Enabling Data Exchange in Two-Agent Interactive Systems Under Privacy Constraints

Elena Veronica Belmega, Lalitha Sankar, H Vincent Poor

Enabling Data Exchange in Two-Agent Interactive Systems under Privacy Constraints

E. Veronica Belmega, *Member, IEEE,* Lalitha Sankar, *Member, IEEE,* and H. Vincent Poor, *Fellow, IEEE*

Abstract—It is advantageous for collecting agents in interconnected systems to exchange information (e.g., functions of their measurements) in order to improve their local processing (e.g., state estimation) because of the typically correlated nature of the data in such systems. However, privacy concerns may limit or prevent this exchange leading to a tradeoff between state estimation fidelity and privacy (referred to as *competitive privacy*). This paper focuses on a two-agent interactive setting and uses a communication protocol in which each agent is capable of sharing a compressed function of its data. The objective of this paper is to study centralized and decentralized mechanisms that can enable and sustain non-zero data exchanges among the agents. A centralized mechanism determines the data sharing policies that optimize a network-wide objective function combining the fidelities and leakages at both agents. Using common-goal games and best-response analysis, the optimal policies are derived analytically and allow a distributed implementation. In contrast, in the decentralized setting, repeated discounted games are shown to naturally enable data exchange (without any central control or economic incentives) resulting from the power to renege on a mutual data exchange agreement. For both approaches, it is shown that non-zero data exchanges can be sustained for specific fidelity ranges even when privacy is a limiting factor. This paper makes a first contribution to understanding how data exchange among distributed agents can be enabled under privacy concerns and the resulting tradeoffs in terms of leakage vs. estimation errors.

Index Terms—competitive privacy, distributed state estimation, non-cooperative games, discounted repeated games

I. Introduction

In many distributed systems, data collection and distributed processing are essential for system-wide reliable operations and control. Examples include critical infrastructures such as the electric power system and water management system, as well as the emerging distributed electronic medical record (EMR) system. In all these systems, while the operations are distributed, the data collected at the various agents are correlated and need to be shared for better local computation and estimation. For example, in the electric power system, data collection and processing are performed at various entities (e.g., utility companies, systems operators, etc.) that are spread out geographically. The interconnectedness of the network, requires that these distributed entities share data amongst themselves to ensure precise estimation and control, and in turn, system stability and reliability. Despite its importance, data sharing in the electric power system is often limited - sometimes with catastrophic consequences [2] - because of competitive interests or security/privacy concerns of the agents.

Although preserving the integrity of telemetry data in such distributed systems is important, this work does not deal with this security problem; rather, the focus is on the problem of enabling data exchange between legitimate agents while ensuring that some aspects of the data (identified as private) are revealed minimally. For example, in distributed cloud-based computing systems (e.g., EMR) as well as in the electric power system, computations have to be done locally and require sharing of public data by the different agents while ensuring that specific data, deemed as private (e.g., personally identifiable information about patients, state information about nodes internal to an agent's region in the electric power system, etc.), is not revealed. We henceforth refer to this problem as *competitive privacy* as in [3]. While one could envision a variety of distributed computations, in this paper, we formalize the problem theoretically as a distributed state estimation problem with privacy constraints. Specifically, we use game-theoretic tools to understand incentives needed to enable non-zero data sharing.

The notion of privacy is predominantly associated with the problem of ensuring that personal data about individuals, stored in a variety of databases or cloud servers, is not revealed. Quantifying the privacy of released data has captured considerable attention from the computer science and information theory research communities leading to two different rigorous frameworks: differential privacy introduced by Dwork et al. [4], [5]; and information-theoretic privacy developed in [6]. The first framework focuses on worst-case guarantees and ignores the statistics of the data, while the latter focuses on average guarantees and is cognizant of the input data statistics; their appropriateness depends on the application at hand. In the information era, however, privacy restrictions also appear in data exchange contexts as detailed above; such a competitive privacy setting was first studied via an information-theoretic framework in [3]. A detailed tutorial on various contexts in which privacy problems are encountered can be found in [7].

For distributed state estimation problems via data exchange, an information-theoretic competitive privacy framework holds the following advantages: (a) it takes into account the statis-

tical nature of the measurements and underlying state (e.g., complex voltage measurements in the grid that are often assumed to be Gaussian distributed); (b) it combines both compression and privacy in one analysis by developing rate and privacy optimal data sharing protocols; and (c) it quantifies privacy over all possible sequence of measurements and system states.

The competitive privacy information-theoretic framework introduced in [3] studies information exchange among two interconnected agents whose privacy concerns limit such data sharing, and thus, impacts the fidelity of the state estimate computed by the agents. The authors proposed a distributed source coding model to quantify the information-theoretic tradeoff between estimate fidelity (distortion via mean-square error), privacy (information leakage), and communication rate (data sharing rate). Every achievable distortion-leakage tuple at both agents represents a vector of opposing quantities. This is so because achieving *minimum distortion* for one agent implies *maximum leakage* for the other; similarly, *minimum leakage* for one agent implies *maximum distortion* for the other. A pertinent question follows: how to choose such a tradeoff in practice?

The objective of this work is to address this question via mechanisms that can enable and sustain specific distortion-leakage tuples in both centralized (a unique decision-maker) and decentralized (each agent has its own individual agenda) settings. Game theory is a mathematical toolbox for studying interactions among strategic agents and has established its value in a wide-variety of fields including wireless communications [8], [9]. While often applied in the non-cooperative and decentralized context, even in centralized settings, game theory can be valuable when devising efficient and distributed algorithms to compute the solution; in fact, these tools can be very useful for solving difficult, non-convex problems that arise in multi-agent models with multiple performance criteria (such as leakage vs. fidelity) as we present in the sequel.

Our first approach assumes a central controller that imposes the data sharing choices of the two agents (e.g., when electric utility companies share their data with a central systems operator). The network-wide objective function captures both the overall leakage of information and the total distortion of the estimates of the two agents via their weighted sum. To circumvent the non-convexity of this objective function, we exploit the parallel between *distributed optimization* problems and *potential games* [10]. The Nash equilibria of the resulting common-goal game are the intersection points of the best-response functions which turn out to be piece-wise affine. Moreover, using game theoretic tools we provide a distributed algorithm - the iterative best-response algorithm - that converges to an optimal solution. Our results show that the central controller can smoothly manipulate the distortion-leakage tradeoff between two extremes: both users share their data fully (*minimum distortion - maximum leakage*) or not at all (*maximum distortion - minimum leakage*). Specifically, not all information-theoretic tuples can arise as outcomes, but only the optimizers of the network-wide objective function.

If there is no central controller (e.g., when agents are two systems operators that need to share data to monitor large parts of the electric grid), each agent chooses its own data sharing strategy to optimize its individual distortion-leakage tradeoff. In [11], we showed that data sharing decreases the distortion of the agent receiving data while the sharing agent only increases its leakage. Thus, when the interaction takes place only once (i.e., one-shot interaction), rational agents have no incentive to share data. Economic incentives overcome this issue [11] and all distortion-leakage tuples can be achieved assuming that agents are paid (by a common moderator) for their information leakage.

In the second part of this paper, we show that pricing is not the only mechanism enabling cooperation. If the agents interact repeatedly over an indeterminate period, *tit-for-tat* type of strategies (i.e., an agent shares its data *as long as* the other agent does the same) turn out to be stable outcomes of the new game. We show that a whole sub-region of distortion-leakage tuples (in between the aforementioned extremes) is achieved without the need for a central authority; effectively, the agents build trust by exchanging data in the long term.

Preliminary results regarding the repeated interaction have been presented in [1]. We provide here a complete analysis and detailed proofs. Moreover, in this current version, we: (i) introduce different discount factors to model individual preferences for present vs. future rewards; (ii) give closed-form bounds on the discount factors; and (iii) illustrate further results.

The paper is organized as follows. In Section II, we introduce the system model and an overview of the most relevant information and game-theoretic concepts and prior results. The common goal non-cooperative game and its Nash equilibria are analyzed in Section III as a simpler alternative to a non-convex centralized problem. In Section IV in which we examine the repeated games framework and study its solutions and achievable distortion-leakage pairs. Numerical results that illustrate the analysis are also provided. We conclude in Section V.

II. SYSTEM MODEL

We consider a network composed of physically interconnected nodes as illustrated in Figure 1. We focus only on a pair of such nodes - called agents - which are capable of communicating and sharing some of their collected data. Each agent observes a sequence of n measurements from which it estimates a set of system parameters, henceforth referred to as states. The measurements at each agent are also affected by the states of the other agent. For the sake of simplicity, we consider a linear approximation model (e.g., model of voltages in the electric power network [12]). Denoting the state and measurement vectors at agent $j \in \{1, 2\}$ as X_j^n and Y_j^n, respectively, the linear model is:

$$
\begin{aligned}
Y_{1,k} &= X_{1,k} + \alpha_1 X_{2,k} + Z_{1,k} \\
Y_{2,k} &= \alpha_2 X_{1,k} + X_{2,k} + Z_{2,k},
\end{aligned}
\tag{1}
$$

where α_1, α_2 are positive parameters. The kth states, $X_{1,k}$ and $X_{2,k}$, for all k, are assumed to be independent and identically distributed (i.i.d.) zero-mean unit-variance Gaussian random variables and the additive zero-mean Gaussian noise variables,
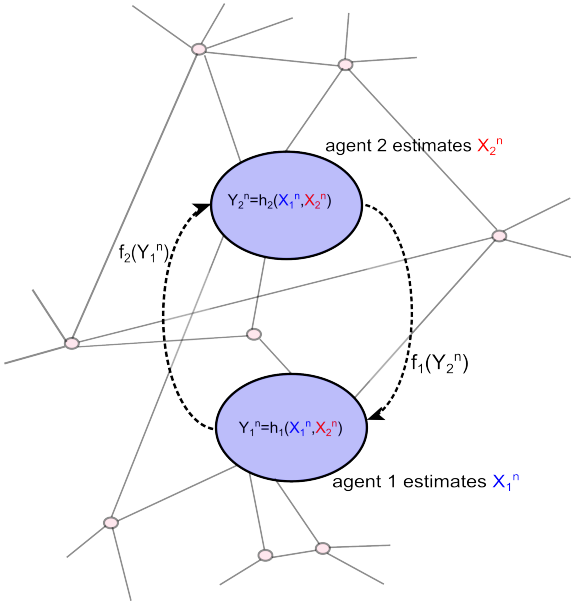
Fig. 1. Network of physically interconnected nodes. We focus on two communicating nodes/agents and exploit the possibility of exchanging information about their local measurements. These measurements depend on both agents' states (i.e., $Y_j^n = h_j(X_1^n, X_2^n)$) as made explicit in (1).

