A Coalitional Game Theoretic Outlook on Distributed Adaptive Parameter Estimation

Nikola Bogdanović, *Member, IEEE,* and Dimitris Ampeliotis, *Member, IEEE*
and Kostas Berberidis, *Senior Member, IEEE*

***Abstract*—In this paper, the parameter estimation problem based on diffusion least mean squares strategies is analyzed from a coalitional game theoretical perspective. Specifically, while selfishly minimizing only their own mean-square costs, the nodes in a network form coalitions that benefit them. Due to its nature, the problem is modeled as a non-transferable game and two scenarios are studied, one where each node's payoff includes only a suitable estimation accuracy criterion and another one in which a graph-based communication cost is also considered. In the former scenario, we first analyze the non-emptiness of the core of the games corresponding to traditional diffusion strategies, and then, the analysis is extended to a recently proposed node-specific parameter estimation setting where the nodes have overlapped but different estimation interests. In the latter scenario, after formulating a coalitional graph game and providing sufficient conditions for its core non-emptiness, we propose a distributed graph formation algorithm, based on merge-and-split approach, which converges to a stable coalition structure.**

***Index Terms*—Adaptive distributed networks, diffusion, LMS, cooperation, node-specific parameter estimation, coalitional game theory, NTU game, core, graph game.**

I. Introduction

Several distributed least mean squares (LMS) algorithms for adaptive parameter estimation over networks were developed in the literature, namely, the consensus, the incremental and the diffusion strategy (see [2]-[7] and references therein). Until recently, the mainstream of existing research in adaptive networks focused on considering that the nodes' interests are absolutely identical, which we refer to as the traditional setting. Presently, research efforts have focused on removing this restriction. For instance, the Node-Specific Parameter Estimation (NSPE) formulation has been introduced and analyzed in [8]-[11], where the nodes are considered to have overlapped but generally different estimation interests. Another useful approach was provided in [12], in which the proposed algorithm deals with the scenario where nodes have numerically similar estimation interests. Also, for the same scenario as in [12], the performance of the diffusion strategy, designed for the traditional setting, has been analyzed in [13]. The main advantage of these approaches with respect to the non-cooperative ones is in an improved estimation accuracy due to a proper cooperation mechanism among nodes. However, the cooperation benefits are usually analyzed from a network-wide perspective. In other words, cooperation is shown to improve the estimation performance metrics averaged over the whole network. Therefore, it remains unclear whether each node in the network benefits from cooperation, or whether the nodes could be organized in certain groups/subnetworks so as to benefit more. In this work, we aim to answer these questions for the NSPE scenario by allowing nodes or groups of nodes to be selfish using concepts from coalitional game theory.

Dimitris Ampeliotis and Kostas Berberidis are with the Computer Engineering and Informatics Department and CTI/RU8, University of Patras, Patras 26500 GREECE e-mail: {ampeliot,berberid}@ceid.upatras.gr. Nikola Bogdanović, as corresponding author, was with the Computer Engineering and Informatics Department and CTI/RU8, University of Patras, Patras 26500 GREECE e-mail: bogdanovic@ceid.upatras.gr. He is now with the Microwave Sensing, Signals and Systems (M3S) group, TU Delft, The Netherlands.

A short and preliminary version of this work appeared in the conference proceedings [1].

The work was partially supported by the European HANDiCAMS project (Grant No. 323944) under the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, and in part by the University of Patras.

A. Game Theory in brief

In general, game theory can be defined as the study of mathematical models related to rational decision-makers in situations involving conflicts of interest and cooperation. Over the last decades, there has been a dramatic growth in both the number of theoretical results and the variety of applications in disciplines such as economics, political sciences, philosophy and more recently, engineering [14],[15]. In contrast to non-cooperative game theory, where the modeling unit is a single player, coalitional game theory, being the focus of this paper, seeks for an optimal coalition structure of players so as to maximize the worth of each coalition. In other words, a coalitional game is differentiated from a non-cooperative game primarily by its focus on what groups of players can achieve together rather than on what individual players can do alone [16]. In accordance with [17], coalitional games can be classified into the following three categories, i.e., canonical coalitional games, coalition formation games and coalitional graph games. The main task in a canonical game is to study stability of the coalition of all players in the game, while in a coalition formation game, one usually analyzes the formation of a stable coalitional structure by also taking cooperation cost into account. In a coalitional graph game, the value of a coalition depends both on the members of the coalition and on the interconnections between them. Also, coalitional games are divided into Transferable Utility (TU) games and Non-Transferable Utility (NTU) games. In the former ones, each group of players is associated with a single number, interpreted as the payoff that is given to the group and that may be distributed in any way among the group members. The

latter represent a more general concept, and NTU games are sometimes referred to as games without side payments [18]. In these games, each coalition cannot necessarily achieve all distributions of some fixed payoff, so that the value of a coalition is a set of payoff vectors used to assign each member of the coalition its own utility [19].

B. Game Theory for adaptive networks

First of all, we should highlight that all game theoretical references addressing the problem of adaptive networks consider the traditional setting only, in which all nodes aim to estimate a set of parameters that are identical to all nodes. Also, most of these references focus on non-cooperative games, where the main solution concept is the celebrated Nash equilibrium (NE). Due to the lack of coordination among the players, the NE can be rather inefficient, so the main question is how to improve the players' payoffs. A way to accomplish this can be through a reputation design strategy, which was employed in [20] to enforce cooperation between rationally selfish nodes in a pairwise one-shot successive game scenario. For the same scenario, an approximative cluster formation protocol was provided in [21], in which the nodes pairwisely choose whether to merge based on the estimation accuracy gain and the inherent communication cost. Another approach to improve the players' payoffs with respect to the ones guaranteed by the Nash equilibrium is correlating the players' choices through some signaling mechanism. This approach has been considered in [22] to tackle the problem of node activation control. Next, in [23], the distributed adaptive filtering problem was addressed using a game-theoretical model inspired by evolutionary biology, i.e., a graphical evolutionary game. More precisely, it was proved that the strategy of using information from nodes with good signal is always an evolutionarily stable strategy.

C. Paper contributions and organization

In this paper, we analyze the distributed adaptive parameter estimation problem using the tools from coalitional game theory. The problem is modeled as a coalitional game with non-transferable utility for which the choice of coalitional actions defines the payoff attainable to each player. The main contributions of this article are listed in the following: 1) the traditional diffusion setting has been analyzed from a coalition game-theoretic perspective, 2) the analysis has been extended to the NSPE setting game, and 3) the NSPE setting has been analyzed as a coalitional graph game.

The structure of the paper is as follows. Section II starts with a brief presentation of the NSPE formulation and the diffusion strategies for both settings, the traditional one and the more general, NSPE setting; then, the distributed adaptive parameter estimation problem with self-interested agents, being in the focus of this paper, is motivated. In Section III, we analyze the aforementioned problem for a scenario where the nodes' payoffs are modeled exclusively using a suitable estimation accuracy criterion. The stability of the coalition of all nodes is studied in games modeling both the traditional and the NSPE setting. Next, Section IV is devoted to the graph games so as to model the problem in question. The nodes' payoffs are now defined as graph-based functions that also take the cooperation cost into account. In Section V, the main findings are verified via computer simulations. Finally, Section VI summarizes the work.

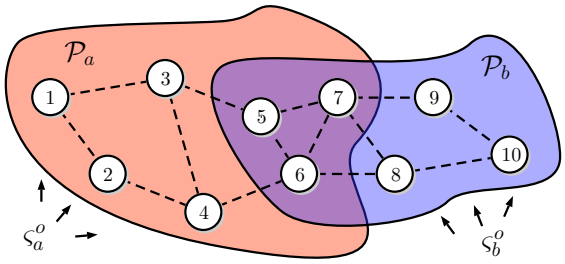


Fig. 1. An illustrative example of an NSPE network with two overlapping phenomena, denoted by ς_a^o and ς_b^o, and with their influence areas given by $\mathcal{P}_a = \{1, \ldots, 7\}$ and $\mathcal{P}_b = \{5, \ldots, 10\}$.

D. Notation

The following notation is used throughout the paper. We use boldface letters for random variables and normal fonts for deterministic quantities. We use the weighted (semi)norm notation $\|x\|_{\Sigma}^2 \triangleq x^H \Sigma x$ with a vector x and a Hermitian positive (semi-)definite matrix Σ. Furthermore, $R_a = E\{\mathbf{a}^H \mathbf{a}\}$ for a random vector $\mathbf{a}$. The notation $\mathrm{diag}\{\cdot\}$ denotes a block-diagonal matrix with arguments on block-diagonal. Additionally, the notation $\mathrm{col}\{\cdot\}$ denotes a column operator stacking arguments on top of each other. Note that 1_L stands for an $L \times 1$ vector of ones. Also, let $\mathbb{R}$ denote the set of all real numbers. $\mathbb{R}^{|\mathcal{N}|}$ is the $|\mathcal{N}|$-dimensional Euclidean space generated by a finite set of players $\mathcal{N}$. An element of $\mathbb{R}^{|\mathcal{N}|}$ is denoted by a vector $\mathrm{x} = (x_k)_{k \in \mathcal{N}}$. For a coalition $S \subseteq \mathcal{N}$, let $\mathrm{x}^S = (x_k)_{k \in S}$ denote the restriction of x on S. For $\mathrm{x}, \mathrm{y} \in \mathbb{R}^{|\mathcal{N}|}$, $\mathrm{y}^S \geq \mathrm{x}^S$ denotes $y_k \geq x_k$ for all $k \in S$ with at least one element satisfying the strict inequality. The symbol $\times$ denotes the Cartesian product.

II. Node-specific parameter estimation- overview

In this section, we initially summarize the NSPE formulation proposed in [8]-[11], and then, the diffusion adaptive strategy in both the NSPE and the traditional setting is provided. Finally, the main questions considered in this paper are posed.

Assume a network of N randomly deployed nodes affected by some phenomena in a different way. For instance, Figure 1 shows two such phenomena, being modeled with unknown deterministic vectors ς_a^o and ς_b^o. Nodes that are influenced by each phenomenon are assumed inter-connected via some topology. The neighborhood of any particular node k, together with node k, is denoted as $\mathcal{N}_k$.

Each node k, at discrete time i, has access to data $\{d_k(i), u_{k,i}\}$ which are time realizations of zero-mean random processes $\mathbf{d}_k(i) \in \mathbb{C}$ and $\mathbf{u}_{k,i} \in \mathbb{C}^{1 \times M_k}$, which are related via the linear model

$$\mathbf{d}_k(i) = \mathbf{u}_{k,i} w_k^o + \mathbf{v}_k(i). \tag{1}$$

where

- w_k^o is the deterministic vector of size M_k that models all phenomena node k is influenced by, and
- the term $\mathbf{v}_k(i) \in \mathbb{C}$ denotes a zero-mean white noise, of variance $\sigma_{v,k}^2$, that is independent of all other variables.

In [8]-[11], nodes willingly cooperate to estimate in a distributed manner the parameter vectors $\{w_k^o\}_{k=1}^N$ that minimize the network-wide mean-square error cost, namely,

$$\underset{\{w_k\}_{k=1}^N}{\text{minimize}} \sum_{k=1}^{N} J_k(w_k) \triangleq \sum_{k=1}^{N} E\left\{|\mathbf{d}_k(i) - \mathbf{u}_{k,i} w_k|^2\right\}. \quad (2)$$

Note that vectors $\{w_k^o\}_{k=1}^N$, i.e., the minimizers of individual costs $\{J_k(w_k)\}_{k=1}^N$, are different in general, yet overlapping. As an example, consider a simple, NSPE setting where there are only two, partially overlapping events (ς_a^o and ς_b^o) in the network, as illustrated in Fig. 1. In this setting, there are some nodes influenced by one of the two events as well as other nodes affected by both events, i.e., by their superposition. The subset of nodes affected by ς_a^o is denoted by $\mathcal{P}_a$ while $\mathcal{P}_b$ stands for the subset of nodes interested in estimating ς_b^o.

