sigma_i(\Theta_i)$ .

**Case 2:** Now we need to show that the inequality (6) must be satisfied for every  $s_i \in \mathcal{S}_i \setminus \sigma(\Theta_i)$ , where  $k_i$  is the same function as in Case 1. Suppose, on the contrary, that there exist  $s_i \in \mathcal{S}_i \setminus \sigma(\Theta_i)$  and  $\theta_{-i} \in \Theta_{-i}$  such that

$$g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})) + \epsilon < \sum_{j \neq i} v_j(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(\sigma_{-i}(\theta_{-i})) \quad (21)$$

with  $\epsilon > 0$ . By incentive compatibility, the  $i$ -th follower can minimize the net cost by acting  $\sigma_i(\theta_i)$ :

$$v_i(g_\zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) \leq v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_i) + g_{\pi_i}(s_i, \sigma_{-i}(\theta_{-i})).$$

By the discussion in Case 1, the equality  $g_{\pi_i}(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})) = \sum_{j \neq i} v_j(g_\zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \theta_j) + k_i(\sigma_{-i}(\theta_{-i}))$  is applicable on the left hand side. Also, by substituting (21) into the right hand side, we obtain

$$\sum_{i=1}^N v_i(g_\zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i})); \theta_i) < \sum_{i=1}^N v_i(g_\zeta(s_i, \sigma_{-i}(\theta_{-i})); \theta_i) - \epsilon.$$

Now, consider an extreme situation in which all cost functions are constant, i.e.,  $v_i(\cdot; \theta_i) = c$  for every  $i \in \{1, \dots, N\}$ . Then, the last inequality leads to  $\epsilon < 0$ , a contradiction.

## B. Proof of Proposition 2

To prove this proposition, assume that,  $\forall n \in \mathbb{N}$ , agent  $i$  follows  $s_i^n \in \mathcal{S}_i^n$  and the rest of the agents follow  $\sigma_{-i}^n(\theta_{-i})$ . Let us define sets

$$\begin{aligned} \mathcal{M} &= \{m \in \mathbb{N} \mid \text{dist}(g_\zeta^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})), \mathcal{Z}) > \beta(n_m)\}, \\ \mathcal{L} &= \{\ell \in \mathbb{N} \mid \text{dist}(g_\zeta^{n_\ell}(s_i^{n_\ell}, \sigma_{-i}^{n_\ell}(\theta_{-i})), \mathcal{Z}) \leq \beta(n_\ell)\}. \end{aligned}$$

Now, we prove the following two claims.

**Claim 1:** If  $\mathcal{M} \neq \emptyset, \forall \epsilon > 0, \exists \bar{m} \in \mathcal{M}$  such that

$$\begin{aligned} v_i(g_\zeta^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})) \\ \leq v_i(g_\zeta^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})) + \epsilon, \end{aligned}$$

for all  $m \in \mathcal{M}$  such that  $m \geq \bar{m}$ .

To prove this claim, note that,  $\forall m \in \mathcal{M}$ , we have

$$\begin{aligned} v_i(g_\zeta^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})) \\ = v_i(g_\zeta^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})); \theta_i) \\ - \inf_{\theta_i \in \Theta_i} \inf_{z \in \mathcal{Z}'} v_i(z; \theta_i) + \sup_{\theta \in \Theta} \sup_{z \in \mathcal{Z}'} \sum_{j=1}^N v_j(z; \theta_j) \\ \geq \sup_{\theta \in \Theta} \sup_{z \in \mathcal{Z}'} \sum_{j=1}^N v_j(z; \theta_j). \end{aligned} \quad (22)$$

Now, because  $g_{\zeta}^{n_m}(\sigma_i^{n_m}(\theta_i), \sigma_{-i}^{n_m}(\theta_{-i})) \in \mathcal{Z}'$ , we get

$$\sup_{\theta \in \Theta} \sup_{z \in \mathcal{Z}'} \sum_{j=1}^N v_j(z; \theta_j) \geq \sum_{j=1}^N v_j(g_{\zeta}^{n_m}(\sigma_i^{n_m}(\theta_i), \sigma_{-i}^{n_m}(\theta_{-i})); \theta_j),$$

which, in combination with (22), gives

$$\begin{aligned} & v_i(g_{\zeta}^{n_m}(\sigma_i^{n_m}(\theta_i), \sigma_{-i}^{n_m}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_m}(\sigma_i^{n_m}(\theta_i), \sigma_{-i}^{n_m}(\theta_{-i})) \\ & \leq v_i(g_{\zeta}^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_m}(s_i^{n_m}, \sigma_{-i}^{n_m}(\theta_{-i})). \end{aligned}$$

This proves Claim 1 by setting  $\bar{m} = \min_{m \in \mathcal{M}} m$  (which is well-defined as  $\mathcal{M} \neq \emptyset$ ).