$Z_{1,k}$ and $Z_{2,k}$, are assumed to be independent of the agent states and of fixed variances σ_1^2 and σ_2^2, respectively.

This model is relevant to direct current (DC) state estimation problems in which the agents (e.g., system operators or energy management entities) need to share their local measurements (e.g., power flow and injections at specific locations) to estimate with high fidelity their local states (e.g., complex voltages).

Agent j can improve the fidelity of its state estimate if the other agent $i \neq j$ decides to share some information regarding its measurements - say $f_j(Y_i^n)$. At the same time, the amount of information agent i wishes to leak about its state is constrained (in the competitive privacy framework of [3]). These conflicting aspects are measured by information-theoretic concepts: the desired fidelity and privacy amount to meeting a distortion (mean-square error) and an information leakage constraint, respectively:

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n}\left(X_{j,k} - \hat{X}_{j,k}\right)^2\right] \leq D_j, \text{ and} \quad (2a)$$

$$\frac{1}{n}I(X_j^n; f_i(Y_j^n), Y_i^n) \leq L_j, \quad (2b)$$

where D_j represents the distortion of estimate $\hat{X}_j^n$ - which depends on the other agent's sharing policy, $f_j(Y_i^n)$ - from the actual state X_j^n, and L_j is the maximum information leakage. The mutual information in (2b) measures the average leakage of information per sample about the private state X_j^n of agent j to the other agent. The other agent can infer information about X_j^n from two sources: (i) its own measurements Y_i^n (1); and (ii) the data shared by agent j, i.e., $f_i\left(Y_j^n\right)$.

Sankar et al. [3] determined the entire region of achievable (D_1, L_1, D_2, L_2) tuples. The authors devised a particular coding scheme - based on quantization and binning techniques

- that satisfies the distortion constraints D_i and achieves the minimal leakage L_j constraint (for both agents). We summarize the resulting achievable distortion-leakage (DL) region in the following theorem.

Theorem 1. *[3] The distortion-leakage tradeoff for a two-agent competitive privacy problem (described above) is characterized by the set of all distortion pairs $(D_1, D_2) \in [D_{\min,1}, D_{\max,1}) \times [D_{\min,2}, D_{\max,2})$ and the information leakage functions, $L_i(D_j)$, $\forall i,j \in \{1,2\}$, $i \neq j$, defined by*

$$L_i(D_j) = \begin{cases} \frac{1}{2}\log\left(\frac{m_i^2}{m_i^2 D_{\min,i} + n_i^2(D_j - D_{\min,j})}\right) & \text{, if } D_j < D_{\max,j} \\ \\ \frac{1}{2}\log\left(\frac{V_j}{V_j - \alpha_j}\right) & \text{, otherwise} \end{cases}$$

(3)

with the following parameters:

$$V_j = 1 + \alpha_j^2 + \sigma_j^2,$$
$$E = \alpha_1 + \alpha_2,$$
$$n_j = \frac{V_i - \alpha_i E}{V_1 V_2 - E^2},$$
$$m_j = \frac{\alpha_j V_i - E}{V_1 V_2 - E^2},$$
$$D_{\min,j} = 1 - \frac{(\alpha_i^2 V_j + V_i - 2\alpha_i E)}{(V_1 V_2 - E^2)},$$
$$D_{\max,j} = 1 - \frac{1}{V_j}.$$

The four dimensional distortion-leakage tuples (D_1, L_1, D_2, L_2) can be fully characterized by all distortion pairs (D_1, D_2) and the corresponding leakage values $L_1(D_2)$ and $L_2(D_1)$ (via the one-to-one bijective functions above).

The maximal and minimal distortions, denoted by $D_{\max,j}$ and $D_{\min,j}$, represent the extreme cases in which the other agent i either sends no information or fully discloses its measurements. If $D_j < D_{\max,j}$, the distortion constraint is non-trivial and agent i has to leak information about its own state. The leakage is increasing with D_j. If $D_j \geq D_{\max,j}$, the distortion constraint is trivial, and agent i does not have to send any data. Its minimum leakage is not zero because agent j can still infer some private data (on agent i state) from its measurements Y_j^n.

Notice that the region contains asymmetric tuples in terms of data sharing. This results from the opposing distortion and leakage components that cannot be optimal simultaneously: *minimum distortion* at one agent corresponds to *maximum leakage* at the other, and *minimum leakage* of one agent corresponds to *maximum distortion* at the other. From this region (which is four dimensional) alone, it is not clear how to choose such a tradeoff tuple. In this paper, the main objective is to study different mechanisms that explain how specific tuples may arise in centralized and decentralized settings.

III. CENTRALIZED SOLUTION VIA COMMON GOAL GAMES

Our first approach is focused on centralized networks in which a central controller (e.g., in the electric power grid or cloud-computing systems) dictates the data-sharing policies of the two agents. The controller wishes to minimize both

the overall estimation fidelity and the information leakage. But, as discussed in the previous section, the two objectives are opposing and they cannot be optimized simultaneously; a network-wide compromise has to be made.

In multi-objective optimization problems, scalarization via the weighted sum of the different objectives is a common technique that provides good tradeoff tuples by solving a simpler scalar problem instead. In some cases (such as convex optimization problems), the tuples obtained by tuning the weights among the objectives are all optimal tradeoffs [13].

The network-wide objective function that captures the trade-off between overall estimation fidelity and leakage - by their weighted sum - can be written as follows:

$$u_{\text{sys}}(D_1, D_2) = -\sum_{j=1}^{2} L_j(D_i) + \frac{q}{2} \log \left(\sum_{j=1}^{2} \overline{D}_j \middle/ \sum_{j=1}^{2} D_j \right),$$

where the leakage of information $L_j(D_i)$ is given in (2b) and $q = \hat{w}/\tilde{w} > 0$ is the ratio of the weighting factors between the two terms.

For homogeneity reasons, the second term in the above equation should relate to logarithmic information measures. We propose to balance the information leakage (in bits/sample) with the overall shared information (also in bits/sample) which is inversely proportional to the distortion [14, Chap. 10]; as the distortions decrease, the information revealed per sample (or communication rate) increases.

The problem reduces to finding the distortion pairs (D_1, D_2) - characterizing the data-sharing policies of both users - that maximize the objective function (5). One can easily check that this function is not always concave on its domain. By using a distributed approach to find the solution, we can overcome this obstacle. Assume each agent controls its own data-sharing policy which impacts directly on the distortion at the other agent. The control parameter (or action) of agent j is denoted by $a_j = D_i$. The agents choices are driven by the same *common goal*, i.e., the network-wide objective function.

We further exploit the parallel between distributed optimization and *potential games* which has several advantages: (i) it allows solution of a non-convex problem in a simpler manner; (ii) it leads to an iterative and distributed procedure that converges to a locally optimal tradeoff tuple; and (iii) the central controller can manipulate this outcome by tuning a scalar parameter alone. The partial shift of intelligence, from the centralized controller towards the agents, paves the way to developing scalable data-sharing policies in more complex networks (with larger numbers of communicating agents).

We model the *common goal game* by $\mathcal{G}_{\text{sys}} = (\mathcal{P}, \{\mathcal{A}_j\}_{j \in \mathcal{P}}, u_{\text{sys}})$ in which $\mathcal{P} \triangleq \{1, 2\}$ designates the set of *players* (the two agents); and $\mathcal{A}_j$ is the set of actions that agent j can take. The payoff function of both players, $u_{\text{sys}} : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}$, is given by

$$u_{\text{sys}}(a_1, a_2) = -\sum_{j=1}^{2} L_j(a_j) + \frac{q}{2} \log \left(\sum_{j=1}^{2} \overline{D}_j \middle/ \sum_{j=1}^{2} a_j \right).$$

The utility function can be re-written using Theorem 1 as

$$u_{\text{sys}}(a_1, a_2) = \frac{1}{2} \log \left(\frac{(\gamma_1 a_1 + \delta_1)(\gamma_2 a_2 + \delta_2)}{(a_1 + a_2)^q} \right) + C_0, \quad (5)$$

where $\gamma_j = (n_j/m_j)^2$, $\delta_j = D_{\min,j} - \gamma_j D_{\min,i}$, and $C_0 = q/2 \, \log(\overline{D}_1 + \overline{D}_2)$. Without loss of generality, the additive constant C_0 and the multiplicative positive constant $1/2$ in the payoff function can be ignored in the following analysis of the Nash Equilibrium (NE) [15].

The non-cooperative game $\mathcal{G}_{\text{sys}}$ falls into a special class called *potential games* [10] that have many interesting properties. Their particularity lies in the existence of a global function - called the *potential function* - that captures the players' incentives to change their actions. In our case, the network-wide objective (5) represents precisely the potential function of the game. Monderer et al. [10] proved that every potential game has at least one NE solution[1]. Also, every local maximizer of the potential is an NE of the game. However, since the potential function is not concave [17], the game may have other NE points (e.g., certain saddle points of the potential function).

To completely characterize the set of all NEs, we study the best-response correspondence defined by:

$$BR_j : \mathcal{A}_i \to \mathcal{A}_j$$
$$\text{s.t. } BR_j(a_i) = \arg \sup_{b_j} u_{\text{sys}}(b_j, a_i),$$

$$BR : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathcal{A}_1 \times \mathcal{A}_2$$
$$\text{s.t. } BR(a_1, a_2) = (BR_1(a_2) \times BR_2(a_1)).$$

The best-response (BR) of agent j to an action a_i played by the other agent i - denoted by $BR_j(a_i)$ - is the optimal choice (payoff maximizing one) of agent j given the action of the other player. The best-response correspondence, $BR(\cdot, \cdot)$, represents the concatenation of both agents' BRs. The optimal action of agent j for fixed choices of the other agents might not be a singleton, hence the correspondence definition (a set-valued function).

Nash [18] showed that the fixed points of the BR correspondence are the NEs. In our case, the BR functions reduce to simply piecewise affine functions. Thus, the game $\mathcal{G}_{\text{sys}}$ can be described as a *Cournot duopoly* interaction [16] in which the set of NE points is completely characterized by intersection points of the best-response functions $BR_1(\cdot)$ and $BR_2(\cdot)$ [19]. Using game theoretic tools, we reduce the non-convex optimization problem to the analysis of intersection points of piecewise affine functions.