For a more general NSPE setting, where there are J phenomena $\varsigma_1^o \dots \varsigma_J^o$, the observation model from (1), at each node k, can be rewritten as

$$\mathbf{d}_k(i) = \sum_{j \in \mathcal{I}_k} \mathbf{u}_{k,i}^{(j)} \varsigma_j^o + \mathbf{v}_{k,i} \quad (3)$$

where

- $\mathbf{u}_{k,i}^{(j)}$ is a row vector, of size $1 \times M_{\varsigma,j}$, corresponding to ς_j^o, and
- $\mathcal{I}_k \subseteq \{1, \dots, J\}$ represents a finite set of indices j related to the vectors ς_j^o that are affecting node k. For instance, in Fig. 1, the set of indices related to the phenomena node 6 is influenced by is $\mathcal{I}_6 = \{a, b\}$.

Now, let $\phi_{k,\varsigma_j}^{(i-1)}$ represent the estimate of ς_j^o at node k and time instant $i-1$, and let us also define the quantity that stacks the estimates with $\phi_{k,\varsigma}^{(i-1)} = \text{col}\Big\{\{\phi_{k,\varsigma_j}^{(i-1)}\}_{j \in \mathcal{I}_k}\Big\}$. For a suitably chosen step size μ_k, a distributed strategy based on diffusion adaptation for the considered NSPE scenario involves two steps [11], i.e., adapting the estimate $\phi_{k,\varsigma}^{(i-1)}$ based on new data

$$\varsigma_k^{(i)} = \phi_{k,\varsigma}^{(i-1)} + \mu_k \, u_{k,i}^H \left[d_k(i) - \sum_{j \in \mathcal{I}_k} u_{k,i}^{(j)} \phi_{k,\varsigma_j}^{(i-1)} \right], \quad (4)$$

followed by the combination step, for each $j \in \mathcal{I}_k$, namely,

$$\phi_{k,\varsigma_j}^{(i)} = \sum_{\ell \in \mathcal{N}_k \cap \mathcal{P}_j} c_{k,\ell}^{\varsigma_j} \, \varsigma_{\ell,j}^{(i)}. \quad (5)$$

In relations above, $\varsigma_{k,j}^{(i)}$ stands for the estimate of ς_j^o at node k and time instant i after the adaptation while $\varsigma_k^{(i)} = \text{col}\Big\{\{\varsigma_{k,j}^{(i)}\}_{j \in \mathcal{I}_k}\Big\}$. The non-negative coefficient $c_{k,\ell}^{\varsigma_j}$ is the (k,ℓ)-th entry of a $|\mathcal{P}_j| \times |\mathcal{P}_j|$ row-stochastic combination matrix C^{ς_j}, satisfying $\sum_{\ell \in \mathcal{N}_k \cap \mathcal{P}_j} c_{k,\ell}^{\varsigma_j} = 1$.

To assess the mean-square estimation performance of an adaptive algorithm, one may use the mean-square deviation (MSD) criterion evaluated in the steady-state, defined below

$$\text{MSD}_k^{(\infty)} = \lim_{i \to \infty} E\|w_k^o - \boldsymbol{\phi}_{k,\varsigma}^{(i)}\|^2. \quad (6)$$

In a more general case, in addition to the estimation performance criterion, it can be considered that communication between the nodes has non-negligible cost, so that each node k aims to minimize

$$\text{MSD}_k^{(\infty)} + \eta_k, \quad (7)$$

where η_k is some normalized communication cost incurred by node k. This scenario will be analyzed in Section IV, while in the following section we will focus on the setting where the cost is negligible. To that aim, we will first analyze traditional diffusion [4]-[7]), which can be seen a special of the NSPE setting. Specifically, in the traditional diffusion setting $w_k^o = w^o$ for all $k \in \{1, 2, \dots, N\}$ in the observation model in (1). In the context of the scenario depicted in Fig. 1, this could happen if both events affected all the nodes, i.e., $|\mathcal{P}_a| = |\mathcal{P}_b| = N$ and $w^o = \text{col}\Big\{\{\varsigma_j^o\}_{j \in \{a,b\}}\Big\}$. Then, relations (4) and (5) reduce to the traditional diffusion strategy:

$$\psi_k^{(i)} = \phi_k^{(i-1)} + \mu_k \, u_{k,i}^H \left[d_k(i) - u_{k,i} \phi_k^{(i-1)} \right] \quad (8)$$

$$\phi_k^{(i)} = \sum_{\ell \in \mathcal{N}_k} c_{k,\ell} \, \psi_\ell^{(i)} \quad (9)$$

where $\psi_k^{(i)}$ and $\phi_k^{(i)}$ are the estimates of w^o at node k at time i after the adaptation and the combination step, respectively, while the non-negative coefficient $c_{k,\ell}$ is the (k,ℓ)-th entry of an $N \times N$ row-stochastic combination matrix C. It should be highlighted that the combination rule has a strong impact on the performance of a diffusion algorithm (in both the NSPE and the traditional setting). Either fixed or adapted over time, there are several combination rules that can be chosen. For instance, the Hastings rule is given as

$$c_{k,l} = \begin{cases} \frac{\sigma_{v,k}^2}{\max\{n_k \cdot \sigma_{v,k}^2, n_l \cdot \sigma_{v,l}^2\}} & l \in \mathcal{N}_k \setminus \{k\} \\ 1 - \sum_{m \in \mathcal{N}_k \setminus \{k\}} c_{k,m} & l = k \end{cases} \quad (10)$$

where n_k is the number of neighbors of node k including the node k itself, i.e., $n_k = |\mathcal{N}_k|$. In contrast to (10), there are rules that satisfy $C^T 1_N = 1_N$ (e.g., Metropolis rule, Laplacian rule etc.), and they are called doubly-stochastic rules.

Now, once the traditional and the NSPE diffusion settings have been reviewed, the main contributions of our work are listed and explained in more detail:

- *The traditional diffusion setting is analyzed from a coalition game-theoretic perspective.* Although some useful discussion related to the cooperation benefits in a traditional diffusion setting appears in the literature [7], as well as some rather relevant works based on non-cooperative game theory [20]-[21], this work (together with its earlier version in [1]) is the first to analyze the problem using the solution concepts from coalitional game theory. Note that this stands in sharp contrast to the aforementioned literature where only single nodes, and not groups of nodes, are allowed to be rationally selfish.

- *The analysis is generalized to the NSPE setting game.* A natural question is whether the coalitional stability of a single-interest network (traditional diffusion) is related to the coalitional stability in the NSPE setting. Under certain conditions, we show that the grand coalition with respect to the whole NSPE network is coalitionally stable iff all grand coalitions with respect to each interest ς_j^o considered separately (as in a traditional diffusion setting) are stable.
- *The NSPE setting is analyzed as a coalitional graph game.* In contrast to our preliminary work [1], where communication topology is a fully connected graph, in Section IV we allow for more general graph structures to model communication costs among the nodes in the network. To that aim, we redefine a solution concept with respect to the NSPE grand coalition, and based on it, we provide sufficient conditions satisfying it. Finally, for a general communication cost setting, we propose a coalition graph formation game.

III. Canonical games for parameter estimation

In this section, the distributed adaptive parameter estimation problem is analyzed for a scenario where the players' payoffs are modeled exclusively using a suitable estimation accuracy criterion, namely, MSD. The nodes of the network represent the players in the game and we define several games for both settings, the traditional one and the NSPE setting. The main focus is to study whether and under which conditions the selfish players will form the grand coalition, i.e., the case where all the nodes of the network cooperate via diffusion strategy through a connected network topology. In all these games, we employ the so-called core [18], [19], a classical solution concept of coalitional game theory. Initially, we argue under which conditions the grand coalition in the traditional setting is stable. Finally, we generalize these findings for the NSPE case by properly decomposing the underlying game.

A. Data assumptions

In order to convey the main ideas of this work, certain assumptions on the data in (1) and (3) are adopted. Specifically, for the traditional setting, the following, common to relevant literature [7], assumptions are made,

A1) the regressors $\mathbf{u}_{k,i} \in \mathbb{C}^{1\times M}$ are temporally white and spatially independent with the autocovariance $R_{u_k} = R_u > 0$ being equal for all nodes,

A2) $\mathbf{v}_k(i) \in \mathbb{C}$ is temporally and spatially white noise, of variance $\sigma^2_{v_k}$, which is independent of $\mathbf{u}_{\ell,j}$ for all k, i, ℓ, j,

A3) the step sizes $\mu_k = \mu$, for all nodes, and are sufficiently small so that the higher-order terms of μ can be ignored.

In case of the NSPE scenario, we further assume that

A4) the regressors $\mathbf{u}_{k,i}^{(j)} \in \mathbb{C}^{1\times M_{\varsigma,j}}$ and $\mathbf{u}_{k,i}^{(j')} \in \mathbb{C}^{1\times M_{\varsigma,j'}}$ are independent for any $j, j' \in \{1,\ldots,J\}$ and $j \neq j'$.

B. MSD game definition

Let us define an NTU coalitional game $(\mathcal{N}, \upsilon)$, where

- $\mathcal{N} = \{1,\ldots,N\}$ is the set of players (nodes) while
- $\upsilon(S) \subseteq \mathbb{R}^N$ is the set of payoff vectors[1] of a coalition $S \subseteq \mathcal{N}$, which are related to the estimation accuracy that the players in a coalition S may achieve.

In particular, if a payoff $x_k(S)$ denotes the k^{th} element of payoff vector $\mathrm{x} \in \upsilon(S)$, then a payoff $x_k(S)$ that represents the maximum estimation accuracy a certain node $k \in S$ can obtain is given by

$$x_k(S) = -\mathrm{MSD}_k(S) = -\lim_{i\to\infty} E\|w^o - \boldsymbol{\phi}_k^{(i)}(S)\|^2. \quad (11)$$

Note that in (11), $\boldsymbol{\phi}_k^{(i)}(S)$ denotes an estimate of an $M \times 1$ vector w^o that a node k may achieve while cooperating via diffusion strategy of (8) and (9), directly or indirectly, with the nodes that also belong to the same network subset $S \subseteq \mathcal{N}$. In other words, coalition S can be seen as a connected component of the network $\mathcal{N}$ to which node k belongs. Such a coalition definition is necessary due to the fact that the estimate of node k is not only influenced by the estimates of its neighbors in $\mathcal{N}_k$ but also by the estimates of other nodes to which its neighbors are connected. In contrast to MSD definition in (6), we emphasize now its dependence on S, while we drop the time index i for notational simplicity. Naturally, each node is rationally selfish in the sense that it aims to maximize its own payoff.

C. MSD game analysis

In the sequel we will give a number of definitions from the literature, so as to analyze the *MSD* game defined in the previous subsection. Firstly, based on [17], we provide the formal definition of a canonical game.