**Claim 2:** If  $\mathcal{L} \neq \emptyset$  and  $|\mathcal{L}| = \infty$ ,  $\forall \epsilon > 0$ ,  $\exists \bar{\ell} \in \mathcal{L}$  such that

$$\begin{aligned} & v_i(g_{\zeta}^{n_{\ell}}(\sigma_i^{n_{\ell}}(\theta_i), \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_{\ell}}(\sigma_i^{n_{\ell}}(\theta_i), \sigma_{-i}^{n_{\ell}}(\theta_{-i})) \\ & \leq v_i(g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) + \epsilon, \end{aligned}$$

for all  $\ell \in \mathcal{L}$  such that  $\ell \geq \bar{\ell}$ .

Let us define

$$\hat{z}^{n_{\ell}} \in \arg \min_{z \in \mathcal{Z}} \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z\|_2, \quad \forall \ell \in \mathcal{L}.$$

Let us show  $\arg \min_{z \in \mathcal{Z}} \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z\|_2 \neq \emptyset$ . Fix an arbitrary  $z_0 \in \mathcal{Z}$  and define  $\mathcal{B} = \{z \in \mathbb{R}^{n_z} \mid \|z - z_0\|_2 \leq 2\|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z_0\|_2\}$ . For all  $z \in \mathcal{Z} \setminus \mathcal{B}$ , we have

$$\begin{aligned} \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z\|_2 & \geq \|z - z_0\|_2 \\ & \quad - \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z_0\|_2 \\ & > \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z_0\|_2, \end{aligned}$$

where the first and the second inequalities, respectively, follow from the triangular inequality and that  $z \in \mathcal{Z} \setminus \mathcal{B}$ . Hence,

$$\begin{aligned} & \inf_{z \in \mathcal{Z}} \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z\|_2 \\ & = \inf_{z \in \mathcal{Z} \cap \mathcal{B}} \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z\|_2, \end{aligned} \quad (23)$$

because, as we showed above, any  $z \notin \mathcal{Z} \cap \mathcal{B}$  results in a strictly larger distance than  $z_0 \in \mathcal{Z} \cap \mathcal{B}$ . Note that  $\mathcal{Z} \cap \mathcal{B}$  is compact because it is bounded (subset of bounded set  $\mathcal{B}$ ) and closed (intersection of two closed sets). Theorem 4.16 in [44, p. 89] shows  $\exists z' \in \mathcal{Z} \cap \mathcal{B}$  that achieves the infimum on the right hand side of (23) and, thus, the infimum on the left hand side of (23). Therefore,  $z' \in \arg \min_{z \in \mathcal{Z}} \|g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) - z\|_2$  because  $z' \in \mathcal{Z}$ . Therefore,  $\hat{z}^{n_{\ell}}$  is well-defined. For any  $\epsilon > 0$ , there exists  $\ell_1 \in \mathcal{L}$  such that for all  $\ell \in \mathcal{L}$  that  $\ell \geq \ell_1$ , we get

$$\left| \sum_{j=1}^N v_j(g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_j) - \sum_{j=1}^N v_j(\hat{z}^{n_{\ell}}; \theta_j) \right| \leq \epsilon/2,$$

because of the continuity of  $v_i(\cdot; \theta_i)$ ,  $\forall i$ , and the fact that  $\lim_{\ell \in \mathcal{L}, \ell \rightarrow \infty} \text{dist}(g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})), \mathcal{Z}) = 0$  since  $\{\beta(n_{\ell})\}_{\ell \in \mathcal{L}}$  is a vanishing sequence (as  $|\mathcal{L}| = \infty$ ). Hence,

$$\begin{aligned} \sum_{j=1}^N v_j(\hat{z}^{n_{\ell}}; \theta_j) & \leq \sum_{j=1}^N v_j(g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_j) + \epsilon/2 \\ & = v_i(g_{\zeta}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_i) \\ & \quad + g_{\pi_i}^{n_{\ell}}(s_i^{n_{\ell}}, \sigma_{-i}^{n_{\ell}}(\theta_{-i})) + \epsilon/2. \end{aligned} \quad (24)$$