We further investigate a refined stability property of NEs, namely, their *asymptotic stability* [16]. This property is important when the game has multiple NEs. In such cases it seems *a priori* impossible to predict which particular NE will be the actual outcome. Nevertheless, if the players update their choices using the best-response dynamics - i.e., the agents sequentially choose their best-response actions to previously

---

[1]Nash equilibrium represents the natural solution concept in non-cooperative games [16] defined as a profile of actions (one action for each agent) that is stable to unilateral deviations. Intuitively, if the players are at an NE, no player has any incentive to deviate and switch its action unilaterally (otherwise, the deviator decreases its payoff value).

observed plays by the others [8] - the outcome of a Cournot duopoly can be predicted exactly, depending on the initial point. To be precise, the asymptotically stable NEs will be the attractors of these dynamics whereas the other NEs will not be observed generically (except when the initial point happens to be one of these NEs). For a more detailed discussion on Cournot duopoly the reader is referred to [16] or [19].

To compute the BRs, we analyze the first-order partial derivatives of the potential function. We distinguish different behaviors depending on the emphasis on either the leakage of information ($q \leq 1$) or estimation fidelity ($q > 1$).

A. Emphasis on the fidelity of state estimation ($q > 1$)

By developing the first-order partial derivatives of the potential function, the best-responses become

$$BR_j(a_i) = \begin{cases} F_j(a_i), & \text{if } D_{\min,i} < F_j(a_i) \leq \overline{D}_i, \\ \overline{D}_i, & \text{if } F_j(a_i) > \overline{D}_i, \\ D_{\min,i} & \text{otherwise,} \end{cases} \quad (6)$$

where $F_j(a_i) = a_i/(q-1) - q\delta_j/((q-1)\gamma_j)$ is an affine function of a_i with parameters $\gamma_j = (n_j/m_j)^2$ and $\delta_j = D_{\min,j} - \gamma_j D_{\min,i}$. The intersection point of the two affine functions $F_1(\cdot)$ and $F_2(\cdot)$ is

$$\begin{cases} a_1^* = \frac{q}{1-(q-1)^2}\left(\frac{\delta_1}{\gamma_1}(q-1) + \frac{\delta_2}{\gamma_2}\right) \\ a_2^* = \frac{q}{1-(q-1)^2}\left(\frac{\delta_2}{\gamma_2}(q-1) + \frac{\delta_1}{\gamma_1}\right). \end{cases} \quad (7)$$

The NE can be completely characterized by the intersection points of the two BR functions in the profile set, i.e., $\Delta \triangleq [D_{\min,2}, \overline{D}_2] \times [D_{\min,1}, \overline{D}_1]$. Noticing that the BRs are piecewise affine functions, the following result is obtained.

Theorem 2. *The game $\mathcal{G}_{\text{sys}}$ has generically a unique or three NEs assuming the central controller puts an emphasis on the overall state estimation fidelity, i.e., $q > 1$. In very specific cases (on the system parameters), the game may have an infinite number of NEs (when the affine functions $F_j(\cdot)$ are identical) or two NEs (when the intersection point (7) lies on the border of Δ).*

Intuitively, if the network parameters α_1 and α_2 are randomly drawn from a continuous distribution, the probability of having infinitely many or two NEs is zero. In general, depending on the relative slopes of the two BRs, the game has a unique NE (given by (7) provided it lies in Δ) or three NEs (one is (7) and the other two lie on the border of Δ). The details of the proof are given in Appendix A. In this case, the NEs of the common goal game are either network-wide optimal or saddle points of the central controller's objective function (also the potential function of the game). However, only the NEs that are optimizers of this objective function are asymptotically stable and can be observed as outcomes of best-response dynamics/algorithms.

B. Emphasis on the overall leakage of information ($q \leq 1$)

As opposed to the previous case, the BR of agent j is a piecewise constant function given as follows:

$$BR_j(a_i) = \begin{cases} \overline{D}_i, & \text{if } \mathbf{C_i^{[q=1]}} \text{ or } \mathbf{C_i^{[q<1]}} \\ D_{\min,i} & \text{otherwise,} \end{cases} \quad (8)$$

with the following conditions

$$\mathbf{C_i^{[q=1]}}: \quad q = 1 \text{ and } a_i > \frac{\delta_i}{\gamma_i},$$

$$\mathbf{C_i^{[q<1]}}: \quad q < 1 \text{ and } F_j(a_i) > \overline{D}_i \text{ or}$$
$$\left(D_{\min,i} < F_j(a_i) \leq \overline{D}_i \text{ and } u_{\text{sys}}(D_{\min,i}, a_i) \leq u_{\text{sys}}(\overline{D}_i, a_i)\right),$$

where $F_j(a_i)$ is defined in (6). The intersection points of such functions switching between the two extremes can only lie on the corner points of Δ.

Theorem 3. *The game $\mathcal{G}_{\text{sys}}$ has either a unique or two NEs assuming the central controller puts an emphasis on the leakage of information ($q \leq 1$). The NEs lie on the four corners of Δ, depending on the system parameters.*

When the game has two NEs, they are always given by the two symmetric extreme corners $(D_{\min,2}, D_{\min,1})$ (both users fully disclose their measurements) and $(\overline{D}_2, \overline{D}_1)$ (no cooperation). Otherwise, either of the four corners can be the outcome of the game, depending on the system parameters. Also, all NEs are asymptotically stable in this case. The proof is omitted as it is tedious and follows simply by analyzing the intersection of piecewise constant functions. In this case, the central controller cannot smoothly manipulate the outcome by tuning $q \in [0, 1]$ and only extreme distortion-leakage pairs are achieved. In the remainder of this section, we focus only on the case $q > 1$, in which the controller puts an emphasis on the estimation fidelity.

C. Numerical results

We assume the target distortions to be equal to the maximum distortions $\overline{D}_j = D_{\max,j}$, $j \in \{1, 2\}$. First, we consider the case in which a unique NE exists and $q > 2$. Fig. 2 illustrates the water-levels of the potential function and the BRs in Δ for the scenario $\alpha_1 = 0.5$, $\alpha_2 = 0.6$ and $\sigma_1^2 = \sigma_2^2 = 0.1$. The NE is the intersection point $(a_1^{\text{NE}}, a_2^{NE}) = (a_1^*, a_2^*) = (0.2559, 0.2542)$ and is asymptotically stable. Using a best-response iteration, the two agents converge always - from any initial point - to the optimal point. If a small perturbation occurs, using the same iterative BR dynamics, the agents will return to this point.

The case in which the game has three NEs is illustrated in Fig. 3 for the scenario $\alpha = 1$, $\alpha_2 = 10$, $\sigma_1^2 = \sigma_2^2 = 0.1$ and $q = 1.2$. The solutions are $(a_1^{\text{NE}}, a_2^{\text{NE}}) \in \{(D_{\min,2}, D_{\min,1}), (\overline{D}_2, \overline{D}_1), (a_1^*, a_2^*)\} \equiv \{(0.1107, 0.0023), (0.9901, 0.5238), (0.2031, 0.1906)\}$.

Analyzing the plot of the BR functions, we can observe that the intersection point (a_1^*, a_2^*) is not asymptotically stable: Assume that a small perturbation moves the agents away from this point. By iterating the best responses, the agents get further away and converge to one of the other NEs. The initial perturbation determines which of the two NEs that are asymptotically stable will be chosen.

Fig. 4 illustrates the NEs depending on the parameter $q \in [0, 100]$ tuned by the central controller. Both scenarios of Fig. 2 and 3 are considered.

By choosing small values of q, the central controller prefers large distortions and small leakage tuples; privacy is enforced

in the network. Larger values of q result in opposite tuples (small distortions and large leakage tuples); cooperation is enabled among selfish agents. In the case of three NEs, the discontinuity at $q = 1$ can be explained by the change in the BR functions; if $q > 1$ they are continuous and piecewise affine; if $q \leq 1$ they are discontinuous Heaviside-type of functions (as seen in Sec. III-B).

We also remark that not all information-theoretic distortion-leakage tuples are achieved at the NE. Only the local maximizers or saddle points of the overall network-wide payoff function are NEs and these tradeoff tuples depend on the system parameters. To achieve different tuples at the NE, other objective functions have to be considered (e.g., the sum of agents' individual payoff functions in (9)).
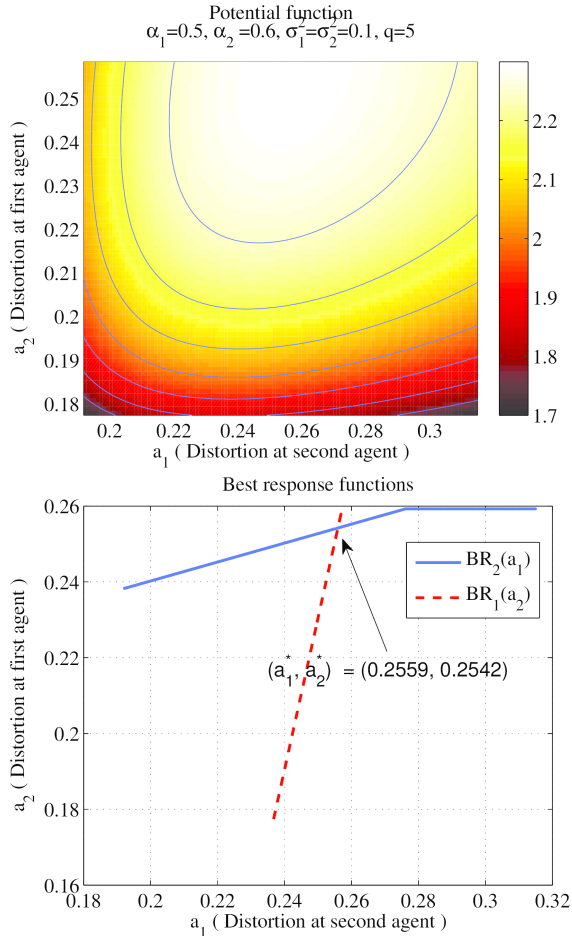

Fig. 2.   Water-levels of the potential (top) and BRs (bottom) as functions of $(a_1, a_2) \in (D_{\min,2}, \overline{D}_2] \times (D_{min,1}, \overline{D}_1]$. The potential has a unique maximum equal to (a_1^*, a_2^*) which is the asymptotically stable NE.