Definition 1. *[17] An NTU game $(\mathcal{N}, \upsilon)$ is canonical[2] if*

- *it is in characteristic form, i.e., the following standard conditions should hold:*
  - *(i) The value $\upsilon(S)$ must be a non-empty closed subset of $\mathbb{R}^N$,*
  - *(ii) The value $\upsilon(S)$ must be comprehensive, i.e., if $\mathrm{x} \in \upsilon(S)$ and $\mathrm{y} \in \mathbb{R}^N$ are such that $y_k \leq x_k\ \forall k \in S$, then $\mathrm{y} \in \upsilon(S)$,*
  - *(iii) The set $\{\mathrm{x}\,|\,\mathrm{x} \in \upsilon(S)\, and\, x_k \geq z_k, \forall k \in S\}$ with $z_k = \max\{y_k\,|\,\mathrm{y} \in \upsilon(\{k\})\} < \infty$, $\forall k \in \mathcal{N}$ must be a bounded subset of $\mathbb{R}^N$,*
- *it possesses the superadditivity property, i.e.,*

$$\upsilon(S_1) \cap \upsilon(S_2) \subseteq \upsilon(S_1 \cup S_2), \quad \forall S_1 \subset \mathcal{N}, S_2 \subset \mathcal{N},\ s.t.\ S_1 \cap S_2 = \emptyset. \quad (12)$$

For a game in characteristic form, the value of a coalition S is determined exclusively by the members of that coalition. Intuitively, the superadditivity property means that if a certain outcome can be attained by the disjoint coalitions S_1 and S_2 when acting separately, then it can also be attained by them when acting in concert. Therefore, it is reasonable to study the

[1] Note that, in the literature, a somewhat different definition of the coalition value can be also found, i.e., $\upsilon_p(S) \subseteq \mathbb{R}^{|S|}$, where $\upsilon(S) = \upsilon_p(S) \times \mathbb{R}^{|\mathcal{N}/S|}$.

[2] In some classical papers, e.g., [18], a *game in characteristic function form* is also assumed to be superadditive; however, in this work we rely on the terminology and classification proposed in [17].

properties of the grand coalition, i.e., the coalition of all nodes. Toward this goal, we will use a classical solution concept of coalitional game theory, namely, the *core* [18]-[19], [24]-[26].

A payoff vector x of a canonical game is said to be in its "core" if no set $S \subset \mathcal{N}$, $S \neq \emptyset$ of players can improve upon it. In other words, no coalition S can provide higher payoffs for all its members than that prescribed by the vector x. Now, we will turn this intuitive description of the core concept into a mathematical definition.

Definition 2. *The core $\mathcal{C}(\mathcal{N}, \upsilon)$ of an NTU canonical game $(\mathcal{N}, \upsilon)$ is the set of payoff vectors defined as*

$$\mathcal{C}(\mathcal{N}, \upsilon) = \left\{ \mathrm{x} \in \upsilon(\mathcal{N}) | \, \forall S \subset \mathcal{N}, \nexists \mathrm{y} \in \upsilon(S), \textit{s.t.}\, \mathrm{y}^S \geq \mathrm{x}^S \right\}. \tag{13}$$

The grand coalition $\mathcal{N}$ is said to be stable iff the core of the corresponding game is not empty, i.e., $\mathcal{C}(\mathcal{N}, \upsilon) \neq 0$.

It is of utmost interest in the analysis of any canonical game to determine whether its core is non-empty, i.e., if the grand coalition is stable. In general, proving the non-emptiness of the core is an $NP-$ complete problem since the number of possible coalition structures grows exponentially with the number of players [27], [24]. A fundamental result in proving that the core is non-empty is related to showing that the analyzed game is balanced. For an NTU game, the balancedness condition is sufficient yet not necessary, as proved by Scarf in [25]; therefore, every balanced NTU game has a non-empty core, but not vice versa. For this reason, Billera [26] and a few more authors (e.g., see [28]) refined the notion of a balanced game. However, for our discussion here, it suffices to consider the balancedness condition from [25]. The formal definition follows.

Definition 3. *[25] An NTU game $(\mathcal{N}, \upsilon)$ is called balanced, if the following inclusion statement holds*

$$\bigcap_{S \in \mathcal{B}} \upsilon(S) \subseteq \upsilon(\mathcal{N}), \quad \forall \mathcal{B}, \tag{14}$$

where $\mathcal{B}$ is a balanced subsets' family, i.e., a family of non-empty, proper subsets of $\mathcal{N}$ where there exist positive weights δ_S for $S \in \mathcal{B}$ such that

$$\sum_{S \in \mathcal{B}, k \in S} \delta_S = 1, \quad \forall k \in \mathcal{N}. \tag{15}$$

The weights δ_S have the property that if any player is selected, the sum of the weights corresponding to those coalitions in $\mathcal{B}$ that contain the individual, must be equal to one. For instance, assume a three person game. Then, an example of a balanced family is the collection of all two player coalitions, since $\delta_{\{1,2\}} = \delta_{\{2,3\}} = \delta_{\{1,3\}} = 0.5$ can be selected as a system of weights [25]. The balancedness condition in (14) states that, for each balanced family $\mathcal{B}$, it should hold that the intersection of the sets of payoff vectors corresponding to each coalition S in $\mathcal{B}$ is contained in $\upsilon(\mathcal{N})$. Note that the condition in (14) is obviously related to the superadditivity property in (12), but neither condition directly implies the other [29].

Now, the former definitions enable us to investigate some properties of the *MSD* game defined in III-B. Specifically, we will analyze the *MSD* game in its two forms, based on the type of the combination rules previously discussed in Sec. II. Note that, under the assumptions made in Subsec. III-A, the MSD achieved at each node k in a connected network of N nodes, can be well-approximated by the following expressions [7]:

1) in case of combination weights using the Hastings rule

$$\mathrm{MSD}_k^{\mathrm{Hastings}}(\mathcal{N}) \approx \frac{\mu M}{2} \cdot \left(\sum_{k=1}^{N} \sigma_{v_k}^{-2} \right)^{-1}, \tag{16}$$

2) in case of a doubly stochastic combination rule

$$\mathrm{MSD}_k^{\mathrm{doubly}}(\mathcal{N}) \approx \frac{\mu M}{2} \cdot \frac{1}{N} \cdot \left(\frac{1}{N} \sum_{k=1}^{N} \sigma_{v_k}^{2} \right). \tag{17}$$

Theorem 1. *The MSD game with the Hastings combination rule, under the assumptions A1-A3, is a canonical one and its core, as defined in* (13)*, is a non-empty one.*

Proof. Clearly, the game is in characteristic form. Furthermore, it is superadditive due to the payoff vectors defined by the MSD expression in (16). Apart from having inherently indivisible payoffs, note also that all nodes in the same coalition achieve the same maximum estimation accuracy, and due to this special property, the superadditivity of this game clearly implies its balancedness. Finally, due to the fact that it is balanced, the MSD game with the Hastings combination rule has a non-empty core. □

Theorem 2. *Under the assumptions A1-A3, the MSD game with a doubly stochastic combination rule cannot be guaranteed to be canonical and its core, as defined in* (13)*, cannot be guaranteed to be non-empty.*

Proof. Assume the case of arbitrarily different noise variances $\{\sigma_{v_k}^2\}_{k=1}^N$ at different nodes. Then, one can easily check that, in general, two disjoint coalitions S_1 and S_2 could be found that do not satisfy the superadditivity property in (12). For instance, this is the case where, for some $m \in \{1, 2\}$, it holds that

$$\frac{\mu M}{2} \cdot \frac{1}{|S_m|^2} \sum_{k \in S_m} \sigma_{v_k}^2 < \frac{\mu M}{2} \cdot \frac{1}{\left(|S_1| + |S_2|\right)^2} \sum_{k \in S_1 \cup S_2} \sigma_{v_k}^2.$$

In other words, the condition above actually states that $-\mathrm{MSD}(S_m) > -\mathrm{MSD}(S_1 \cup S_2)$, for some $m \in \{1, 2\}$. Similarly, one could also find two disjoint coalitions S_1 and S_2, where $S_1 \cup S_2 = \mathcal{N}$, that do not satisfy superadditivity. This implies that the core of the game with the setting above, as defined in (13), is empty since at least one coalition may deviate from the grand coalition and provide higher payoffs for all of its members. Thus, in the case of arbitrary noise variances, the MSD game with a doubly stochastic combination rule cannot be guaranteed to have a non-empty core. □

Remark 1: Note that, for sufficiently similar noise variances at all nodes, it may happen that the superadditivity property is satisfied even for a doubly stochastic rule, and thus the corresponding game has a non-empty core. To realize this,

let us assume an extreme case where all nodes have exactly the same noise variances, i.e., $\sigma^2_{v_k} = \sigma^2_v$. Then, the grand coalition would form since the noise variances averaged over any coalition are exactly the same and due to the fact that $1/N < 1/|S|$, for any $S \subset \mathcal{N}$.

Remark 2: Note that the core (and the coalitional stability) can be defined not only with respect to the grand coalition, but also with respect to any coalitional structure $\mathcal{R}$ (all nodes partitioned into some disjoint coalitions), as in the seminal paper by Aumann [30]. This problem can be classified as a static coalition formation game [17], where an external factor may impose a certain coalitional structure, and the main objective is to study this structure's stability. As mentioned before, the number of possible coalition structures grows exponentially with the number of players; thus, it is reasonable to examine one or very few coalition structures. For instance, assume that there exists an $\mathcal{X} = \{X_1, X_2\}$ composed of two disjoint groups of nodes, X_1 and X_2, where within each group the variances are the same or sufficiently similar, i.e., σ^2_{v,X_m}, for $m \in \{1, 2\}$. Now, for a doubly stochastic combination rule, if the variances of the two groups are different enough, i.e.,

$$\frac{\sigma^2_{v,X_m}}{|X_m|} < \frac{|X_m| \cdot \sigma^2_{v,X_m} + |X_\ell| \cdot \sigma^2_{v,X_\ell}}{\left(|X_m| + |X_\ell|\right)^2}, \quad \ell \neq m \tag{18}$$

is satisfied for some $m \in \{1, 2\}$, then the core can be shown to be non-empty for this particular $\mathcal{X}$, i.e., $\mathcal{C}(\mathcal{N}, \upsilon, \mathcal{X}) \neq \emptyset$.[3]

D. NSPE setting generalization: NSPE-MSD game

This subsection is concerned with a coalitional game-theoretical analysis of an NSPE setting. As surveyed in Sec. II, an NSPE setting assumes that not all nodes have exactly the same interests; to put it differently, several phenomena may affect the nodes in a different way (or at different scales).

We leave aside for the moment the formal definition of an NSPE-MSD game. The focus is put instead on the natural question whether the coalitional stability of a single-interest network (traditional diffusion) is related to the coalitional stability in an NSPE setting. The answer is positive, under certain assumptions.

To show this, the first step is to check if there is a way to decouple the estimation processes related to ς^o_j and $\varsigma^o_{j'}$. This could significantly facilitate the analysis of the NSPE-MSD game.

Note that, even under assumptions A1-A2 and A4, in the general case there is an inherent coupling between the estimation processes related to ς^o_j and $\varsigma^o_{j'}$, where $j, j' \in \{1, \ldots, J\}$ and $j \neq j'$. Specifically, this coupling is due to the influence of higher order data moments that are multiplied by μ^2 in the expressions for mean-square performance. Furthermore, this effect occurs regardless of the cooperation strategy. For instance, for a classical, standalone LMS filter see discussion in [31]. On the other hand, for cooperative NSPE settings, an interested reader may consult the papers in [10]-[11]. However, for sufficiently small step sizes (A3) , this effect can be ignored. Based on [11], we will show this for the diffusion NSPE case that is of interest in this paper.

Theorem 3. *Under assumptions A1-A4, for $k \in \mathcal{P}_j \cap \mathcal{P}_{j'}$, the estimation processes of ς^o_j and $\varsigma^o_{j'}$ can be considered uncoupled, i.e.,*

$$\mathrm{MSD}_k(\mathcal{N}) \approx \sum_{j \in \mathcal{I}_k} \mathrm{MSD}_{k,\varsigma_j}(\mathcal{P}_j) \tag{19}$$

where $\mathrm{MSD}_{k,\varsigma_j}(\mathcal{P}_j)$ is the MSD at node k related to estimation of ς^o_j.