Note that, by construction,  $\hat{z}^{n_{\ell}} \in \mathcal{Z}$ ,  $\forall \ell \in \mathcal{L}$ . Hence, for any  $\epsilon > 0$ , there exists  $\ell_2 \in \mathcal{L}$  such that for all  $\ell \in \mathcal{L}$  that  $\ell \geq \ell_2$ , we get

$$\begin{aligned} & v_i(g_{\zeta}^{n_{\ell}}(\sigma_i^{n_{\ell}}(\theta_i), \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_i) + g_{\pi_i}^{n_{\ell}}(\sigma_i^{n_{\ell}}(\theta_i), \sigma_{-i}^{n_{\ell}}(\theta_{-i})) \\ & = \sum_{j=1}^N v_j(g_{\zeta}^{n_{\ell}}(\sigma_i^{n_{\ell}}(\theta_i), \sigma_{-i}^{n_{\ell}}(\theta_{-i})); \theta_j) \\ & \leq \sum_{j=1}^N v_j(\hat{z}^{n_{\ell}}; \theta_j) + \epsilon/2, \end{aligned} \quad (25)$$

where the last inequality is because  $\{g_{\zeta}^{n_{\ell}} \circ \sigma^{n_{\ell}}\}_{\ell \in \mathcal{L}}$  is asymptotically efficient as  $|\mathcal{L}| = \infty$  (note that, by definition, if a sequence of decision rules is asymptotically efficient, every infinite subsequence of it is also asymptotically efficient). Combining (24) and (25) while setting  $\bar{\ell} = \max(\ell_1, \ell_2)$  proves Claim 2.

Now, we are ready to prove the statement of this proposition. If  $|\mathcal{L}| < \infty$  (which implies that  $|\mathcal{M}| = \infty$  as, by definition,  $\mathcal{M} \cup \mathcal{L} = \mathbb{N}$  and  $\mathcal{M} \cap \mathcal{L} = \emptyset$ ), the proposition follows from Claim 1 and by setting  $n_0 = \max\{n_{\max_{\ell \in \mathcal{L}} \ell}, n_{\bar{m}}\}$  (in Definition 5). Otherwise, if  $|\mathcal{L}| = \infty$ , the proof follows from Claims 1 and 2 and by setting  $n_0 = \max\{n_{\bar{\ell}}, n_{\bar{m}}\}$ .

### C. Proof of Proposition 3

First, note that the payments introduced in Algorithm 1 is of the form introduced in Proposition 1 when setting  $k_i(s_{-i}^n) = -\sum_{j \neq i} \hat{v}_j^{-i,n}$  for all  $1 \leq i \leq N$ . Note that this quantity does not depend on follower  $i$ 's actions. Second, for each  $n \in \mathbb{N}$ , the mechanism  $M^n$  is single fault tolerant, since the outcome of the mechanism is guaranteed to be feasible (i.e.,  $\zeta^n(\theta) \in \mathcal{Z}$ ) due to the operation in line 17 of Algorithm 1. Third, by the convergence property of the dual decomposition algorithm and the continuity of cost functions, we have

$$\lim_{n \rightarrow \infty} \sum_{i=1}^N v_j(\zeta_i^n(\theta); \theta_i) = \min_{Rz=c} \sum_{i=1}^N v_i(z_i; \theta_i).$$

Hence, the sequence of decision rules  $\{\zeta^n\}_{n \in \mathbb{N}}$  is asymptotically efficient. Now, the rest follows from Proposition 1.

### D. Proof of Proposition 5

Note that  $\mathcal{Z} = \{z \mid B^{\top} z = 0\}$ . For any  $\hat{z} \in \mathbb{R}^N$ ,  $\text{dist}(\hat{z}, \mathcal{Z}) = \|B(B^{\top} B)^{-1} B^{\top} \hat{z}\|_2$  as  $(I - B(B^{\top} B)^{-1} B^{\top}) \hat{z}$  is the projection of  $\hat{z}$  into  $\mathcal{Z}$ . Following [45, p. 74], we get

$$\begin{aligned} \lambda_{\max}(I - \alpha B^{\top} B) & \leq \lambda_{\max}(I) + \lambda_{\max}(-\alpha B^{\top} B) \\ & = \lambda_{\max}(I) - \alpha \lambda_{\min}(B^{\top} B) \\ & \leq 1 - \rho, \end{aligned}$$

and

$$\begin{aligned} \lambda_{\min}(I - \alpha B^{\top} B) & \geq \lambda_{\min}(I) + \lambda_{\min}(-\alpha B^{\top} B) \\ & = \lambda_{\min}(I) - \alpha \lambda_{\max}(B^{\top} B) \\ & \geq 0. \end{aligned}$$