IV. DISCOUNTED REPEATED GAMES

In distributed networks, the need for continual monitoring makes repeated interactions among agents inevitable; e.g., the control of the electric power network depends on the state estimation performed periodically by distributed entities that interact with each other over and over. Such a repeated interaction may build trust among agents leading to sustained information exchange.
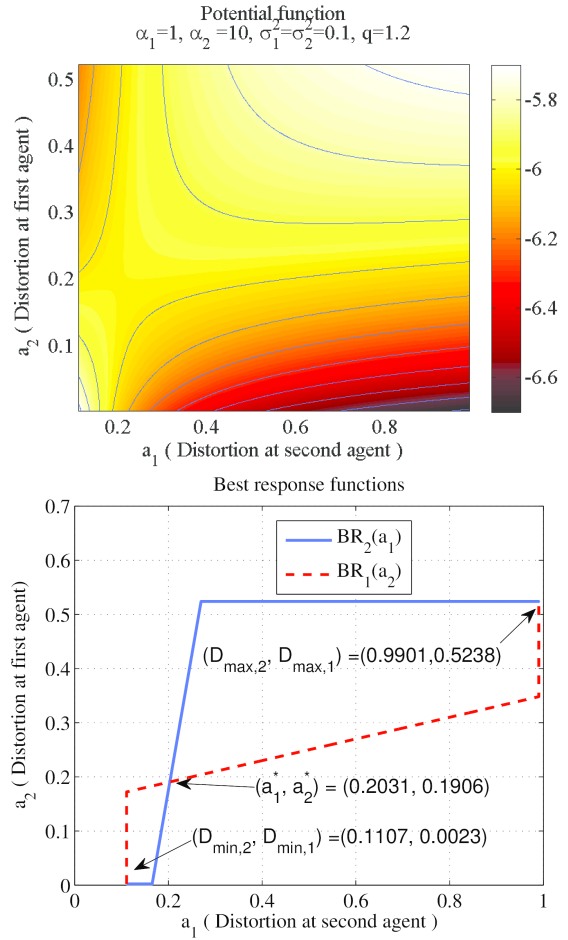

Fig. 3.   Water-levels of the potential (top) and BRs (bottom) as functions of $(a_1, a_2) \in (D_{\min,2}, \overline{D}_2] \times (D_{min,1}, \overline{D}_1]$. The potential has two local maxima $(D_{\min,2}, D_{\min,1})$, $(\overline{D}_2, \overline{D}_1)$ and one saddle point (a_1^*, a_2^*). The saddle point is a NE not asymptotically stable whereas the other two are asymptotically stable NE.

As opposed to the previous section, we do not assume the presence of a central controller. Rather, we exploit the repetition aspect to achieve distortion-leakage tuples beyond the no-data sharing case naturally without economic incentives.

One-shot game and pricing: We start with a brief overview of the non-cooperative game introduced in [11]. Consider the tuple $\mathcal{G} = (\mathcal{P}, \{\mathcal{A}_j\}_{j \in \mathcal{P}}, \{u_j\}_{j \in \mathcal{P}})$, where the set of players and their action sets are identical to the game described in Sec. III. The difference lies in the individual payoff functions: u_j, $\forall j \in \mathcal{P}$, the j-th of which measures the satisfaction of agent j and depends on agent j's own action choice but also on the others' choices. As opposed to the common-goal game, each agent cares only about its own leakage of information and state estimation fidelity. Thus, the payoff function of agent j, $u_j : \mathcal{A}_j \times \mathcal{A}_i \to \mathbb{R}$, is given by

$$u_j(a_j, a_i) = -L_j(a_j) + \frac{q_j}{2} \log \left(\frac{\overline{D}_j}{a_i} \right). \qquad (9)$$

The second term represents the information rate of the data received from the other agent depending on $a_i = D_j$, i.e., the distortion of agent j. The weight $q_j = \hat{w}_j / \tilde{w}_j$ is the ratio between the emphasis on leakage vs. state estimation fidelity of agent j.

Achievable distortion pairs at the NE by tuning q
$\alpha_1=0.5$, $\alpha_2=0.6$, $\sigma_1^2=\sigma_2^2=0.1$

Achievable distortion pairs at the NE by tuning q
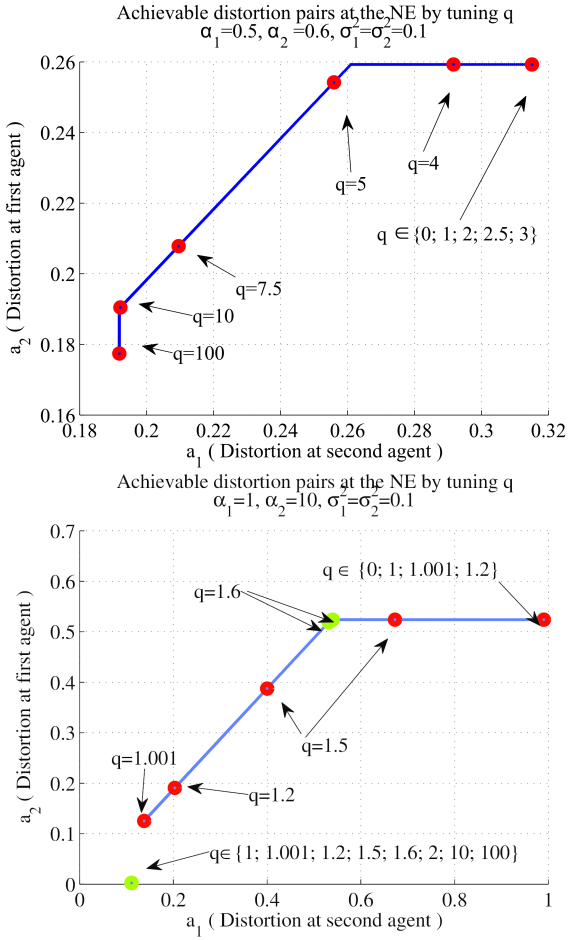$\alpha_1=1$, $\alpha_2=10$, $\sigma_1^2=\sigma_2^2=0.1$

Fig. 4. Achievable distortion pairs at the NE obtained by tuning $q \in [0, 100]$ for the scenarios of Fig. 2 and Fig. 3. Not all distortions pairs can be achieved at the NE of the common-goal game. The distortion pairs that are achieved are the points which correspond to either the local maxima or saddle points of the system-wide objective function.

Maximizing the utility in (9) with respect to (w.r.t.) a_j is equivalent to minimizing only the first term: the leakage of information. Indeed, the second term is a result of the data shared by the other agent i, and hence, not in control of agent j. The game simplifies into two simple decoupled optimization problems; each agent chooses to stay silent (minimizing its leakage of information). The only rational outcome is the *maximum distortion - minimum leakage* extreme for both agents.

Remark 1. *The one-shot game $\mathcal{G}$ is somewhat similar to the classical prisoners' dilemma [16] (which is a discrete game as opposed to our continuous game): each agent has a strictly dominant strategy[2] which is that of not sharing any data (beyond the minimum requirement).*

In [11], we show that any tuple in the information-theoretic region is achievable provided the agents are appropriately rewarded via pricing. The modified payoff functions that

---

[2]A strictly dominant strategy is an action that is the best choice of an agent independent from the others' choices.

include such pricing are, for $j \in \mathcal{P}$,

$$\tilde{u}_j(a_j, a_i) = u_j(a_j, a_i) + \frac{p_j}{2} \log\left(\frac{\overline{D}_i}{a_j}\right). \quad (10)$$

The drawback of such a pricing technique - which rewards an agent proportionality to its data sharing rate - is the implicit presence of a mediator (central controller or self-regulating market) which can manipulate the outcome by tuning the prices $p_j > 0$. In the following, we show that repetition enables cooperation among selfish agents - without any centralized interference.

We assume that the agents interact with each other multiple times under the same conditions, i.e., they play the same non-cooperative game $\mathcal{G}$ repeatedly. The total number of rounds is denoted by $T \geq 1$. Two cases are distinguished in function of the available knowledge of T: (i) perfect knowledge of T - both agents know in advance when their interaction ends; and (ii) imperfect or statistical knowledge of T - the agents do not know the precise ending of their interaction.

In both cases, we study the possibility of enabling and sustaining cooperation by allowing the agents to make only credible commitments, i.e., commitments on which they have incentives to follow through. The equilibrium concept we investigate here is a refinement of the Nash equilibrium, i.e., *subgame perfect equilibrium*, defined in the sequel.

A. Strategies, Payoffs and Subgame Perfect Equilibria

We introduce some useful notation and definitions. These tools are necessary for a clear understanding of the solutions arising in repeated games.

We assume that the game $\mathcal{G}$ described above is played several times. Repeated games differ from one-shot games by allowing players to observe the history of the game and condition their current play on past actions. The history at the end of stage $t \geq 1$ is denoted by $h^{(t+1)} = (a^{(1)}, \ldots, a^{(t)})$, where $a^{(\tau)} = (a_1^{(\tau)}, a_2^{(\tau)})$ represents the agents' play or action profile at stage τ. The set of all possible histories at the end of stage t is denoted by $\mathcal{H}^{(t+1)}$ such that $\mathcal{H}^{(1)}$ denotes the void set. We can now formally define a repeated game.

Definition 1. *A repeated game is a sequence of non-cooperative games given by the tuple $\mathcal{G}_R^{(T)} = (\mathcal{P}, \{\mathcal{S}_j\}_{j\in\mathcal{P}}, \{v_j\}_{j\in\mathcal{P}}, T)$, where $\mathcal{P} \triangleq \{1,2\}$ is the set of players (the two agents); $\mathcal{S}_j$ is the strategy set of agent j; and v_j is the payoff function which measures the satisfaction of agent j for any strategy profile.*

As opposed to the one-shot game, we have to make a clear distinction between an action - the choice of an agent at a specific moment (or stage of the game) - and a strategy that describes the agents' behavior for the whole duration of the game. A strategy of an agent is a contingent plan devising its play at each stage t and for any possible history $h^{(t)}$; more precisely it is defined as follows.