Proof. The proof and an illustrative example are given in Appendix A. □

Now, let us recast the problem in a game-theoretic form, i.e.,

- the set of players is given as $\mathcal{N}^{\text{NSPE}} = \{\mathcal{N}^{(1)}, \ldots, \mathcal{N}^{(J)}\}$, where $\mathcal{N}^{(j)} = \mathcal{P}_j$ and $j \in \{1, \ldots, J\}$,
- the coalition value $\upsilon(S) \in \mathbb{R}^{|\mathcal{N}^{(1)}| + \ldots + |\mathcal{N}^{(J)}|}$ is the set of payoff vectors $\upsilon(S) = \upsilon_1\left(S \cap \mathcal{N}^{(1)}\right) \times \ldots \times \upsilon_J\left(S \cap \mathcal{N}^{(J)}\right)$, where $\upsilon_j\left(S \cap \mathcal{N}^{(j)}\right) \in \mathbb{R}^{|\mathcal{N}^{(j)}|}$, $\forall j \in \{1, \ldots, J\}$. The maximum payoff $x^{(j)}_k$, that a node $k \in \mathcal{P}_j$ may achieve in a coalition $S \cap \mathcal{N}^{(j)}$, while estimating parameters ς_j, is given by

$$x^{(j)}_k\left(S \cap \mathcal{N}^{(j)}\right) = -\mathrm{MSD}_{k,\varsigma_j}\left(S \cap \mathcal{N}^{(j)}\right). \tag{20}$$

The definitions above can be seen as a rigorous formalization of the intuition that each player should be defined as a node per each estimation task that is within its interest. In other words, a single node may represent more than one player. Also, an $\mathcal{N}^{\text{NSPE}}$ can be seen as the natural generalization of the notion of grand coalition, for an NSPE setting. Finally, the foremost results of this subsection are summarized in the subsequent theorem.

Theorem 4. *The NSPE-MSD game with the Hastings combination rule, under the assumptions A1-A4, is a canonical one and its core $\mathcal{C}\left(\mathcal{N}^{NSPE}, \upsilon\right)$ is non-empty. Also, under the same assumptions, the core $\mathcal{C}\left(\mathcal{N}^{NSPE}, \upsilon\right)$ of a doubly stochastic combination rule NSPE-MSD game is non-empty iff the core $\mathcal{C}\left(\mathcal{N}^{(j)}, \upsilon_j\right)$ of the underlying MSD game $\left(\mathcal{N}^{(j)}, \upsilon_j\right)$ with a doubly stochastic rule is non-empty for each $j \in \{1, \ldots, J\}$.*

Proof. The proof of the first statement is a generalization of the proof of Thm 1, based on the decoupling of the estimation processes and the core's composition property in [18]. Specifically, the core of an NTU game $(\mathcal{N}, \upsilon)$, which is the composition of two games $\left(\mathcal{N}^{(1)}, \upsilon_1\right)$ and $\left(\mathcal{N}^{(2)}, \upsilon_2\right)$ that have no interconnection, is the Cartesian product of the cores of the two component games, i.e., $\mathcal{C}\left(\mathcal{N}, \upsilon\right) = \mathcal{C}\left(\mathcal{N}^{(1)}, \upsilon_1\right) \times \mathcal{C}\left(\mathcal{N}^{(2)}, \upsilon_2\right)$. In the same vein, for J games, $\mathcal{C}\left(\mathcal{N}, \upsilon\right) = \mathcal{C}\left(\mathcal{N}^{(1)}, \upsilon_1\right) \times \ldots \times \mathcal{C}\left(\mathcal{N}^{(J)}, \upsilon_J\right)$. Next, it has been shown that, under the assumptions A1-A4, the estimation processes in an NSPE setting are decoupled. Thus, the underlying games with respect to each estimation task will not be interconnected, and since each one has a non-empty core if the Hastings rule is used, the first statement of this theorem holds. Finally, the

[3]The analysis of a game with an explicit cooperation cost and the problem of finding a stable coalition structure will be addressed in Sec. IV.

proof of the second statement follows via similar reasoning that is used to prove the first statement. □

IV. Coalitional graph game for parameter estimation

In the previous section, we have studied the *(NSPE-)MSD* games, in which there was no explicit cost for cooperation. In this section, by assuming non-negligible communication cost, we formulate a coalitional graph game to model the NSPE problem in which the nodes' individual payoffs are related to the specific graph that interconnects them[4]. Initially, by relying on [32], the core of a graph game will be introduced as a solution concept, and then, sufficient conditions for its non-emptiness and stability of the grand coalition in our game will be provided. Finally, for a general case in which the grand coalition is not formed, a coalitional graph formation game will be proposed, inspired by a so-called merge-and-split approach [33].

A. NSPE-MSD-Graph-COMM game definition

Let the graph $G_S = (S, \mathcal{E})$ stand for a graph on coalition S, where S and $\mathcal{E}$ are sets of vertices and edges, respectively. Also, let the collection Γ denote the set of all graphs on coalitions $S \subseteq \mathcal{N}$ that can appear in the game. Then, we define an NTU coalitional graph game (Γ, υ) where υ is now a graph-based function, assigning to every graph in Γ a set of payoff vectors. Here, we model the maximum payoff $x_k^{(j)}$ achieved by each node k in a coalition S^j connected by some graph G_{S^j}, per each estimation task $j \in \{1, \ldots, J\}$ where $k \in \mathcal{P}_j$, as:

$$x_k^{(j)}(G_{S^j}) = -\mathrm{MSD}_{k,\varsigma_j}(S^j) - \eta_k^{(j)}(G_{S^j}). \tag{21}$$

In the relation above, the term $\eta_k^{(j)}(G_{S^j})$ represents the total communication cost that is incurred by a node k to establish a coalition S^j, and is given by

$$\eta_k^{(j)}(G_{S^j}) = \sum_{\substack{l \in S^j \\ (k,l) \in G_{S^j}}} \varepsilon^j(k,l)\,, \tag{22}$$

where $\varepsilon^j(k,l)$ is the communication cost required to connect nodes k and l in a communication graph G_{S^j}.

Note that the players in $S \subseteq \mathcal{N}$ and the communication links between them are represented by the vertices and the edges in some graph $G_S \in \Gamma$, respectively. Also, due to the nature of our game, the graphs are assumed to be undirected, and the players can only cooperate if they are connected.

B. NSPE-MSD-Graph-COMM game analysis

To study the game defined above, we will first comment on the objectives and solution concepts developed in the literature on coalitional graph games. A coalitional graph game model was first introduced by Myerson in [34]. Ever since, it has been adopted and adapted in different contexts and applications, see [17],[32],[35],[36]. In this subsection, we will look for a sufficient condition for the stability of the grand coalition, that is, a graph connecting all players in our game.

The core of a graph game can be seen now as the set of all payoff vectors attainable for some allowed graph connecting all players that are not dominated by any coalition connected through a graph in the collection of allowed graphs [32]. In other words, a payoff vector lies in the core of the graph game if there is no possible graph on a coalition which can make all of its members better off. The formal definition follows.

Definition 4. *The core $\mathcal{C}(\Gamma, \upsilon)$ of an NTU graph game (Γ, υ) is the set of payoff vectors $\mathrm{x} \in \mathbb{R}^N$ satisfying, for at least one $G_{\mathcal{N}} \in \Gamma$, that $\mathrm{x} \in \upsilon(G_{\mathcal{N}})$ and that it does not exist a coalition $S \subset \mathcal{N}$, a graph $G_S \in \Gamma$ and a vector $\mathrm{y} \in \upsilon(G_S)$ such that $y_k \geq x_k$, $\forall k \in S$.*

Next, based on [32] as well as on particularities of the NSPE problem, we define sufficient conditions for the non-emptiness of the core of our graph game.

Definition 5. *An NTU graph game (Γ, υ), with power measure function m, has a non-empty (balanced) core $\mathcal{C}(\Gamma, \upsilon)$ if, $\forall j \in \{1, \ldots, J\}$, it holds that*

(i) *$\forall k \in \mathcal{N}^{(j)}$, the graph $G_{\{k\}} \in \Gamma$ and for some $z_k \in \mathbb{R}$, the set $\upsilon(G_{\{k\}})$ is given by $\upsilon(G_{\{k\}}) = \{\mathrm{x} \in \mathbb{R}^{|\mathcal{N}^{(j)}|} \,|\, x_k \leq z_k\}$,*

(ii) *$\forall S^j \subseteq \mathcal{N}^{(j)}$ and $\forall G_{S^j} \in \Gamma$, the set $\{(x_k)_{k \in S^j} \in \mathbb{R}^{|S^j|} \,|\, \mathrm{x} \in \upsilon(S) \text{ and } x_k \geq z_k, \forall k \in S^j\}$ must be bounded,*

(iii) *$\forall S^j \subseteq \mathcal{N}^{(j)}$ and $\forall G_{S^j} \in \Gamma$, the set $\upsilon(G_{S^j})$ is closed and comprehensive, i.e., if $\mathrm{x} \in \upsilon(G_{S^j})$ and $\mathrm{y} \in \mathbb{R}^{|\mathcal{N}^{(j)}|}$ are such that $y_k \leq x_k$ $\forall k \in S^j$, then $\mathrm{y} \in \upsilon(G_{S^j})$,*

(iv) *the game is graph balanced, i.e., if, for at least one $G_{\mathcal{N}^{(j)}} \in \Gamma$, it holds that*

$$\cap_{p=1}^{r} \upsilon(F^p) \subseteq \upsilon(G_{\mathcal{N}^{(j)}}), \quad \forall \mathcal{F}, \tag{23}$$

where $\mathcal{F} = \{F^1, \ldots, F^r\}$ is a graph balanced family in $\Gamma^{(j)}$, i.e., a family of r graphs in $\Gamma^{(j)}$ where, for a given power measure function m, there exist positive numbers λ_p, with $\sum_{p=1}^{r} \lambda_p = 1$, such that

$$\sum_{p=1}^{r} \lambda_p m(F^p) = \frac{1}{N} \cdot 1_N. \tag{24}$$

In the definition above, a power measure m on some graph G_{S^j} is simply a vector $m(G_{S^j}) \in \mathbb{R}^{|\mathcal{N}^{(j)}|}$ with the elements satisfying $\sum_{k \in G_{S^j}} m_k(G_{S^j}) = 1$ and being zeros elsewhere. For instance, a power measure may reflect the relative position of each player within the graph. Also, in the definition above, $\Gamma^{(j)}$ represents all allowed graphs related to an estimation task j which can appear in the game.

Let us define now some useful quantities for obtaining sufficient conditions for the grand coalition stability in our graph game. Firstly, let the term $gain_{min}^{(j)}$ correspond to the minimum possible gain that some coalition S_1 can obtain after

[4]In our preliminary work on this topic [1], we have analyzed coalitional formation game in the NSPE setting where a broadcast communication cost has been assumed. Here, a more general communication cost, namely, a graph-based function, is assumed. Therefore, we will detail a stability concept applicable to such a game.

merging with another coalition S_2, while estimating ς_j^o, as

$$gain_{min}^{(j)} = \min_{\substack{S_1^j, S_2^j \subset \mathcal{N} \\ S_1^j \cap S_2^j = \emptyset}} \left\{ gain^{(j)}(S_1, S_2) \right\} \tag{25}$$

where

$$gain^{(j)}(S_1, S_2) = \mathrm{MSD}_{k,\varsigma_j}(S_1^j) - \mathrm{MSD}_{k,\varsigma_j}(S_1^j \cup S_2^j). \tag{26}$$

Next, consider that the edge with the highest communication cost among the nodes in $\mathcal{P}_j$ is given as

$$\varepsilon_{max}^{(j)} = \max_{\substack{(k,l) \in G_{S^j} \\ \forall G_{S^j} \in \Gamma}} \varepsilon^j(k,l). \tag{27}$$

Theorem 5. *The core $\mathcal{C}(\Gamma, \upsilon)$ of the NSPE-MSD-Graph-COMM game with Hastings combination rule, under the assumptions A1-A4, is non-empty if*

$$gain^{(j)}\left(\mathcal{P}_j \setminus \{k\}, \{k\}\right) \geq \varepsilon_{max}^{(j)}, \qquad \forall j \in \{1, \ldots, J\}. \tag{28}$$

where k is the node with the worst variance in a set $\mathcal{P}_j$, i.e., $\sigma_{v,k}^2 = \max_{l \in \mathcal{P}_j} \{\sigma_{v,l}^2\}$. Furthermore, for each estimation task j, there exists at least one graph in that core with a tree structure.

