

Repeated games for privacy-aware distributed state estimation in interconnected networks

Elena Veronica Belmega, Lalitha Sankar, H. Vincent Poor

▶ **To cite this version:**

Elena Veronica Belmega, Lalitha Sankar, H. Vincent Poor. Repeated games for privacy-aware distributed state estimation in interconnected networks. 6th International Conf. on NETwork Games, COntrol and OPtimization (NETGCOOP 2012), Nov 2012, Avignon, France. pp.64-68. hal-01309006

HAL Id: hal-01309006
https://hal.science/hal-01309006

Submitted on 28 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Repeated games for privacy-aware distributed state estimation in interconnected networks

E. V. Belmega[†], L. Sankar[*], and H. V. Poor[‡]
[†] ETIS / ENSEA - Université Cergy-Pontoise - CNRS, Cergy-Pontoise, France
[*] Arizona State University, Tempe, AZ 85281
[‡] Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544
Email: belmega@ensea.fr, lalithasankar@asu.edu, poor@princeton.edu

***Abstract*—The conflict between cooperation in distributed state estimation and the resulting leakage of private state information (competitive privacy) is studied for a system composed of two interconnected agents. The distributed state estimation problem is studied using an information theoretic rate-distortion-leakage tradeoff model and a repeated non-cooperative game framework. The objective is to investigate the conditions under which the repetition of the agents' interaction enables data sharing among the agents beyond the minimum requirement. In the finite horizon case, similarly to the one-shot interaction, data sharing beyond the minimum requirement is not a credible commitment for either of the agents. However, non-trivial mutual data sharing is sustainable in the long term, i.e., in the infinite horizon case.**

***Index Terms*—Competitive privacy, rate-distortion-leakage tradeoff, subgame perfect equilibrium**

I. Introduction

Classical state estimation requires a central coordinator that estimates the state using measurements from all the nodes in the network. Several distributed approaches to this problem have been considered [1] [2]. However, these approaches are essentially based on two-tier hierarchical models in which each local node estimates independently, and then, at a higher level, a central coordinator receives the estimation results from the individual areas and coordinates them to obtain a system-wide solution. However, this model does not scale with increasing measurement rates due to communication and reliability challenges inherent in systems with one coordination center. This requirement of estimating the state precisely and often is driving the need for a *fully distributed approach* wherein the local control centers *interactively* estimate the system state as a whole.

In the context of the electrical power-grid and the state-estimation problem at the regional transmission organisation (RTO) level, the authors of [3] introduce a *competitive privacy* notion to capture the conflict between sharing data for distributed estimation (utility/benefit to all RTOs) and withholding data for economic and end-user privacy reasons. Assuming that the RTOs are willing to cooperate, they present an information-theoretic rate-distortion-leakage framework for a two-RTO network and show that each RTO has to tolerate a level of information leakage proportional to the fidelity (distortion) required at the other RTO. These concepts are very general and can be applied to any interconnected system composed of several autonomous and possibly competing agents. A natural question that arises is whether such a cooperation can be achieved via incentives.

In [4], the authors address this question by introducing a pricing-based objective function. This objective function is essentially a linear combination between the leakage of information caused by sharing data and a pricing function that awards the agent proportionally to the fidelity improvement at the other agent's side. We show that strictly non-zero pricing incentives are required to achieve non-trivial target distortions. Furthermore, all distortion levels in the rate-distortion-leakage achievable tradeoff region can be achieved by appropriate pricing offered to balance out the negative effect of information leakage.

However, in several scenarios, this pricing technique may appear artificial and not realistic (e.g., in the absence of a centralized regulator that pays the agents in exchange for their sharing of data, in a self-regulatory market, etc.). In this paper, we investigate whether better distortion levels can be achieved as a result of repeated interactions among selfish agents using a non-cooperative repeated games framework. Two different cases are investigated: finite and infinite horizon repeated games. If the agents know exactly when the last stage of their interaction takes place, non-trivial data exchange cannot be sustained and the same distortion pair as the one-stage game is achieved. If the agents are unsure of the ending of their interaction, non-trivial data exchange can be sustainable in the long term. However, it turns out that not all the achievable distortion levels can be achieved as a result of the long-term interaction. More specifically, we have observed that very asymmetric distortion pairs cannot be sustainable. This means that, in the long term interaction, one agent has incentive to share data beyond the minimum requirement if and only if the other agent is doing the same thing.

II. System Model

The system under investigation is basically the one given in [3], i.e., we assume a linearized noisy Gaussian measurement

This work was supported in part by the U. S. National Science Foundation under Grant CCF-1016671.

model at the agents' level given as follows:

$$\begin{aligned} Y_1 &= X_1 + \alpha X_2 + Z_1 \\ Y_2 &= \beta X_1 + X_2 + Z_2, \end{aligned} \tag{1}$$

where $\alpha > 0$ and $\beta > 0$. The states, X_1 and X_2, are assumed to be independent standard (zero-mean unit-variance) Gaussian random variables and the additive zero-mean Gaussian noises, Z_1 and Z_2, are assumed to be independent each other and of the agent states and to have fixed variances σ_1^2 and σ_2^2. Agent j, $j \in \{1, 2\}$, encodes (quantizes) its measurement vector at a rate (in bits per measurement) R_j. The encoding is subject to the satisfaction of two constraints: a distortion requirement on the state estimate at the other agent (cooperation) and a leakage requirement (privacy) on the information leaked about its state to the other agent. Let D_j and L_j denote respectively the fidelity (distortion) and leakage requirements of the estimate at agent j where D_j is the mean square error between the original and reconstructed state vectors and L_j is the average mutual information between the state vector at agent j and the revealed data and measurement vector at the other agent.