Definition 2. *A pure strategy for player j, s_j, is a sequence of causal functions $\{s_j^{(t)}\}_{1 \leq t \leq T}$ such that $s_j^{(t)} : \mathcal{H}^{(t)} \to [D_{\min,i}, \overline{D}_i]$, and $s_j^{(t)}(h^{(t)}) = a_j^{(t)} \in [D_{\min,i}, \overline{D}_i]$.*

The set of strategies, denoted by $\mathcal{S}_j$, is the set of all possible sequences of functions given in Definition 2, such that, at each stage of the game, every possible history of play $h^{(t)}$ is mapped into a specific action in $\mathcal{A}_j$ to be chosen at this stage.

In repeated games, the agents wish to maximize their averaged payoffs over the entire game horizon. We assume that agents discount future payoffs: present payoffs are more important than future promises.

Definition 3. *The discounted payoff function of player j given a joint strategy $s = (s_1, s_2)$ is given by*

$$v_j(s) = (1 - \rho_j) \sum_{t=1}^{T} \rho_j^{t-1} u_j(a^{(t)}), \tag{11}$$

where $a^{(t)}$ is the action profile induced by the joint strategy s, $u_j(\cdot)$ is the payoff function in (9), and $\rho_j \in (0, 1)$ is the discount factor of player j.

The Nash equilibrium concept for repeated games is defined similarly to that for one-shot games (any strategy profile that is stable to unilateral deviations). Some Nash equilibria of repeated games may rely on *empty threats* [16] of suboptimal play at histories that are not expected to occur (under the players' rationality assumption). Thus, we focus on a subset of Nash equilibria that allow players to make only commitments that they have incentives to follow through on: *the subgame perfect equilibria.*

Before defining this concept, we have to define subgames. Given any history $h^{(t)} \in \mathcal{H}^{(t)}$, the game from stage t onwards, is a subgame denoted by $\mathcal{G}_R(h^{(t)})$. The final history for this subgame is denoted by $h^{(T+1)} = (h^{(t)}, a^{(t)}, \ldots, a^{(T)})$. The strategies and payoffs are functions of the possible histories consistent with $h^{(t)}$. Any strategy profile s of the whole game induces a strategy $s|h^{(t)}$ on any subgame $\mathcal{G}_R(h^{(t)})$ such that for all j, $s_j|h^{(t)}$ is the restriction of s_j to the histories consistent with $h^{(t)}$.

Definition 4. *A subgame perfect equilibrium, $s^* = (s_1^*, s_2^*)$, is a strategy profile (in a repeated game with observed history) such that, for any stage and any history $h^{(t)} \in \mathcal{H}^{(t)}$, the restriction $s^*|h^{(t)}$ is a Nash equilibrium for the subgame $\mathcal{G}_R(h^{(t)})$.*

This equilibrium concept is a refinement of the NE because it is required to be an NE in every possible subgame aside from the entire history game. We analyze this solution concept for two different cases depending on the available knowledge of the end stage: perfect knowledge and imperfect or statistical knowledge of T.

B. Perfect knowledge of end stage

We assume first that the agents know in advance the value of T, i.e., when the game ends precisely. We show that data-sharing beyond the minimum requirement cannot be enabled in this case. This result follows as a corollary to Theorem 4.1 in [16, Chapter 4.2-4.3], i.e., the *one-stage-deviation principle* applied to the repeated prisoners' dilemma:

Proposition 1. *Assuming the agents know perfectly the value of T, the discounted repeated game $\mathcal{G}_R^{(T)}$ has a unique subgame perfect equilibrium s^* described by "no data sharing beyond the minimum requirement" at each stage of the game and for both agents:*

$$s_j^{(t),*} = \overline{D}_i, \ \forall t \in \{1, \ldots, T\}, \forall j \in \mathcal{P}. \tag{12}$$

The proof is omitted as it follows similarly to the repeated prisoners' dilemma (using an extension of the backward induction principle to dominance solvable games [16]). The key element is the strict dominance principle: a rational player will never choose an action that is strictly dominated. The same result remains true if the discounted payoffs are replaced with average payoffs, $v_j(s) = \frac{1}{T} \sum_{t=1}^{T} u_j(a^{(t)})$. Moreover, this result extends to a general class called *dynamic games* in which the system parameters ($\alpha_1^{(t)}, \alpha_2^{(t)}, \sigma_1^{(t)}, \sigma_2^{(t)}$) may vary at every stage of the game. The same reasoning holds since, at any stage of the game, the action corresponding to "no data sharing beyond the minimum requirement" is the strictly dominating one.

The only distortion-leakage tuple that can be achieved is the *maximum distortion-minimum leakage* one - similar to the one-shot game. The main reason why cooperation is not sustainable is that agents know precisely when their interaction ends. Next, we consider that the agents interact over an indeterminate period (they are unsure of the precise ending).

C. Imperfect knowledge of end stage

We now assume that the players do not know the value of T (the end stage). The discount factor ρ_j can be interpreted as the agent's belief (or probability) that the interaction goes on (see [20] and references therein). The probability that the game stops at stage t is then $(1 - \rho_j)\rho_j^{t-1}$. The discounted payoff (11) represents an expected or average utility. Thus, we assume that agent j know ρ_j which models its belief on the interaction continuing or not, at every stage (the probability that the game goes on).

The strategy of playing the one-shot NE at every stage is a subgame perfect equilibrium in this case as well.

Theorem 4. *Assuming imperfect knowledge of the end stage and that $D_{\min,j} > 0$ for all $j \in \mathcal{P}$, in the discounted repeated game $\mathcal{G}_R^{(\rho)} = (\mathcal{P}, \{\mathcal{S}_j\}_{j \in \mathcal{P}}, \{v_j\}_{j \in \mathcal{P}})$, the strategy "do not share any information beyond the minimum requirement" at each stage of the game and for both agents is a subgame perfect equilibrium, i.e.*

$$s_j^{(t),*} = \overline{D}_i, \ \forall t \geq 1, \forall j \in \mathcal{P}. \tag{13}$$

The details of the proof are reported in Appendix B. Unlike the case of perfect knowledge of T, we show that this is not the only possible outcome and other distortion-leakage pairs can be achieved.

Inspired from the repeated prisoners' dilemma, our objective is to show that non-zero exchange of information can be sustainable. Consider the action profiles $(D_2^*, D_1^*) \in [D_{\min,2}, \overline{D}_2) \times [D_{\min,1}, \overline{D}_1)$ which perform strictly better than
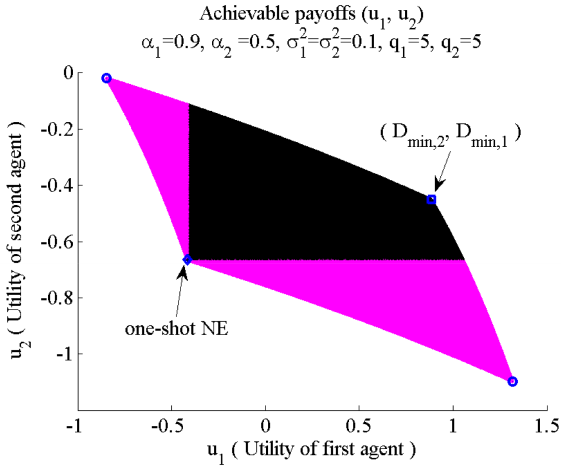
Fig. 5. The set of all payoff pairs. The low-left corner is the one-shot Nash equilibrium. The darker area is the subset of payoff pairs strictly better than the one-shot NE $q_1 = q_2 = 5$.

the one-shot NE for both agents:

$$\begin{cases} u_1(D_2^*, D_1^*) & > & u_1(\overline{D}_2, \overline{D}_1) \\ u_2(D_1^*, D_2^*) & > & u_2(\overline{D}_1, \overline{D}_2). \end{cases} \quad (14)$$

Such tuples may be expected to represent long term contracts or agreements between rational agents. Other tuples will never be acceptable: By not sharing any data, an agent is guaranteed at least the one-shot NE payoff value. In the game theoretic literature, these utility pairs are also known as *individually rational* payoffs [21].

These payoffs can be visualised in Fig. 5 for the scenario $\alpha_1 = 0.9$, $\alpha_2 = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.1$, $\overline{D}_j = D_{\min,j} + 0.5(D_{\max,j} - D_{\min,j})$, $q_1 = q_2 = 5$. The plotted area represents the set of all payoff pairs. The four corner points represent the four extremes: $(\overline{D}_2, \overline{D}_1)$ (the low-left corner: the one-shot NE), $(D_{\min,2}, \overline{D}_1)$ (the upper-left corner: the most advantageous for agent 2 - it shares nothing while agent 1 fully discloses its data), $(\overline{D}_2, D_{\min,1})$ (the low-right corner: the most advantageous for agent 1) and $(D_{\min,2}, D_{\min,1})$ (the upper-right corner: both agents fully disclose their data, maximizing their leakage). The darker area (in black) represents the subset of pairs satisfying (14). The lighter area (in magenta) represents the payoff pairs rejected by one or both rational players.

To gain more insight into these achievable agreement points, we write explicitly the payoff functions expressions in (9):

$$\begin{cases} u_j(\overline{D}_i, \overline{D}_j) & = & -L_j(\overline{D}_i) \\ u_j(D_i^*, D_j^*) & = & -L_j(D_i^*) + \frac{q_j}{2} \log\left(\frac{\overline{D}_j}{D_j^*}\right). \end{cases} \quad (15)$$

Data-sharing beyond the minimal requirement has two opposing effects: i) the leakage terms increase ($L_j(\overline{D}_i) < L_j(D_i^*)$, $\forall\ D_i^* < \overline{D}_i$); and ii) the estimation fidelity terms increase ($\log(\overline{D}_j/D_j^*) > 0$, $\forall\ D_j^* < \overline{D}_j$). Thus, the pairs (D_2^*, D_1^*) represent the tuples that result in an increase in the state estimation fidelity that overcomes the loss caused by the leakage for both agents.

Intuitively, the greater the emphasis on the state estimation terms, the larger the region of achievable agreement points is.

We also observe that the achievable distortion pairs satisfying the conditions in (14) must be relatively symmetric distortion pairs. Otherwise said, both agents have to share their data for the agreement to be acceptable to both parties.

Unlike the one-shot game or the determined horizon repeated game (the agents have perfect knowledge of T), the commitment of sharing data resulting in any distortion pair (D_2^*, D_1^*) is sustainable under some conditions on the discount factor. If the probability of the game stopping is small enough, then the commitment of playing (D_2^*, D_1^*) is credible and, thus, sustainable to rational agents.