Proof. Clearly, $\mathrm{MSD}_{k,\varsigma_j}(G_{S^j}) = \mathrm{MSD}_{k,\varsigma_j}(S^j)$, for all $G_{S^j} \in \Gamma$. Also, observe that the *NSPE-MSD-Graph-COMM* game is graph balanced if

$$gain_{min}^{(j)} \geq \varepsilon_{max}^{(j)}, \qquad \forall j \in \{1, \ldots, J\}. \tag{29}$$

Next, due to following property of the Hastings combination rule

$$\mathrm{MSD}_{k,\varsigma_j}(\mathcal{P}_j) \leq \mathrm{MSD}_{k,\varsigma_j}(S_2^j), \quad \forall\, S_2^j \subseteq \mathcal{P}_j, \tag{30}$$

it holds that

$$gain_{min}^{(j)} = gain^{(j)}\left(\mathcal{P}_j / \{k\}, \{k\}\right), \tag{31}$$

if $\sigma_{v,k}^2 = \max_{l \in \mathcal{P}_j} \{\sigma_{v,l}^2\}$.

Now, assume that a graph $G_{\mathcal{N}}$ whose payoff vector $\mathrm{x} \in \upsilon(G_{\mathcal{N}})$ lies in the core $\mathcal{C}(\Gamma, \upsilon)$ is not a tree. This implies that at least one communication link, e.g., (l, i), can be removed while still preserving the same coalition size connected through a graph $G'_{\mathcal{N}}$. Clearly, $x_k(G'_{\mathcal{N}}) > x_k(G_{\mathcal{N}})$, for $k \in \{l, i\}$, due to (21), while the best payoffs all other nodes may achieve remain unchanged. Accordingly, the non-empty core must contain payoff vectors corresponding to a tree graph. □

We remark that the sufficiency conditions in (28) are rather conservative. Observe that obtaining more specific and sharper sufficient conditions for the core non-emptiness of a graph game is highly dependent on the set of allowed graphs Γ, both on the number of graphs and their allowed structures. For instance, assume that only a single star topology network $G_{\mathcal{N}}$ and its subgraphs are allowed in a graph game. Also, assume that k stands now for the central node in both $G_{\mathcal{P}_j}$, $\forall j \in \{1, \ldots, J\}$, while l denotes a leaf node. Then, sufficient conditions for its core to be non-empty can be written as

$$gain^{(j)}\left(\{k\}, \mathcal{P}_j \setminus \{k\}\right) \geq \sum_{l \in \mathcal{P}^j} \varepsilon^j(k,l) \tag{32}$$

and

$$gain^{(j)}\left(\{l\}, \mathcal{P}_j \setminus \{l\}\right) \geq \varepsilon^j(k,l), \quad \forall l \in \mathcal{P}_j \setminus \{k\}. \tag{33}$$

For this specific example, the conditions above not only are sharper than the one in (28) but also necessary. In fact, one should check the benefits of the grand coalition only with respect to the fully non-cooperative case.

C. NSPE-MSD Graph Formation game based on Merge-Split

Due to the fact that the grand coalition will not form in general, this subsection is devoted to answering the question of which coalition graphs will form. Note that finding stable coalition groups involves exponential computational complexity with respect to the number of players [27]. In addition to this, here we also need to find a proper coalition graph connecting the nodes in each coalition group, which increases the complexity even further. Thus, the main challenge here is to devise an efficient algorithm that allows nodes to make and break coalitions in a distributed way. Specifically, the focus will be put on the so-called merge-split strategy, being firstly presented in [33], and investigated in the context of several applications [17], [24], [37]-[39]. To that aim, let us make the following definitions.

Definition 6. *A collection of coalitions $\mathcal{S}$ is the set of mutually disjoint coalitions, i.e., $\mathcal{S} = \{S_1, \ldots, S_q\}$, where $S_n \subset \mathcal{N}$ for $n = 1, \ldots, q$. If a collection $\mathcal{S}$ comprises all the nodes of $\mathcal{N}$, then the collection $\mathcal{S}$ is a partition of $\mathcal{N}$. Finally, a collection of graphs on a collection $\mathcal{S}$ is the set of specific graphs on each coalition in $\mathcal{S}$, i.e., $\mathcal{G}_{\mathcal{S}} = \{G_{S_1}, \ldots, G_{S_q}\}$.*

Definition 7. *The Pareto order operator $\triangleright$ for comparing two collections of graphs on the same collection $\mathcal{S} = \{S_1, \ldots, S_q\}$, i.e., $\mathcal{G}_{\mathcal{S}}$ and $\mathcal{G}'_{\mathcal{S}}$, is defined as follows*

$$\mathcal{G}_{\mathcal{S}} \triangleright \mathcal{G}'_{\mathcal{S}} \longleftrightarrow x_k(\mathcal{G}_{\mathcal{S}}) \geq x_k(\mathcal{G}'_{\mathcal{S}}), \quad \forall k \in \mathcal{S} \tag{34}$$

with at least one node satisfying the strict inequality $>$. Similarly, for two collections of graphs on two different collections $\mathcal{X} = \{X_1, \ldots, X_m\}$ and $\mathcal{S} = \{S_1, \ldots, S_q\}$, where $\mathcal{X}$ and $\mathcal{S}$ are partitions of the same subset of nodes $\mathcal{A} \subseteq \mathcal{N}$, the notation $\mathcal{G}_{\mathcal{X}} \triangleright \mathcal{G}'_{\mathcal{S}}$ means that $x_k(\mathcal{G}_{\mathcal{X}}) \geq x_k(\mathcal{G}'_{\mathcal{S}})$, $\forall k \in \mathcal{X}, \mathcal{S}$, with at least one node satisfying the strict inequality.

Now, a distributed graph formation algorithm based on the merge, rearrange and split steps is provided in the table on the following page. In the algorithm, note that J separate graph formation games are given, an $|\mathcal{N}^{(j)}|-$player game for each estimation task j, where $j \in \{1, \ldots, J\}$.

Initially, a merge step is perfomed based on the Pareto order. In other words, a merge of some coalitions is possible if the payoff of each node involved either increases or remains the same but it never decreases. Practically, this step can be implemented as follows. Starting from a collection of graphs related to a network partition $\mathcal{S}^{(j)}$, for each estimation task $j \in \{1, \ldots, J\}$, some coalition S_i^j, connected through

Algorithm 1 Merge-rearrange-split for NSPE diffusion

Initial state: Start with some set of graphs $\{G_{S_1^j}, \ldots, G_{S_q^j}\}$ where $\mathcal{S}^{(j)} = \{S_1^j, \ldots, S_q^j\}$ is a partition of $\mathcal{N}^{(j)}$, $\forall j \in \{1, \ldots, J\}$.

Proposed Graph Formation for NSPE Algorithm:

Phase I - Adaptive coalitional graph formation:

repeat

a) $\forall j \in \{1, \ldots, J\}$, merge any set of coalition graphs $\{G_{S_1^j}, \ldots, G_{S_q^j}\}$ if it holds $G_{\cup_{n=1}^q S_n^j} \triangleright \{G_{S_n^j}\}_{n=1}^q$ and if the newly formed graph exists in the set of allowed graphs, i.e., $G_{\cup_{n=1}^q S_n^j} \in \Gamma$. Thus, it is possible to merge only with those coalition(s) to which its member nodes have direct communication link(s).

b) $\forall j \in \{1, \ldots, J\}$, rearrange the graph $G_{\cup_{n=1}^q S_n^j}$ to $G'_{\cup_{n=1}^q S_n^j}$ on any coalition $\bigcup_{n=1}^q S_n^j$, if there exists $G'_{\cup_{n=1}^q S_n^j}$ such that $G'_{\cup_{n=1}^q S_n^j} \triangleright G_{\cup_{n=1}^q S_n^j}$. The search should exclude the leaf nodes that are connected by the edges which are of the minimum costs among all edges that are allowed to these leaf nodes.

c) $\forall j \in \{1, \ldots, J\}$, split any coalition graph $G'_{\cup_{n=1}^q S_n^j}$ if it holds $\{G'_{S_n^j}\}_{n=1}^q \triangleright G'_{\cup_{n=1}^q S_n^j}$ and if all corresponding graphs are allowed, i.e., $\{G'_{S_n^j}\}_{n=1}^q \in \Gamma$.

until merge-rearrange-split iteration terminates.

Phase II - NSPE diffusion strategy: Run the adaptation (4) and the combination step (5) at each time instant i.

Repeat periodically: To account for time-varying noise variances and/or changing communication costs during the network operation.

graph $G_{S_i^j}$, starts the merging process by performing pairwise negotiations with other coalitions (after e.g., random pairing) to which there are possible communication links. In the negotiation process, some coordination among the nodes in the same coalition is necessary in general, and can be realized, for instance, through the coalition head. In the case where the Hastings combination rule is employed, note that the negotiation process is easier to perform since the merge decision depends only on the two nodes which are to be paired; other nodes in their coalitions will not be worse off since they do not have additional communication cost and the estimation accuracy will not be decreased irrespectively of the negotiation decision. In case where the coalitions decide to merge, based on the Pareto order, a new edge is added between the graphs. Next, the newly formed coalition proceeds with the search for merging until it is possible. The merge search is repeated for all other coalitions from $\mathcal{S}^{(j)}$ that have not been merged yet. In the rearrange step, each coalition aims at finding a tree that Pareto dominates the current graph. Two types of situations may occur. Firstly, if the graph $G_{\cup_{n=1}^q S_n^j}$ has at least one cycle, then every edge is checked to be deleted, starting from the one with the maximum cost, until $G_{\cup_{n=1}^q S_n^j}$ has no cycles while still being connected. Secondly, if graph $G_{\cup_{n=1}^q S_n^j}$ is a tree, then a tree that dominates it , according to the Pareto order, is being sought. This step can be realized through the coalition head that knows the topology and communication costs within the coalition. Finally, the split step, for every previously formed coalition $\bigcup_{n=1}^q S_n^j$, searches through every partition of $\bigcup_{n=1}^q S_n^j$ whose graphs are allowed and are subgraphs of $G_{\cup_{n=1}^q S_n^j}$. To realize this step, in addition to topology information the coalition head should have knowledge of $\sigma_{v_k}^2$ related to the nodes in the coalition being examined.

Note that the search in all three steps, and thereby the complexity of the algorithm, is limited by the set Γ. Thus, the complexity can be significantly reduced for the trees with several leafs.

The algorithm performs several cycles of these three steps until it converges to a stable graph structure, which can be defined as the graph structure where no coalitions have an incentive to pursue coalitional graph formation through merge, arrange or split steps. This is closely related to the so-called $\mathbb{D}_{hp}$-stability concept, proposed in [33] for the original merge-split algorithm. Specifically, a partition $\mathcal{S}$ is $\mathbb{D}_{hp}$-stable, if, for the partition $\mathcal{S}$, no coalition has an incentive to split or merge [33], [37]. Furthermore, as every iteration of merge-split terminates, a resulting partition from such iterations cannot be subject to any further merge or split; thus, every partition resulting from the merge-split algorithm is $\mathbb{D}_{hp}$-stable [37].