Now, note that

$$B^{\top} z^{k+1} = B^{\top} (I - \alpha B B^{\top}) z^k = (I - \alpha B^{\top} B) B^{\top} z^k,$$

which results in

$$\begin{aligned} \|B^\top z^{k+1}\|_2^2 &\leq \|I - \alpha B^\top B\|_2^2 \|B^\top z^k\|_2^2 \\ &\leq \lambda_{\max}(I - \alpha B^\top B)^2 \|B^\top z^k\|_2^2 \\ &\leq (1 - \rho)^2 \|Rz^k\|_2^2, \end{aligned}$$

where the second inequality follows from [45, p. 133]. This shows that

$$\|B^\top z^k\|_2^2 \leq (1 - \rho)^{2k} \|B^\top z^0\|_2^2 \leq (1 - \rho)^{2k} \sup_{q \in \Theta} \|B^\top q\|_2^2.$$

Define  $\beta(n) = (1 - \rho)^n \|B(B^\top B)^{-1}\|_2 \sup_{q \in \Theta} \|B^\top q\|_2$ . Evidently,  $\lim_{n \rightarrow \infty} \beta(n) = 0$  since  $0 < \rho \leq 1$  (as  $\lambda_{\min}(B^\top B) > 0$  because  $\mathcal{G}$  is a tree). Furthermore, since  $\mathbf{1}^\top z^k = \mathbf{1}^\top z^0$  following the update dynamics in Algorithm 2 (see [46]), we get  $\lim_{k \rightarrow \infty} z^k = \left(\frac{1}{N} \sum_{i=1}^N \theta_i\right) \mathbf{1}$ . Therefore,  $\{g_\zeta^n \circ \sigma^n\}_{n \in \mathbb{N}}$  is asymptotically efficient. Now, the rest of the proof follows from applying Proposition 2.