Theorem 5. *Assuming imperfect knowledge of the end stage in the discounted repeated game $\mathcal{G}_R = (\mathcal{P}, \{\mathcal{S}_j\}_{j \in \mathcal{P}}, \{v_j\}_{j \in \mathcal{P}})$ and for any agreement profile $(D_2^*, D_1^*) \in [D_{\min,2}, \overline{D}_2) \times [D_{\min,1}, \overline{D}_1)$ that meets the conditions (14), if the discount factors are bounded as follows:*

$$1 > \rho_j > \frac{2[L_j(D_i^*) - L_j(\overline{D}_i)]}{q_j \log\left(\overline{D}_j/D_j^*\right)}, \quad (16)$$

and $D_{\min,j} > 0$ for all $j \in \mathcal{P}$, then the following strategy is a subgame perfect equilibrium: For all j, "agent j shares data at the agreement point D_i^* in the first stage and continues to share data at this agreement point if and as long as the other player i shares data at the agreement point D_j^*. If any player has ever defected from the agreement point, then the players do not cooperate beyond the minimum requirement from this stage on."

A detailed proof is given in Appendix C. This result assesses that both agents can achieve better distortion levels than the one-shot NE naturally, without the interference of a central authority or economic incentives. The optimal strategy is a *tit-for-tat* type of policy: *Each agent fulfils its part of the agreement and shares data if and as long as the other party does the same.* The *threat point* or punishment mechanism that guarantees the sustainability of such agreements is fully decentralized and involves stopping all data exchanges from the moment one of the agents defects.

Any distortion pair (D_i^*, D_j^*) in (14) is achievable *in the long term*, provided the discount factors are large enough. The lower bound in (16) depends on the agents' emphasis on leakage vs. fidelity. Larger emphasis on the leakage of information ($q_j \leq 1$) implies larger discount factors. Thus, smaller ending probability (or a longer expected interaction) is needed to sustain data sharing when agents are more sensitive to privacy concerns. This lower bound also depends on the specific agreement pair (D_2^*, D_1^*). It is again a compromise: Smaller distortion agreements imply larger leakages of information, thus, larger discount factors.

In conclusion, the minimum expected length of the interaction needed to sustain an agreement depends on the agents' tradeoffs between the leakage of information and state estimation fidelity resulting from their data exchange.

Theorem 5 may be extended to the case in which the parameters change at each stage of the game. However, the conditions on the discount factor would be much stricter. A different approach should be investigated in such general

dynamic games. This issue falls outside the scope of the present work and is left for future investigation.

D. Numerical results

We focus on the scenario $\alpha_1 = 0.9$, $\alpha_2 = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.1$ and $\overline{D}_j = D_{\max,j}$ for $j \in \mathcal{P}$. The minimum and maximum distortions are $D_{\min,1} = 0.3088$, $\overline{D}_1 = 0.3926$, $D_{\min,2} = 0.2183$ and $\overline{D}_2 = 0.2388$. For simplicity, we assume that both agents have the same belief on the end stage of the game, i.e., $\rho_1 = \rho_2 = \rho$.

If the agents put an emphasis on leakage (e.g., $q_1 = q_2 = 1$, $q_1 = 1, q_2 = 2$, or $q_1 = 2, q_2 = 1$), there is *no distortion pair* (D_2^*, D_1^*) that strictly improves both players' payoffs compared to the one-shot NE $(\overline{D}_2, \overline{D}_1)$. This means that the improvement in an agent's estimation fidelity from the data shared by the other agent is overcome by the loss of privacy incurred by the agreement point.

If the agents put more emphasis on their estimation fidelities, the region of agreements (D_2^*, D_1^*) becomes non-trivial. Figure 6 illustrates this region in the cases: i) $q_1 = 2$, $q_2 = 2$; ii) $q_1 = 1$, $q_2 = 5$; and iii) $q_1 = 5$, $q_2 = 5$. The colored regions represent all the possible agreements sustainable in the long term, whereas the white regions represent the distortion points that cannot be achieved. In each of these figures, the upper-right corner represents the minimum cooperation requirement $(\overline{D}_2, \overline{D}_1)$.

Very asymmetric distortion pairs (the upper-left and lower-right regions) are not achievable in the long term; a rational user will only agree to fulfil equitable data-sharing agreements. In other words, either both players share information at a non-trivial rate or neither of them does. The higher the emphasis on state estimation fidelity, the larger the agreement region and lower the distortion levels achieved: The minimal distortion pair $(D_{\min,2}, D_{\min,1})$ is only sustainable in the third case $(q_1 = q_2 = 5)$ when the emphasis on the estimation fidelity is high enough for both agents.

We can observe a symmetry regarding the values of ρ needed to sustain a given agreement pair. The fairer or more symmetric distortion pairs require a shorter expected game duration to be sustainable. The most unfair distortion pairs (the border points on the region of sustainable agreements) require the longest expected game duration; close to one probability of the game continuing. Beyond these edges, the difference between what an agent shares and what it receives in return is unacceptable, even in a long term interaction.

V. CONCLUSION AND FUTURE WORK

Data sharing among physically interconnected nodes/agents of a network improves their local state estimation. When privacy also plays a role, enabling non-zero data exchange often requires incentives.

In a centralized setting, we have shown that the central controller can manipulate the data sharing policies of the agents by tuning a single parameter - depending on the emphasis between leakage vs. estimation fidelity. A whole range of outcomes can be chosen in between two extremes: both agents fully disclose their measurements (*minimum distortion - maximum leakage*), and both agents stay silent (*maximum distortion - minimum leakage*).

If the network lacks a central controller and the agents are driven only by their individual agendas, we have proven that non-zero data sharing cannot be an outcome. Rational agents cannot trust each other in sharing data when the interaction takes place only once or in a finite number of rounds. However, if the agents interact repeatedly in the long term - over an undetermined number of rounds - then a whole region of outcomes is achieved depending on the agents' emphasis on leakage vs. state estimation fidelity. There is a symmetry in this achievable region as rational agents agree only on *tit-for-tat* data sharing policies. This result, i.e., long term repetition enables data exchange, follows from the underlying assumption that agents can perfectly observe the past plays (the history of the game) and condition their present choices on these observations. In practice, this implies important signalling among the agents which is an interesting topic for future work.

This work is a step towards understanding incentive mechanisms needed to enable data sharing between two distributed agents that otherwise would not cooperate because of privacy concerns. One possible generalization of this work is to find incentive mechanisms for the case of three or more connected agents. The primary limiting reason restricting this generalization is the lack of an information-theoretically optimal distortion-leakage data sharing protocol. The problem is still under investigation and the main challenges include finding a data sharing protocol that (a) allows an arbitrary ordering of the agents' communications; and (b) takes into account all prior communications at any agent in order to generate data to be shared with all other agents in the next round. These challenges lead to a combinatorial explosion in the number of possible data sharing schemes. However, it is possible to consider simple sharing schemes that are analogous to the two-user case (such as pairwise interactions) and develop incentive mechanisms for such schemes.

More generally, this work can also be tailored to specific applications, such as data sharing between local entities in the electric grid, to determine the appropriate incentive mechanism. On the one hand, data sharing between utilities is often monitored by central coordinating agencies (systems operators), and thus, the framework of common-goal games can be appropriate. On the other hand, data sharing between wide-area balancing authorities can be studied by tailoring the repeated games framework.

In conclusion, tools from classical and distributed optimization jointly with game theoretic tools are useful in a variety of distributed data sharing applications. This work illustrates one such application of distributed state estimation with privacy constraints. However, the incentive mechanisms developed here (for both centralized and decentralized networks) have a wide range of applications.
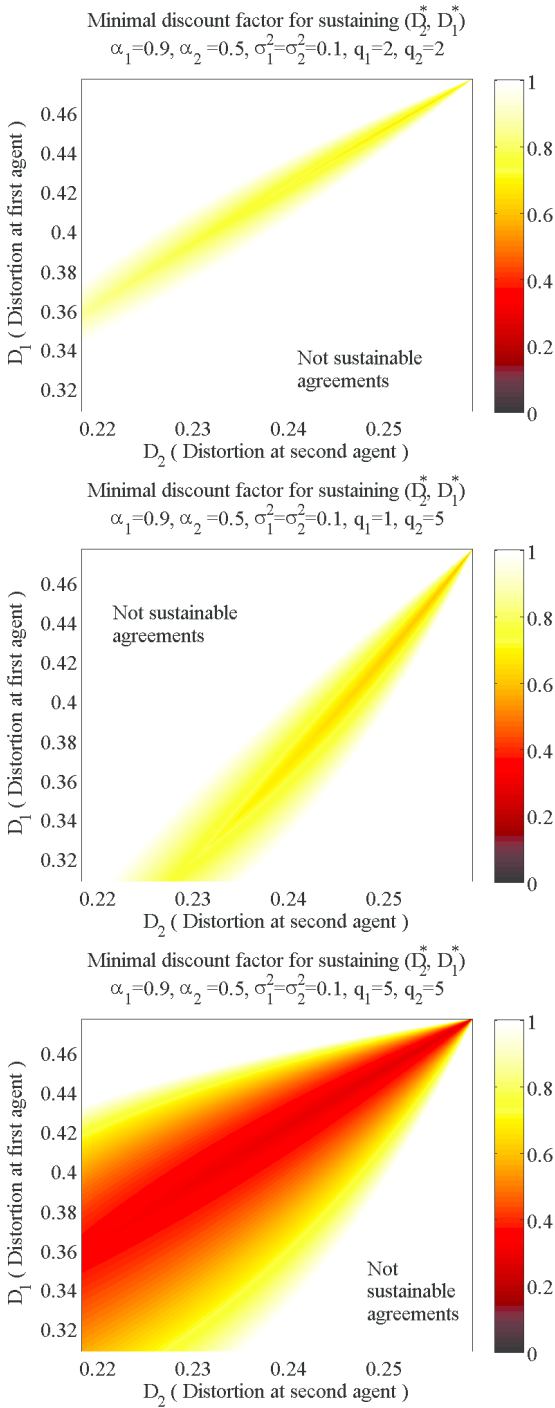
Fig. 6. The subset of all possible pairs (D_2^*, D_1^*) that are achieved and the minimal discount factor ρ needed to sustain them in the long-term interaction for the cases: i) $q_1 = 2, q_2 = 2$; ii) $q_1 = 1, q_2 = 5$; and iii) $q_1 = 5, q_2 = 5$.