Therefore, our proposed merge-rearrange-split algorithm also terminates for any initialization, and the coalition structure always converges to a stable coalition structure where no player has incentive to leave its coalition graph. To realize this, note first that in the specific case where only complete graphs are allowed by Γ, the proposed algorithm reduces to the classical merge-split. On the other hand, in the general case, there is an additional degree of freedom w.r.t. the original merge-split, and it is related to finding a proper graph on a given coalition. Due to the fact that this degree of freedom is also optimized using the Pareto order, where in each graph change no node in the graph is worse off, the proposed algorithm terminates and converges to the coalition structure where no coalition (graph) has an incentive to split, rearrange or merge.

V. Simulations

In this section, we initially provide some computer simulations that verify the main findings presented in Sec. III. Later, the graph game results of the previous section are illustrated by simulating the proposed graph formation algorithm.

A. Validation of Theorems 1-2 and 4-5

Let us remind that in order to prove Thm. 1-2 and Thm. 4-5, the theoretical expressions for MSD in the steady state (16)-(17) have been used. Here, we confirm the results by firstly simulating a traditional diffusion LMS setting, and then, its NSPE-based counterpart.

We consider a network of $N = 15$ nodes. The measurements follow the observation model (1) with $w_k^o = w^o$ and $M_k = M = 8$, $\forall k \in \mathcal{N}$. By following assumption

A3, the stepsize in the adaptation step (8) is chosen to be sufficiently small, i.e., $\mu = 5 \cdot 10^{-3}$. The regressors $\mathbf{u}_{k,i}$ are zero mean Gaussian random variables with autocovariance $R_u = I$. Next, the noise variances for nodes 5, 6 and 11 are $\sigma^2_{v_5} = \sigma^2_{v_6} = \sigma^2_{v_{11}} = 7 \cdot 10^{-3}$, while for all other nodes they are chosen from the interval $(0.05, 0.2)$. The network is connected by a topology where $4 \leq n_k \leq 6$ for any node $k \in \mathcal{N}$, see Fig. 9. For each plot, the results are averaged over 50 randomly initialized independent experiments. The steady-state curves are generated by running the algorithms for 3 000 iterations. The quantities of interest, namely, MSD at each node, are then obtained by averaging the last 300 samples of the corresponding learning curves. For each independent experiment, the deterministic vector w^o is randomly selected from the interval $(0, 1)$ and then normalized such that its norm equals 1.

For the diffusion LMS, we consider the following coalition structures,

a) the grand coalition, i.e., where nodes cooperate all together,
b) two disjoint coalitions of nodes specified by $X_1 = \{5, 6, 11\}$ and $X_2 = \mathcal{N} \setminus X_1$,
c) random coalition structure with at least two coalitions.

Also, we include the LMS-based non-cooperative strategy in the comparison.

Figure 3 depicts the comparison in the case where the traditional diffusion LMS employs the Hastings combination rule (10). Although there is a significant diversity among the noise variances at different nodes, it can be observed that the grand coalition is beneficial for all nodes.

On the other hand, in this simulation setting the grand coalition does not bring improvement to all nodes in the case of doubly-stochastic combination weights, as shown in Fig 4. Here, as a doubly stochastic rule, we use the Metropolis rule given by

$$c_{k,l} = \begin{cases} \frac{1}{\max\{n_k, n_l\}} & l \in \mathcal{N}_k \setminus \{k\} \\ 1 - \sum_{m \in \mathcal{N}_k \setminus \{k\}} c_{k,m} & l = k \end{cases}. \quad (35)$$

Note that the disjoint coalitions structure (b) is the most beneficial for coalition X_1, while the grand coalition is the most beneficial for the nodes in X_2. However, since X_1 can deviate from the grand coalition, the grand coalition is not stable. Furthermore, it can be observed that the disjoint coalitions structure (b) is stable, since no node or group of nodes can leave coalition structure $\mathcal{X} = \{X_1, X_2\}$ and make all its members better off.

In Remarks 1 and 2 of Sec. III-C, it has been suggested that, for sufficiently similar noise variances of all nodes, the grand coalition can be stable. Therefore, we now simulate the non-cooperative LMS and the traditional diffusion LMS for the grand coalition structure (a) and two disjoint coalitions X_1 and X_2 as in (b), for the doubly stochastic combination rule in (35). Specifically, Figure 5 presents the results related to node 5 (which is in X_1) for different noise variance diversity coefficient defined as $\sigma^2_{v,X_2}/\sigma^2_{v,X_1}$. The noise variance σ^2_{v,X_2} takes different values such that the coefficient in the range from 10 to 2. It can be observed that when the diversity

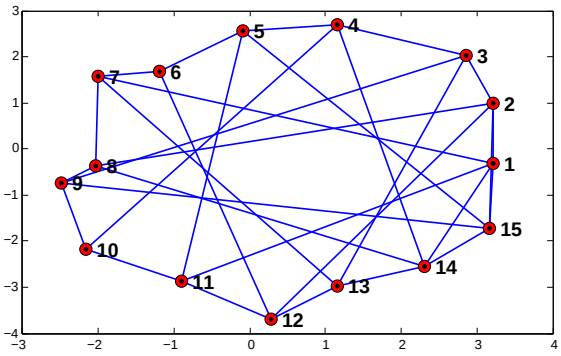


Fig. 2. Network topology.

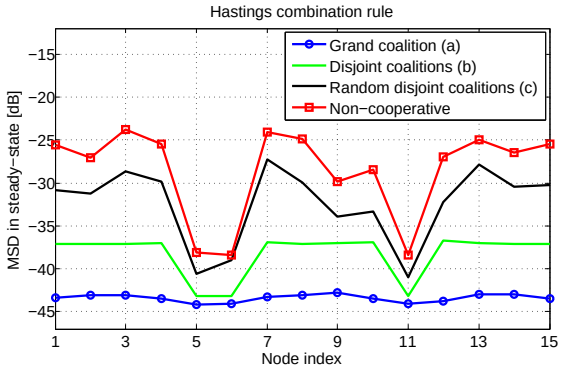


Fig. 3. Steady-state MSD per node for the Hastings combination rule in the traditional setting, i.e., $w^o_k = w^o$, $\forall k \in \mathcal{N}$.

coefficient is below 6, the grand coalition is more beneficial than the disjoint coalition structure (b) and thus, it is stable. Note also that, for $|X_1| = 3$ and $|X_2| = 12$, the condition in (18), which characterize when the disjoint coalition structure (b) can be stable, reduces to $6 \cdot \sigma^2_{v,X_1} < \sigma^2_{v,X_2}$; thus, Fig. 5 verifies this condition as well.

Next, Figure 6 analyzes an NSPE scenario with a simulation setting similar to the one for Figs. 3 and 4, yet the results for Hastings and doubly stochastic rule are plotted together. Here, there are two estimation tasks, i.e., ς^o_a and ς^o_b, and their areas of influence are set to be $\mathcal{P}_a = \{1, \ldots, 9\}$ and $\mathcal{P}_b = \{7, \ldots, 15\}$, respectively. For each independent experiment, vectors ς^o_a and ς^o_b are randomly selected from the interval $(0, 1)$ and then normalized. The filter lengths $M_{\varsigma,a}$ and $M_{\varsigma,b}$ are chosen to be the same, i.e., $M_{\varsigma,a} = M_{\varsigma,b} = 8$. Similarly to the traditional diffusion scenario in Figs. 3 and 4, it is shown that the grand coalition for the Hastings case is stable for both estimation tasks, while the core $\mathcal{C}\left(\mathcal{N}^{\text{NSPE}}, v\right)$ for the doubly stochastic rule is empty in this simulation setting.

B. Graph game

In this subsection, we simulate the proposed NSPE merge-rearrange-split protocol so as to illustrate its effectiveness and also to verify the discussion in Sec. IV.

The communication cost between the nodes k and ℓ is modeled by a simple exponential model, i.e., $\varepsilon(k, \ell) = \beta \cdot e^{(r_{k,\ell}/r_0)}$,

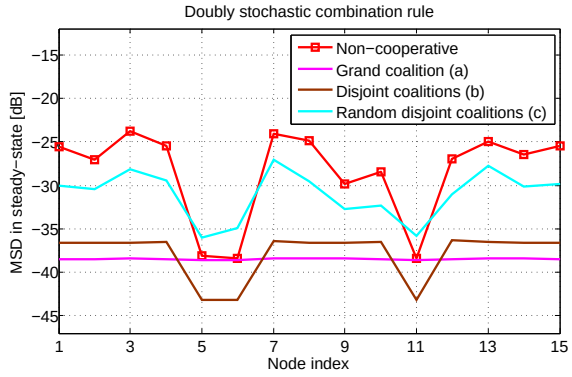


Fig. 4. Steady-state MSD per node for a doubly stochastic combination rule in the traditional setting, i.e., $w_k^o = w^o$, $\forall k \in \mathcal{N}$.

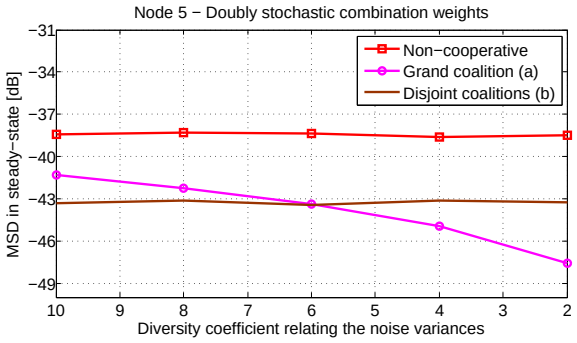


Fig. 5. Steady-state MSD of node 5 for different variance similarity coefficients related for the doubly stochastic combination rule in the traditional setting.

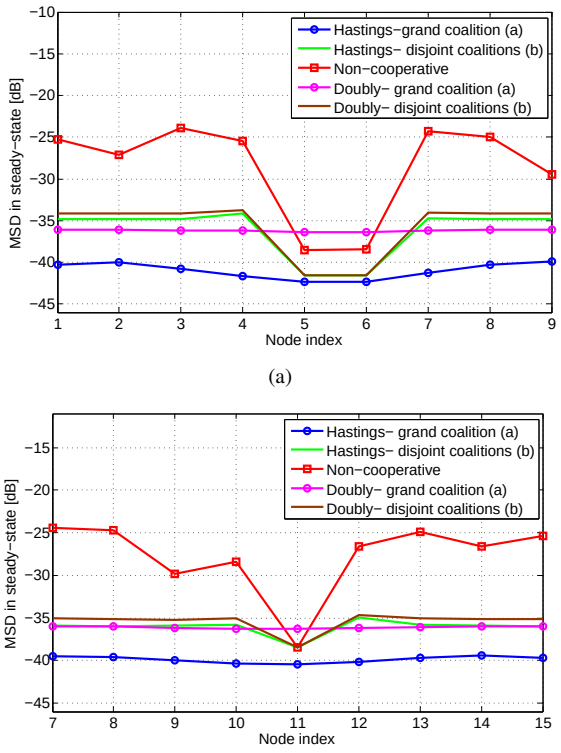


Fig. 6. Steady-state MSD per node for the NSPE setting with (a) estimation of ς_a^o and (b) estimation of ς_b^o.

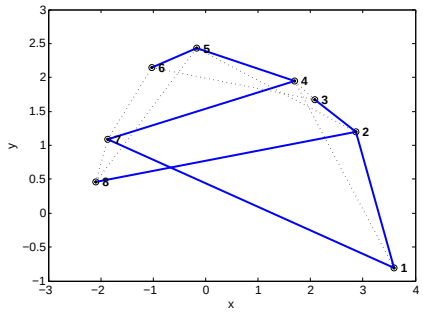
where r_0 is a reference distance, $r_{k,\ell}$ denotes the distance between the nodes k and ℓ, while β is a normalization coefficient. To establish a coalition S^j through some graph G_{S^j}, the total communication cost $\eta_k^{(j)}(G_{S^j})$ for each node $k \in S^j$ is the sum of its active communication links, as in (22). We consider a network of $N = 15$ nodes where each node k has a noise variance $\sigma_{v_k}^2$ between 0.1 and 0.6. There are two vectors of parameters to be estimated, i.e., ς_a^o and ς_b^o, with $\mathcal{P}_a = \{1, \ldots, 8\}$ and $\mathcal{P}_b = \{6, \ldots, 15\}$. We set $\mu = 0.001$ and $M_{\varsigma,a} = M_{\varsigma,b} = 10$. We consider the Hastings combination policy so we use the MSD expressions provided in (48).