## REFERENCES

- [1] T. Tanaka, F. Farokhi, and C. Langbort, "A faithful distributed implementation of dual decomposition and average consensus algorithms," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013.
- [2] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems (TOPLAS)*, vol. 4, no. 3, pp. 382–401, 1982.
- [3] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford University Press, 1995.
- [4] R. Myerson, *GAME THEORY*. Harvard University Press, 1997.
- [5] M. O. Jackson, "Mechanism theory," in *Optimization and Operations Research* (U. Derigs, ed.), Encyclopedia of Life Support Systems, Oxford, UK: EOLSS Publishers, 2003.
- [6] M. O. Jackson, "A crash course in implementation theory," *Social Choice and Welfare*, vol. 18, no. 4, pp. 655–708, 2001.
- [7] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. New York, NY, USA: Cambridge University Press, 2008.
- [8] V. Krishna, *Auction Theory*. Elsevier Science, 2009.
- [9] N. Nisan and A. Ronen, "Computationally feasible VCG mechanisms," in *Proceedings of the 2nd ACM Conference on Electronic Commerce*, pp. 242–252, 2000.
- [10] D. Lehmann, L. I. O’Callaghan, and Y. Shoham, "Truth revelation in approximately efficient combinatorial auctions," *Journal of the ACM*, vol. 49, pp. 577–602, sep 2002.
- [11] A. Archer and É. Tardos, "Truthful mechanisms for one-parameter agents," in *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pp. 482–491, IEEE, 2001.
- [12] J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker, "A BGP-based mechanism for lowest-cost routing," *Distributed Computing*, vol. 18, no. 1, pp. 61–72, 2005.
- [13] G. Kotsalis and J. S. Shamma, "Robust synthesis in mechanism design," in *Proceedings of the 49th IEEE Conference on Decision and Control*, pp. 225–230, dec. 2010.
- [14] N. Nisan and A. Ronen, "Algorithmic mechanism design," in *Proceedings of the 31st Annual ACM Symposium on Theory of computing*, pp. 129–140, 1999.
- [15] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge University Press, 2007.
- [16] J. R. Marden and A. Wierman, "Distributed welfare games with applications to sensor coverage," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pp. 1708–1713, IEEE, 2008.
- [17] T. Tanaka, A. Z. W. Cheng, and C. Langbort, "A dynamic pivot mechanism with application to real time pricing in power systems," in *Proceedings of the American Control Conference*, pp. 3705–3711, 2012.
- [18] N. Li and J. R. Marden, "Designing games for distributed optimization," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 2, pp. 230–242, 2013.
- [19] D. Monderer and M. Tennenholtz, "Distributed games: From mechanisms to protocols," in *AAAI/IAAI*, pp. 32–37, 1999.
- [20] J. Feigenbaum, C. H. Papadimitriou, and S. Shenker, "Sharing the cost of multicast transmissions," *Journal of Computer and System Sciences*, vol. 63, pp. 21–41, 2001.
- [21] J. Feigenbaum and S. Shenker, "Distributed algorithmic mechanism design: recent results and future directions," in *DIAL-M*, pp. 1–13, 2002.
- [22] D. C. Parkes and J. Shneidman, "Distributed implementations of Vickrey–Clarke–Groves mechanisms," in *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multi Agent Systems*, pp. 261–268, 2004.
- [23] A. Petcu, B. Faltings, and D. C. Parkes, "MDPOP: Faithful distributed implementation of efficient social choice problems," in *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 1397–1404, 2006.
- [24] R. Cavallo, D. C. Parkes, and S. Singh, "Optimal coordinated planning amongst self-interested agents with private state," *arXiv preprint arXiv:1206.6820*, 2012.
- [25] J. Green and J.-J. Laffont, "Characterization of satisfactory mechanisms for the revelation of preferences for public goods," *Econometrica: Journal of the Econometric Society*, pp. 427–438, 1977.
- [26] C. Papadimitriou, *Computational Complexity*. Addison-Wesley, 1995.
- [27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [28] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," *Lecture notes of EE3920, Stanford University, Autumn, 2003*. [http://www.stanford.edu/class/ee364b/lectures/subgrad\\_method\\_notes.pdf](http://www.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf).
- [29] H. R. Varian, *Intermediate Microeconomics: A Modern Approach*. W. W. Norton & Company, 2010.
- [30] S. Hao, "A study of basic bidding strategy in clearing pricing auctions," in *Power Industry Computer Applications, 1999. PICA'99. Proceedings of the 21st 1999 IEEE International Conference*, pp. 55–60, IEEE, 1999.
- [31] A. R. Kian, J. B. Cruz, and R. J. Thomas, "Bidding strategies in oligopolistic dynamic electricity double-sided auctions," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 50–58, 2005.
- [32] R. Johari and J. N. Tsitsiklis, "Efficiency loss in a network resource allocation game," *Mathematics of Operations Research*, vol. 29, no. 3, pp. 407–435, 2004.
- [33] R. B. Myerson and M. A. Satterthwaite, "Efficient mechanisms for bilateral trading," *Journal of economic theory*, vol. 29, no. 2, pp. 265–281, 1983.
- [34] V. Krishna and M. Perry, "Efficient mechanism design," (*Technical Report*) Pennsylvania State University, 1998.
- [35] J. R. Marden and A. Wierman, "Overcoming limitations of game-theoretic distributed control," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pp. 6466–6471, IEEE, 2009.
- [36] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [37] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995.
- [38] W. Ren, R. W. Beard, and E. M. Atkins, "A survey of consensus problems in multi-agent coordination," in *Proceedings of the American Control Conference*, pp. 1859–1864, 2005.
- [39] M. G. Rabbat, R. D. Nowak, and J. A. Bucklew, "Generalized consensus computation in networked systems with erasure links," in *Proceedings of the 6th IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 1088–1092, 2005.
- [40] A. Rantzer, "Dynamic dual decomposition for distributed control," in *Proceedings of the American Control Conference*, pp. 884–888, 2009.
- [41] E. Camponogara, D. Jia, B. H. Krogh, and S. Talukdar, "Distributed model predictive control," *IEEE Control Systems Magazine*, vol. 22, no. 1, pp. 44–52, 2002.
- [42] A. N. Venkat, I. A. Hiskens, J. B. Rawlings, and S. J. Wright, "Distributed mpc strategies with application to power system automatic generation control," *Control Systems Technology, IEEE Transactions on*, vol. 16, no. 6, pp. 1192–1206, 2008.
- [43] R. R. Negenborn, B. De Schutter, and J. Hellendoorn, "Multi-agent model predictive control for transportation networks: Serial versus parallel schemes," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 3, pp. 353–366, 2008.
- [44] W. Rudin, *Principles of Mathematical Analysis*. International Series in Pure and Applied Mathematics, McGraw-Hill, 3 ed., 1976.
- [45] H. Lütkepohl, *Handbook of Matrices*. Wiley, 1996.
- [46] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.