APPENDIX A
PROOF OF THEOREM 2

Before providing the proof, we start by fully characterizing the set of NEs. Three cases are distinguished depending on the q parameter that determines the relative slopes of the BR functions.

1) If $q > 2$, then there is a unique and asymptotically stable NE. If the intersection point of the affine functions $F_1(\cdot)$

and $F_2(\cdot)$ denoted by (a_1^*, a_2^*) with

$$\begin{cases} a_1^* = \frac{q}{1-(q-1)^2}\left(\frac{\delta_1}{\gamma_1}(q-1) + \frac{\delta_2}{\gamma_2}\right) \\ a_2^* = \frac{q}{1-(q-1)^2}\left(\frac{\delta_2}{\gamma_2}(q-1) + \frac{\delta_1}{\gamma_1}\right). \end{cases} \quad (17)$$

lies in the interior of Δ, then it is the NE of the game. Otherwise, the NE lies on the border of Δ.

2) If $q = 2$, then we have two different situations. If the condition $\delta_1/\gamma_1 + \delta_2/\gamma_2 \neq 0$ holds, then there is a unique and asymptotically stable NE lying on the border of Δ. If on the contrary $\delta_1/\gamma_1 + \delta_2/\gamma_2 = 0$, then $F_i(a_j) \equiv F_j^{-1}(a_j)$. In this case, if this affine function intersects Δ non-trivially, then the game has an infinite number of NEs which are not asymptotically stable. Otherwise, the unique NE lies on the border and is asymptotically stable.

3) If $q < 2$, then there are two or three different NEs provided that the intersection point in (7) lies in the interior or on the border of Δ: this intersection point is the only asymptotically unstable equilibrium. The other one or two NEs lie on the corners of Δ, $(D_{\min,2}, D_{\min,1})$ and $(\overline{D}_2, \overline{D}_1)$). Otherwise, there is a unique NE which lies on the border of Δ and is asymptotically stable.

Intuitively, the scalar threshold equal to 2 for the parameter q comes from the relative order among the two slopes of the BR functions. If $q = 2$, then the two slopes are identical and equal to one. In any other case, the slopes of the two curves are different in the same axis system $\widetilde{a_1 O a_2}$ (since one of the two curves would have to be inverted). The relative slopes of the two curves greatly influence their intersection points and, thus, the set of NEs.

The proof follows an approach similar to that in [19] for the power allocation game over non-overlapping frequency bands in the interference relay channel and assuming a zero-delay scalar amplify-and-forward relaying protocol. We investigate the NEs of the game $\mathcal{G}_{\text{sys}}$ when $q > 1$ and their asymptotic stability. A necessary and sufficient condition that guarantees the asymptotic stability of a certain NE, say $(a_1^{\text{NE}}, a_2^{\text{NE}})$, is related to the relative slopes of the BRs [16]:

$$\left|\frac{\partial BR_1}{\partial a_2}(a_2)\frac{\partial BR_2}{\partial a_1}(a_1)\right| < 1 \quad (18)$$

for all (a_1, a_2) in an open neighbourhood of $(a_1^{\text{NE}}, a_2^{\text{NE}})$.

The analysis of NE is based on the analysis of intersection points of the two BR functions in (6).

First, we analyze all the possible cases in which the intersection points between the affine functions $F_1(\cdot)$ and $F_2(\cdot)$ are outside the interval Δ or on the two corners: $(D_{\min,2}, \overline{D}_1)$ or $(\overline{D}_2, D_{\min,1})$. In these cases, the NE is unique and it lies on the border of Δ. These cases correspond to: (i) $F_1(\overline{D}_1) \leq D_{\min,2}$ or $F_2(D_{\min,2}) \geq \overline{D}_1$, (ii) $F_1(D_{\min,1}) \geq \overline{D}_2$ or $F_2(\overline{D}_2) \leq D_{\min,1}$ and the corresponding analysis will not be reported here as it is tedious and similar to the following more interesting one. The more interesting case is when

$F_1(D_{\min,1}) < \overline{D}_2$, $F_1(\overline{D}_1) > D_{\min,2}$, $F_2(D_{\min,2}) < \overline{D}_1$ and $F_2(\overline{D}_2) > D_{\min,1}$. This means that, if the curves $F_1(\cdot)$ and $F_2(\cdot)$ intersect, the intersection point or points lie in Δ and are NEs of the game under study. We have again three sub-cases:

a) If $q = 2$: then the two functions $F_1(\cdot)$ and $F_2^{-1}(\cdot)$ have the same slope (equal to one) and thus they are parallel.

- If $\delta_1/\gamma_1 = -\delta_2/\gamma_2$, then the two functions are the same. All the points on these curves that intersect Δ are NEs of the game. Therefore, we have an infinite number of NEs. The asymptotic stability condition is not met because

$$\left| \frac{\partial BR_1}{\partial a_2}(a_2) \frac{\partial BR_2}{\partial a_1}(a_1) \right| = 1,$$

for all these NEs.

- If $\delta_1/\gamma_1 \neq -\delta_2/\gamma_2$, then the two BR function intersect on the border of Δ in a unique asymptotically stable point for which

$$\left| \frac{\partial BR_1}{\partial a_2}(a_2) \frac{\partial BR_2}{\partial a_1}(a_1) \right| = 0.$$

b) If $q > 2$: then the NE is unique and a detailed discussion follows depending on the signs of the following inequalities: $F_1(D_{\min,1}) \lesseqgtr D_{\min,2}$, $F_1(\overline{D}_1) \lesseqgtr \overline{D}_2$, $F_2(D_{\min,2}) \lesseqgtr D_{\min,1}$ and $F_2(\overline{D}_2) \lesseqgtr \overline{D}_1$ and also on the relative positions of the intersection points between the two $F_j(\cdot)$ functions and the border of Δ. We will detail only one of these cases.
If $F_1(D_{\min,1}) \geq D_{\min,2}$, $F_1(\overline{D}_1) \leq \overline{D}_2$, $F_2(D_{\min,2}) \geq D_{\min,1}$ and $F_2(\overline{D}_2) \leq \overline{D}_1$, then the two BR functions coincide on Δ with the two functions $F_j(\cdot)$. The unique NE is given by the intersection point (a_1^*, a_2^*) of $F_1(\cdot)$ and $F_2(\cdot)$ such that

$$\begin{aligned} a_1^* &= \frac{q}{1-(q-1)^2} \left(\frac{\delta_1}{\gamma_1}(q-1) + \frac{\delta_2}{\gamma_2} \right) \\ a_2^* &= \frac{q}{1-(q-1)^2} \left(\frac{\delta_2}{\gamma_2}(q-1) + \frac{\delta_1}{\gamma_1} \right). \end{aligned} \tag{19}$$

It is easy to see that

$$\left| \frac{\partial BR_1}{\partial a_2}(a_2) \frac{\partial BR_2}{\partial a_1}(a_1) \right| < 1$$

and, thus, the NE is asymptotically stable.

c) If $q < 2$: then the discussion follows similarly depending on the signs of the following inequalities: $F_1(D_{\min,1}) \lesseqgtr D_{\min,2}$, $F_1(\overline{D}_1) \lesseqgtr \overline{D}_2$, $F_2(D_{\min,2}) \lesseqgtr D_{\min,1}$ and $F_2(\overline{D}_2) \lesseqgtr \overline{D}_1$ and also on the relative intersection points between the $F_j(\cdot)$ functions with the border of Δ.

APPENDIX B
PROOF OF THEOREM 4

The backward induction argument is no longer valid since agents do not know which stage is the final one. Instead, we apply the *one-stage-deviation principle* for discounted repeated games that are uniformly bounded in each stage [16]. This principle states that a strategy profile $s^* = (s_1^*, s_2^*)$ is subgame perfect if and only if there is no player j and strategy $\hat{s}_j$ that agrees with s_j^* except at a single stage τ and history $h^{(\tau)}$, and such that $\hat{s}_j|h^{(\tau)}$ is a better response than $s_j^*|h^{(\tau)}$ in the subgame $\mathcal{G}_R^{(\rho)}(h^{(\tau)})$.

First, we have to check the uniform boundedness condition on the stage payoffs. Indeed, we can show that the stage payoffs in (9) are bounded as follows:

$$\begin{aligned} |u_j(a_j^{(t)}, a_i^{(t)})| &\leq L_j(a_j^{(t)}) + \frac{q_j}{2} \log \left(\overline{D}_j / a_i^{(t)} \right) \\ &\leq (1 + q_j) \frac{1}{2} \log \left(1/D_{\min,j} \right). \end{aligned}$$

Given that $D_{\min,j} < 1$, under the mild assumptions that $D_{\min,j} > 0$ and that q_j is finite, the stage payoffs are uniformly bounded.

Second, we have to check whether unilateral deviation in a single stage from the strategy in (13) can be profitable. If not, then the strategy is a subgame perfect equilibrium. Assume that player j deviates at time τ and history $h^{(\tau)}$ by choosing $\hat{s}_j^{(\tau)}(h^{(\tau)}) = \hat{D}_i \in (D_{\min,i}, \overline{D}_i)$ at stage τ. From then on, this strategy conforms to s^*, i.e., $\hat{s}_j^{(t)} \equiv s_j^{*,(t)}$, for all $t > \tau$. This means that the leakage of information of player j will increase at stage τ and therefore its payoff will decrease: $u_j(\hat{D}_i, \overline{D}_j) < u_j(\overline{D}_i, \overline{D}_j)$. This implies directly that $v_j(\hat{s}_j|h^{(\tau)}, s_i^*|h^{(\tau)}) < v_j(s_j^*|h^{(\tau)}, s_i^*|h^{(\tau)})$. Therefore, no agent has any interest in deviating at any single stage and the plan defined in (13) represents a subgame perfect equilibrium.

As opposed to the case in which the perfect knowledge of T is available, the discounted payoffs play a crucial role in the one-stage-deviation principle and, thus, this proof is not readily applicable in the case in which a uniform average of the stage-payoffs is considered. Also, similarly to the case of perfect knowledge of T, this result extends to a general class of dynamic games in which the system parameters change at every stage of the game.