In Fig. 7, an outcome of a single realization of the proposed protocol is presented. For both estimation tasks, all possible links between the nodes are denoted by the dotted lines. On the other hand, the graphs which are the outcome of the graph formation algorithm for both estimation tasks are given by the solid lines. Normalization coefficient β is set such that the condition in (28) is satisfied. It can be noticed that the resulting graphs for both estimation tasks are indeed the trees.

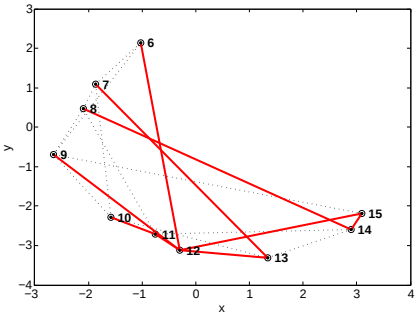
Figure 8 plots the number of nodes in coalition graphs resulting from the NSPE merge-rearrange-split algorithm as a function of the communication cost, which is characterized by normalization coefficient β. For both estimation tasks, the maximum and the average coalition size are presented. The results have been averaged over 50 experiments where we randomly set $r_{k,\ell}/r_0$ between 0 and 6. On one hand, it can be observed that for no, or relatively small, communication costs, the number of nodes connected in coalition graphs that estimate ς_a^o and ς_b^o are $|\mathcal{P}_a| = 8$ and $|\mathcal{P}_b| = 10$, respectively. On the other hand, by increasing the communication costs, the graphs on coalitions get split, and finally, they reduce to the non-cooperative nodes.

VI. Conclusions

In this article, a distributed adaptive parameter estimation problem has been modeled as a non-transferable coalitional game. Initially, we have studied the parameter estimation problem via traditional diffusion strategy as a canonical game; afterwards, the analysis has been extended to NSPE setting. The analysis has focused on the Hastings combination rule and on the group of combination policies that are doubly stochastic. For both, the traditional and the NSPE scenarios, the grand coalition has been shown to be stable coalition structure for the Hastings policy, while some insights have been provided for the doubly stochastic ones. Then, we have proposed a graph game for the NSPE setting that accounts for the communication costs, which are modeled as graph-based functions, and its stability has been studied. Afterwards, by relying on the merge-split approach, a distributed algorithm for graph formation has been proposed. Finally, indicative computer simulations have verified the theoretical analyses.



(a)



(b)

Fig. 7. Tree coalition graph structure for: (a) estimation of ς_a^o and (b) estimation of ς_b^o.

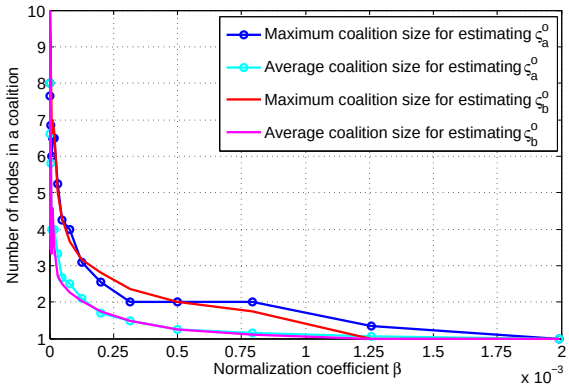


Fig. 8. Maximum and average coalition size versus communication cost for the NSPE network with $\mathcal{P}_a = \{1, \ldots, 8\}$ and $\mathcal{P}_b = \{6, \ldots, 15\}$.

Appendix A
Proof of Theorem 3

Firstly, we need to provide some network recursions related to mean-square performance of a diffusion NSPE algorithm. To this aim, we define the weight-error vectors $\tilde{\boldsymbol{\phi}}_{k,\varsigma}^{(i)} = w_k^o - \boldsymbol{\phi}_{k,\varsigma}^{(i)}$, where we use notation introduced in Sec. II. Next, we stack the weight-error vectors from all nodes, at each time i, as follows

$$\tilde{\boldsymbol{\phi}}_i = \mathrm{col}\{\tilde{\boldsymbol{\phi}}_{1,\varsigma}^{(i)}, \ldots, \tilde{\boldsymbol{\phi}}_{N,\varsigma}^{(i)}\}.$$

Similarly to [11], by subtracting both sides of (4) and (5) from w_k^o, and using the observation model in (3), we obtain the following network-wide error recursion, i.e.,

$$\tilde{\boldsymbol{\phi}}_i = \breve{\mathcal{C}}\,(I - \mu\boldsymbol{\mathcal{D}}_i)\tilde{\boldsymbol{\phi}}_{i-1} - \mu\,\breve{\mathcal{C}}\,\boldsymbol{\mathcal{V}}_i\,. \tag{36}$$

In the relation above, the vector $\boldsymbol{\mathcal{V}}_i$ is of dimension $\breve{M} \times 1$, where $\breve{M} = \sum_{k=1}^N M_k$, and is given by

$$\boldsymbol{\mathcal{V}}_i = \mathrm{col}\{\mathbf{u}_{1,i}^H \mathbf{v}_{1,i}, \ldots, \mathbf{u}_{N,i}^H \mathbf{v}_{N,i}\}, \tag{37}$$

while $\boldsymbol{\mathcal{D}}_i$ is an $\breve{M} \times \breve{M}$ block-diagonal matrix, defined below

$$\boldsymbol{\mathcal{D}}_i = \mathrm{diag}\{\mathbf{u}_{1,i}^H \mathbf{u}_{1,i}, \ldots, \mathbf{u}_{N,i}^H \mathbf{u}_{N,i}\}. \tag{38}$$

Finally, the extended combination matrix $\breve{\mathcal{C}}$ in (36) has the following form

$$\breve{\mathcal{C}} = \mathrm{col}\left\{\{C_1^{\varsigma_j}\}_{j\in\mathcal{I}_1}, \ldots, \{C_N^{\varsigma_j}\}_{j\in\mathcal{I}_N}\right\}, \tag{39}$$

where the corresponding blocks are given as

$$C_k^{\varsigma_j} = \begin{bmatrix} C_{k1}^{\varsigma_j} & C_{k2}^{\varsigma_j} & \ldots & C_{kN}^{\varsigma_j} \end{bmatrix} \tag{40}$$

with

$$C_{k\ell}^{\varsigma_j} = \begin{cases} \begin{bmatrix} 0_{M_{\varsigma,j}\times M'_\ell} & c_{k,\ell}^{\varsigma_j} I_{M_{\varsigma,j}} & 0_{M_{\varsigma,j}\times(M_\ell - M'_\ell - M_{\varsigma,j})} \end{bmatrix} \\ \qquad\qquad\qquad\qquad \text{if } \ell \in \mathcal{P}_j, \text{and} \\ 0_{M_{\varsigma,j}\times M_\ell} \qquad\qquad \text{if } \ell \notin \mathcal{P}_j, \end{cases} \tag{41}$$

and M'_ℓ in (41) defined as $M'_\ell = \sum_{j'\in\mathcal{I}_{j\ell}} M_{\varsigma,j'}$ where $\mathcal{I}_{j\ell} = \{j' \in \mathcal{I}_\ell : j' < j\}$.

Now, after equating the weighted norm of (36), taking the expectation and relying on assumptions A1-A2, and A4, we get the following relation

$$\begin{aligned} E\|\tilde{\boldsymbol{\phi}}_i\|_\Sigma^2 = E\left\{\|\tilde{\boldsymbol{\phi}}_{i-1}\|^2_{(I-\mu\boldsymbol{\mathcal{D}}_i)^H\,\breve{\mathcal{C}}^T\,\Sigma\,\breve{\mathcal{C}}\,(I-\mu\boldsymbol{\mathcal{D}}_i)}\right\} \\ + E\left\{\mu^2 \boldsymbol{\mathcal{V}}_i^H\,\breve{\mathcal{C}}^T\,\Sigma\,\breve{\mathcal{C}}\,\boldsymbol{\mathcal{V}}_i\right\}, \end{aligned} \tag{42}$$

where Σ is an arbitrary $(\breve{M} \times \breve{M})$ Hermitian nonnegative-definite matrix that we are free to choose, so as to be able to express the MSD performance criterion.

Now, based on analyses in [11], [5], for sufficiently small step sizes we have

$$E\left\{(I - \mu\boldsymbol{\mathcal{D}}_i)^H\,\breve{\mathcal{C}}^T\,\Sigma\,\breve{\mathcal{C}}\,(I - \mu\boldsymbol{\mathcal{D}}_i)\right\} \approx \mathcal{B}^H\Sigma\mathcal{B}\,, \tag{43}$$

where $\mathcal{B} = \breve{\mathcal{C}}\,(I - \mu\mathcal{R}_U)$ with $\mathcal{R}_U = \mathrm{diag}\Big\{\{R_{u,k}\}_{k=1}^N\Big\}$, while the autocovariance matrix of each node k has the following structure $R_{u,k} = \mathrm{diag}\Big\{\{R_u^{(j)}\}_{j\in\mathcal{I}_k}\Big\}$. Thus, the relation in (42) can be written as

$$E\|\tilde{\boldsymbol{\phi}}_i\|_\Sigma^2 = E\|\tilde{\boldsymbol{\phi}}_{i-1}\|^2_{\mathcal{B}^H\Sigma\mathcal{B}} + \mathrm{Tr}\,(\mathcal{Y}\Sigma)\,. \tag{44}$$

where $\mathcal{Y} = \mu^2\,\breve{\mathcal{C}}\cdot\mathrm{diag}\Big\{\{\sigma_{v,k}^2 R_{u,k}\}_{k=1}^N\Big\}\cdot\breve{\mathcal{C}}^T$. Finally, by using the Neumann series property, i.e., $(I - X)^{-1} = \sum_{m=0}^\infty X^m$, we get

$$\lim_{i\to\infty} E\|\tilde{\boldsymbol{\phi}}_i\|_\Sigma^2 = \sum_{m=0}^{\infty} \mathrm{Tr}\,\left(\mathcal{B}^m\mathcal{Y}(\mathcal{B}^H)^m\Sigma\right). \tag{45}$$

Let us focus now on analyzing MSD at some node k, where $k \in \mathcal{P}_j \cap \mathcal{P}_{j'}$ and $j, j' \in \{1, \ldots, J\}$ and $j \neq j'$. Specifically, let us on focus on the MSD with respect to a single interest, for

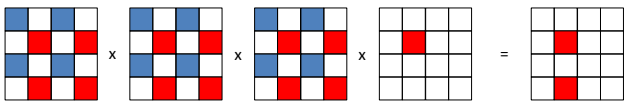


Fig. 9. The product of $\mathcal{B}^m\mathcal{Y}(\mathcal{B}^H)^m\Sigma$ in (46).

instance, ς_j^o. To that aim, we set the selection matrix Σ to be an $\breve{M}\times\breve{M}$ zero matrix, except its block between $(\sum_{\ell=1}^{k-1}M_\ell+M_k')$-th and $(\sum_{\ell=1}^{k}M_\ell-M_k'-M_{\varsigma,j})$-th rows and columns which is set to be $I_{M_{\varsigma,j}}$. Now, it can be checked that due to the peculiar structure that matrices $\mathcal{B}$ and $\mathcal{Y}$ have, it holds that

$$\mathcal{B}^m\mathcal{Y}(\mathcal{B}^H)^m\Sigma = f\left(C^{\varsigma_j}, R_u^{(j)}, \left\{\sigma_{v,k}^2\right\}_{k\in\mathcal{P}_j}, M_{\varsigma,j}\right), \quad (46)$$

where $f(\cdot)$ is some matrix-valued function. In other words, the selection matrix Σ defined above selects only some columns of the term $\mathcal{B}^m\mathcal{Y}(\mathcal{B}^H)^m$ with the non-zero blocks corresponding only to the terms related to estimating ς_j^o, and none related to $\varsigma_{j'}^o$ where $j\neq j'$.