APPENDIX C
PROOF OF THEOREM 5

We use the one-stage deviation principle similarly to the proof of Theorem 4. Assume that no agent deviates in any subgame from the agreement point. In this case, the discounted long-term payoff of player j is equal to $u_j(D_i^*, D_j^*)$, i.e., the instantaneous payoff achieved at the agreement point. If a player j deviates at stage τ by choosing $\hat{s}_j^{(\tau)} = D_i > D_i^*$ and then onwards conforms to the strategy by choosing D_i, its discounted payoff is

$$\begin{aligned} &(1-\rho_j) \sum_{t=0}^{\tau-2} \rho_j^t u_j(D_i^*, D_j^*) + (1-\rho_j)\rho_j^{\tau-1} u_j(D_i, D_j^*) + \\ &(1-\rho_j)\rho_j^\tau \sum_{t=0}^{+\infty} \rho^t u_j(\overline{D}_i, \overline{D}_j) = u_j(D_i^*, D_j^*) - \\ &\rho_j^{\tau-1} \left[u_j(D_i^*, D_j^*) - u_j(D_i, D_j^*) + \right. \\ &\left. \rho_j \left(u_j(D_i, D_j^*) - u_j(\overline{D}_i, \overline{D}_j) \right) \right]. \end{aligned} \tag{20}$$

Notice that $u_j(D_i, D_j^*) - u_j(D_i^*, D_j^*) > 0$ (by sharing more information, the leakage term for player j increases and its payoff decreases), and $u_j(D_i, D_j^*) - u_j(\overline{D}_i, \overline{D}_j) > 0$ (from the previous observation and condition (14)) for any $D_i > D_i^*$. Under the following sufficient condition on the discount factor:

$$1 > \rho_j > \max_{D_i \in (D_i^*, \overline{D}_i]} \left\{ \frac{u_j(D_i, D_j^*) - u_j(D_i^*, D_j^*)}{u_j(D_i, D_j^*) - u_j(\overline{D}_i, \overline{D}_j)} \right\}, \tag{21}$$

this discounted payoff for the deviator in (20) is less than the payoff of no deviation $u_j(D_i^*, D_j^*)$.

Now, let us assume that a deviation has occurred. At stage τ and for any history $h^{(\tau)}$ after this deviation, if player j were to deviate from the prescribed strategy and choose $\hat{s}_j^{(\tau)} = D_i < \overline{D}_i$ and then conform from this stage onwards, its leakage term would increase and its payoff in stage τ would be strictly less than if it had not deviated. Thus, no player has any incentive to deviate at any single stage of the game and for any history of play and the strategy described in this theorem is a subgame perfect equilibrium for the discounted repeated game in which the end stage of the game is not known.

To complete the proof, we have to show that the sufficient condition in (21) is equivalent to the one in (16). First, from (3) we observe that the leakage function $L_j(D_i)$ is strictly decreasing with D_i (the smaller distortion at agent i the bigger the leakage term), and, thus, we have $\frac{\mathrm{d}L_j}{\mathrm{d}D_i} < 0$. Second, by replacing the payoff functions expressions in (9) we have

$$\frac{u_j(D_i, D_j^*) - u_j(D_i^*, D_j^*)}{u_j(D_i, D_j^*) - u_j(\overline{D}_i, \overline{D}_j)} = \frac{[L_j(D_i^*) - L_j(D_i)]}{[L_j(\overline{D}_i) - L_j(D_i)] + \frac{q_j}{2}\log\left(\overline{D}_j/D_j^*\right)}.$$

We compute the derivative of the right-side term w.r.t. D_i and obtain

$$\frac{\mathrm{d}}{\mathrm{d}D_i} \frac{L_j(D_i^*) - L_j(D_i)}{L_j(\overline{D}_i) - L_j(D_i) + \frac{q_j}{2}\log\left(\frac{\overline{D}_j}{D_j^*}\right)} =$$
$$-\frac{L_j(\overline{D}_i) - L_j(D_i^*) + \frac{q_j}{2}\log\left(\frac{\overline{D}_j}{D_j^*}\right)}{\left[L_j(\overline{D}_i) - L_j(D_i)] + \frac{q_j}{2}\log\left(\frac{\overline{D}_j}{D_j^*}\right)\right]^2} \left(\frac{\mathrm{d}L_j}{\mathrm{d}D_i}\right). \quad (22)$$

From the fact that $\frac{\mathrm{d}L_j}{\mathrm{d}D_i} < 0$, equations (14) and (15), we obtain that the derivative in (22) is strictly positive, and, thus

$$\max_{D_i \in (D_i^*, \overline{D}_i]} \left\{ \frac{u_j(D_i, D_j^*) - u_j(D_i^*, D_j^*)}{u_j(D_i, D_j^*) - u_j(\overline{D}_i, \overline{D}_j)} \right\} = 2\frac{L_j(D_i^*) - L_j(\overline{D}_i)}{q_j \log\left(\frac{\overline{D}_j}{D_j^*}\right)}.$$

REFERENCES

[1] E. V. Belmega, L. Sankar, and H. V. Poor, "Repeated games for privacy-aware distributed state estimation in interconnected networks," in *Proc. IEEE Intl. Conf. on Network Games, Control and Optimization (NETGCOOP)*, Avignon, France, Nov. 2012.

[2] *Mandatory Reliability Standards for Interconnection Reliability Operating Limits*, Federal Energy Regulatory Commission, Mar. 2011.

[3] L. Sankar, S. K. Kar, R. Tandon, and H. V. Poor, "Competitive privacy in the smart grid: An information-theoretic approach," in *Proc. IEEE Intl. Conf. Smart Grid Communications*, Brusells, Belgium, Oct. 2011.

[4] C. Dwork, "Differential privacy," in *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*. Springer, 2006, pp. 1–12.

[5] ——, "Differential privacy: A survey of results," in *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation*, ser. TAMC'08. Springer-Verlag, 2008, pp. 1–19.

[6] L. Sankar, S. R. Rajagopalan, and H. V. Poor, "Utility-privacy tradeoffs in databases: An information-theoretic approach," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, pp. 838–852, Jun. 2013.

[7] L. Sankar, W. Trappe, K. Ramchandran, H. V. Poor, and M. Debbah, "The role of signal processing in meeting privacy challenges: An overview," *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 95–106, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6582707

[8] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks: Fundamentals and Applications*. Academic Press, 2011.

[9] Z. Han, D. Niyato, W. Saad, T. Basar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press, 2012.

[10] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, pp. 124–143, 1996.

[11] E. V. Belmega, L. Sankar, H. V. Poor, and M. Debbah, "Pricing mechanisms for cooperative state estimation," in *Proc. International Symposium on Communications, Control and Signal Processing (IS-CCSP)*, Rome, Italy, May 2012, pp. 1–4.

[12] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York: CRC Press, 2004.

[13] S. Boyd and L. Vandenberghe, "Convex optimization," *Cambridge University Press*, 2004.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley Interscience, 2006.

[15] J. Weibull, *Evolutionary Game Theory*. MIT Press, 1997.

[16] D. Fudenberg and J. Tirole, *Game Theory*. The MIT Press, 1991.

[17] A. Neyman, "Correlated equilibrium and potential games," *Int. Journal of Game Theory*, vol. 26, pp. 223–227, 1997.

[18] J. F. Nash, "Equilibrium points in n-points games," *Proc. of the Nat. Academy of Sciences*, vol. 36, no. 1, pp. 48–49, Jan. 1950.

[19] E. V. Belmega, B. Djeumou, and S. Lasaulce, "Power allocation games in interference relay channels: Existence analysis of Nash equilibria," *EURASIP Journal on Wireless Communications and Networking (JWCN)*, pp. 1–20, Nov. 2010.

[20] M. L. Treust and S. Lasaulce, "A repeated game formulation of energy-efficient decentralized power control," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 2860–2869, Sep. 2010.

[21] R. Aumann and S. Hart, Eds., *Handbook of Game Theory with Economic Applications*, ser. Handbook of Game Theory with Economic Applications. Elsevier, 1992, vol. 1.

E. Veronica Belmega received the engineer diploma (M.Sc.) from the University Politehnica of Bucharest, Romania, in 2007. She obtained the M.Sc. and Ph.D. degrees, both from the Université Paris-Sud 11, Orsay, France, in 2007 and 2010 respectively. In 2010-2011, she was a post-doctoral researcher in a joint project between the Alcatel-Lucent Chair on Flexible Radio in Supélec and Princeton University. She is currently an assistant professor at ETIS/ENSEA - UCP - CNRS, Cergy-Pontoise, France. She was one of the ten recipients of the L'Oréal - UNESCO - French Academy of Science Fellowship: "For young women doctoral candidates in science" in 2009. Her interests lie in resource allocation in wireless communications, state estimation in electrical power grid using tools from game theory and distributed optimization.

Lalitha Sankar received the B.Tech degree from the Indian Institute of Technology, Bombay, the M.S. degree from the University of Maryland, and the Ph.D degree from Rutgers University. She is presently an Assistant Professor in the Department of Electrical, Computer, and Energy Engineering at Arizona State University. Prior to this, she was an Associate Research Scholar at Princeton University. Following her doctorate, Dr Sankar was a recipient of a three year Science and Technology Teaching Postdoctoral Fellowship from the Council on Science and Technology at Princeton University. Prior to her doctoral studies, she was a Senior Member of Technical Staff at AT&T Shannon Laboratories. Her research interests include cyber-security and information privacy in distributed and cyber-physical systems, especially the smart grid, privacy of electronic data, network information theory and its applications to data privacy. For her doctoral work, she received the 2007–2008 Electrical Engineering Academic Achievement Award from Rutgers University. She received the IEEE Globecom 2011 Best Paper Award for her work on privacy of side-information in multi-user data systems. Dr. Sankar received the 2014 National Science Foundation CAREER award.

H. Vincent Poor (S'72, M'77, SM'82, F'87) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Michael Henry Strater University Professor of Electrical Engineering and Dean of the School of Engineering and Applied Science. Dr. Poor's research interests are in the areas of stochastic analysis, statistical signal processing, and information theory, and their applications in wireless networks and related fields such as social networks and smart grid. Among his publications in these areas is the recent book *Mechanisms and Games for Dynamic Spectrum Allocation* (Cambridge University Press, 2014).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and a foreign member of Academia Europaea and the Royal Society. He is also a fellow of the American Academy of Arts and Sciences, the Royal Academy of Engineering (U. K.), and the Royal Society of Edinburgh. He received the Technical Achievement and Society Awards of the IEEE Signal Processing Society in 2007 and 2011, respectively. Recent recognition of his work includes the 2014 URSI Booker Gold Medal, and honorary doctorates from several universities in Europe and Asia, including an honorary D.Sc. from Aalto University in 2014.