To visualize this, let us consider a small example. Without loss of generality, assume that there are two nodes and two estimation tasks for each of them. Then, assume that all terms related to the first task are denoted by blue, while all terms related to the second task are denoted by red color. Also, all zero-blocks are white. Next, let us choose Σ so as to find MSD at node 1 for the second task. Then, due to the structure of the combination matrix and block-diagonal covariance matrix, the first three terms in the product $\mathcal{B}^m\mathcal{Y}(\mathcal{B}^H)^m\Sigma$ have similar chess-like structure. Finally, as depicted, the MSD at node 1 for the second task will have the terms that are related only to this task; in other words, the terms of other task have no influence on it. The tasks are decoupled.

Therefore, we conclude that, under assumptions A1-A4, for $k\in\mathcal{P}_j\cap\mathcal{P}_{j'}$, the estimation processes of ς_j^o and $\varsigma_{j'}^o$ can be considered uncoupled, i.e.,

$$\mathrm{MSD}_k(\mathcal{N}) \approx \sum_{j\in\mathcal{I}_k}\mathrm{MSD}_{k,\varsigma_j}(\mathcal{P}_j) \quad (47)$$

where $\forall j\in\{1,\ldots,J\}$ and for instance, for the Hastings combination rule, $\mathrm{MSD}_{k,\varsigma_j}(\mathcal{P}_j)$ is given by

$$\mathrm{MSD}_{k,\varsigma_j}(\mathcal{P}_j) \approx \begin{cases} \frac{\mu M_{\varsigma,j}}{2}\cdot\left(\sum_{k\in\mathcal{P}_j}\sigma_{v_k}^{-2}\right)^{-1} & \text{if } k\in\mathcal{P}_j,\\ 0 & \text{otherwise.}\end{cases} \quad (48)$$

This concludes the proof.

REFERENCES

[1] N. Bogdanović, D. Ampeliotis, and K. Berberidis, "Coalitional game theoretic approach to distributed adaptive parameter estimation," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015.

[2] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2365–2382, 2009.

[3] S. Kar and J. M. F. Moura, "Consensus + innovations distributed inference over networks: cooperation and sensing in networked systems," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 99–109, 2013.

[4] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, 2007.

[5] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1035–1048, 2010.

[6] S. Chouvardas, K. Slavakis, and S. Theodoridis, "Adaptive robust distributed learning in diffusion sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4692–4707, 2011.

[7] A. H. Sayed, S.-Y. Tu, J. Chen, X. Zhao, and Z. Towfic, "Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 155–171, May 2013.

[8] N. Bogdanović, J. Plata-Chaves, and K. Berberidis, "Distributed incremental-based LMS for node-specific parameter estimation over adaptive networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 5425–5429.

[9] ——, "Distributed diffusion-based LMS for node-specific parameter estimation over adaptive networks," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 7223–7227.

[10] ——, "Distributed incremental-based LMS for node-specific adaptive parameter estimation," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5382–5397, Oct 2014.

[11] J. Plata-Chaves, N. Bogdanović, and K. Berberidis, "Distributed diffusion-based LMS for node-specific adaptive parameter estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3448–3460, July 2015.

[12] J. Chen, C. Richard, and A. H. Sayed, "Multitask diffusion adaptation over networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4129–4144, Aug 2014.

[13] ——, "Diffusion LMS over multitask networks," *IEEE Transactions on Signal Processing*, vol. 63, no. 11, pp. 2733–2748, June 2015.

[14] O. Morgenstern and J. Von Neumann, "Theory of games and economic behavior," 1953.

[15] R. Myerson, *Game theory*. Harvard university press, 2013.

[16] Y. Shoham and K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.

[17] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 77–97, 2009.

[18] R. J. Aumann, "The core of a cooperative game without side payments," *Transactions of the American Mathematical Society*, pp. 539–552, 1961.

[19] B. Peleg and P. Sudhölter, *Introduction to the theory of cooperative games*. Springer Science & Business Media, 2007, vol. 34.

[20] C. Yu, M. van der Schaar, and A. Sayed, "Reputation design for adaptive networks with selfish agents," *2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 160–164, Jun. 2013.

[21] ——, "Cluster formation over adaptive networks with selfish agents," in *2013 Proceedings of the 21st European Signal Processing Conference (EUSIPCO)*, 2013, pp. 1–5.

[22] O. Gharehshiran, V. Krishnamurthy, and G. Yin, "Distributed energy-aware diffusion least mean squares: Game-theoretic learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 821–836, Oct 2013.

[23] C. Jiang, Y. Chen, and K. R. Liu, "Distributed adaptive networks: A graphical evolutionary game-theoretic view," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5675–5688, 2013.

[24] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "Coalitional games for distributed collaborative spectrum sensing in cognitive radio networks," in *IEEE INFOCOM 2009*, April 2009, pp. 2114–2122.

[25] H. Scarf, "The core of an N person game," *Econometrica: Journal of the Econometric Society*, pp. 50–69, 1967.

[26] L. Billera, "Some theorems on the core of an N-game without side-payments," *SIAM Journal on Applied Mathematics*, vol. 18, no. 3, pp. 567–579, 1970.

[27] S. Sen and P. S. Dutta, "Searching for optimal coalition structures," in *Fourth International Conference on MultiAgent Systems, Proceedings*. IEEE, 2000, pp. 287–292.

[28] A. Predtetchinski and P. J.-J. Herings, "A necessary and sufficient condition for non-emptiness of the core of a non-transferable utility game," *Journal of Economic Theory*, vol. 116, no. 1, pp. 84–92, 2004.

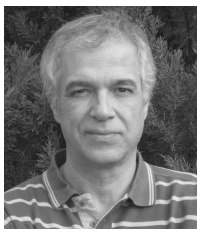
[29] L. Shapley and H. Scarf, "On cores and indivisibility," *Journal of mathematical economics*, vol. 1, no. 1, pp. 23–37, 1974.

[30] R. J. Aumann and J. H. Dreze, "Cooperative games with coalition structures," *International Journal of game theory*, vol. 3, no. 4, pp. 217–237, 1974.
[31] A. H. Sayed, *Adaptive filters*. Wiley-IEEE Press, 2011.
[32] P. J.-J. Herings, G. van der Laan, and T. Dolf, "Cooperative games in graph structure," Maastricht University, Maastricht Research School of Economics of Technology and Organization (METEOR), Research Memorandum 011, 2002. [Online]. Available: http://EconPapers.repec.org/RePEc:unm:umamet:2002011
[33] K. Apt and A. Witzel, "A generic approach to coalition formation," *International Game Theory Review*, vol. 11, no. 03, pp. 347–367, 2009.
[34] R. B. Myerson, "Graphs and cooperation in games," *Mathematics of operations research*, vol. 2, no. 3, pp. 225–229, 1977.
[35] E. Stai, S. Papavassiliou, and J. S. Baras, "A coalitional game based approach for multi-metric optimal routing in wireless networks," in *2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 1930–1934.
[36] T. Wang, L. Song, and Z. Han, "Coalitional graph games for popular content distribution in cognitive radio vanets," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 8, pp. 4010–4019, 2013.
[37] W. Saad, Z. Han, M. Debbah, and A. Hjørungnes, "A distributed coalition formation framework for fair user cooperation in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4580–4593, September 2009.
[38] W. Wang, B. Kasiri, J. Cai, and A. Alfa, "Distributed cooperative multi-channel spectrum sensing based on dynamic coalitional game," in *2010 IEEE Global Telecommunications Conference (GLOBECOM 2010)*, Dec 2010, pp. 1–5.
[39] L. Mashayekhy and D. Grosu, "A merge-and-split mechanism for dynamic virtual organization formation in grids," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 540–549, March 2014.

Nikola Bogdanović (S'10-M'16) was born in Niš, Serbia, in 1983. He received the B.Sc. and the M.Sc degree in Electrical and Computer Engineering from the Faculty of Electronic Engineering in Niš, Serbia. He obtained the Ph.D. degree in Signal Processing and Communications from the University of Patras, Greece, in 2014.

He was a Marie Curie Ph.D. fellow at the Computer Technology Institute (CTI) in Patras, Greece, where he was then working as a postdoctoral fellow. Afterwards, as a Marie Curie postdoc fellow, he joined the Microwave Sensing, Signals and Systems (MS3) group at TU Delft, the Netherlands, where is currently employed. His research interests include adaptive filtering, distributed signal processing with application to estimation and tracking, wireless sensor networks, game theory, and radar management.

Dimitris Ampeliotis (S'04-M'09) was born in Athens, Greece, in 1979. He received the Diploma degree in computer engineering and informatics, the MSC degree in signal processing, and the Ph.D. degree in signal processing and communications, all from the University of Patras, Greece, in 2002, 2004 and 2009, respectively. Since 2003, he has been a Researcher with the Computer Technology Institute (CTI) and Press "Diophantus", Patras, Greece. For the academic years 2010-14 he worked as an Adjunct Lecturer at the Technological Educational Institution of western Greece. During 2010-11 he also worked as an Adjunct Lecturer at the Computer Engineering and Informatics Department, University of Patras, Greece. His research interests include distributed algorithms for estimation problems (source localization, dictionary learning), cooperative game theory, turbo equalization, blind source separation and medical image processing. Dr. Ampeliotis is a member of the Technical Chamber of Greece, a member of EURASIP and a member of the IEEE.



Kostas Berberidis (S'87-M'90-SM'07) received the Diploma degree in electrical engineering from DUTH, Greece, in 1985, and the Ph.D. degree in signal processing and communications from the University of Patras, Greece, in 1990. During 1991, he worked at the Signal Processing Laboratory of the National Defense Research Center. From 1992 to 1994 and from 1996 to 1997, he was a researcher at the Computer Technology Institute (CTI), Patras, Greece. In period 1994/95 he was a Postdoctoral Fellow at CCETT/CNET, Rennes, France. Since December 1997, he has been with the Computer Engineering and Informatics Department (CEID), University of Patras, where he is currently a Professor, and Head of the Signal Processing and Communications Laboratory. Also, since 2008, he has been Director of the Signal Processing & Communications Research Unit of the Computer Technology Institute and Press "Diophantus". His research interests include adaptive signal processing, distributed processing and learning, signal processing for communications, and wireless sensor networks.

Prof. Berberidis has served or has been serving as a member of scientific and organizing committees of several international conferences, as Associate Editor for the IEEE Transactions on Signal Processing and the IEEE Signal Processing Letters, as a Guest Editor for the EURASIP JASP and as Associate Editor for the EURASIP Journal on Advances in Signal Processing. Since February 2010, he has been serving as Chair of the Greece Chapter of the IEEE Signal Processing Society. Also, currently, he is a member of the "Signal Processing Theory and Methods" Technical Committee of the IEEE Signal Processing Society and the "Signal Processing for Communications and Electronics" Technical Committee of the IEEE Communications Society. Moreover, since August 2015 he has been a member of the EURASIP Special Area Teams "Theoretical and Methodological Trends for Signal Processing" and "Signal Processing for Multisensor Systems". He is a Member of the Technical Chamber of Greece, a member of EURASIP, and a Senior Member of the IEEE. He has been elected a member of the Board of Directors of EURASIP, with his term starting January 1, 2017.