We summarize the resulting rate-distortion-leakage (RDL) achievable region in the following theorem.

Theorem 1: [3] The rate-distortion-leakage tradeoff is the set of all $(R_1, R_2, D_1, D_2, L_1, L_2)$ tuples that satisfy the following: for agent j, $j \in \{1, 2\}$, and $i \in \{1, 2\} \setminus \{j\}$,

- $D_i < D_{\max,i}$*:*

$$R_j \geq \frac{1}{2} \log \left(\frac{c_j m_j^2}{D_i - D_{\min,i}} \right), \text{ and} \tag{2}$$

$$L_j \geq \frac{1}{2} \log \left(\frac{m_j^2}{m_j^2 D_{\min,j} + n_j^2 (D_i - D_{\min,i})} \right); \tag{3}$$

- $D_i \geq D_{\max,i}$*:* $R_j = 0$ *and* $L_j = \log\left(V_i/(V_i - q_j)\right)/2$,

where $q_1 = \beta$, $q_2 = \alpha$, $V_1 = 1 + \alpha^2 + \sigma_1^2$, $V_2 = 1 + \beta^2 + \sigma_2^2$, $E = \alpha + \beta$, $c_j = \frac{V_1 V_2 - E^2}{V_i}$, $n_1 = \frac{V_2 - \beta E}{V_1 V_2 - E^2}$, $n_2 = \frac{V_1 - \alpha E}{V_1 V_2 - E^2}$, $m_1 = \frac{\alpha V_2 - E}{V_1 V_2 - E^2}$, $m_2 = \frac{\beta V_1 - E}{V_1 V_2 - E^2}$ *and*

$$D_{\min,1} = 1 - \frac{(\beta^2 V_1 + V_2 - 2\beta E)}{(V_1 V_2 - E^2)}, \tag{4a}$$

$$D_{\min,2} = 1 - \frac{(V_1 + \alpha^2 V_2 - 2\alpha E)}{(V_1 V_2 - E^2)}, \tag{4b}$$

$$D_{\max,j} = 1 - \frac{1}{V_j}. \tag{4c}$$

In what follows, we focus on the lower-bound of this achievable region where only the minimal encoding rate R_j required to achieve the distortion level D_i is considered, i.e. the inequalities in (2) and (3) are satisfied with equality.

In [4], the authors consider the different encoding rates in the RDL tradeoff region as the actions that agents can take. Here, we will use a formulation that makes the analysis clearer and easier. Given that the encoding rate R_j is a bijection with respect to (w.r.t.) the distortion at the other agent D_i, we assume that the action agent j can take is $a_j \equiv D_i \in \mathcal{A}_j \triangleq (D_{\min,i}; \overline{D}_i]$. This simply means that the decision of agent j on *how much information to divulge to the other agent* directly tunes the state estimation fidelity of the other agent i. The term $\overline{D}_i$ denotes the maximum allowed distortion level of the state estimate of agent i, which can be equal to or less than $D_{\max,i}$ in Theorem 1.

Remark: It is an abuse of notation not to distinguish between the limits $a_j \longrightarrow D_{\min,i}$ and $a_j = D_{\min,i}$. This is not rigorous, but makes the analysis simpler and clearer. To be rigorous, one must carefully consider that full-cooperation regimes between the agents are achievable only asymptotically, when the encoding rates go to infinity.

The objective function of each agent is composed of two opposing quantities: the leakage and the distortion functions. Therefore we consider the objective function of agent j, $u_j : \mathcal{A}_j \times \mathcal{A}_i \to \mathbb{R}$, as the weighted sum of these two opposing quantities:

$$u_j(a_j, a_i) = -\tilde{w}_j L_j(a_j) + \hat{w}_j \log\left(\frac{\overline{D}_j}{a_i}\right) \tag{5}$$

where $L_j(a_j)$ represents the leakage of information at agent j as a function of a_j, i.e., the distortion at agent i, and a_i is the distortion at agent j. The two weighting factors, i.e., $\tilde{w}_j > 0$ and $\hat{w}_j > 0$ capture the emphasis of each agent on its leakage of information versus its fidelity of the state estimation.

III. Repeated non-cooperative games

The non-cooperative game we wish to analyze is described by the tuple $\mathcal{G} = (\mathcal{P}, \{\mathcal{A}_k\}_{k\in\mathcal{P}}, \{u_k\}_{k\in\mathcal{P}})$, where $\mathcal{P} \triangleq \{1, 2\}$ is the set of *players* (the two agents); $\mathcal{A}_k$ is the set of actions that agent k can take and u_k is the payoff function which measures the satisfaction of agent k for any strategy profile. The game is simply a decoupled optimization problem and the optimal solution corresponds to the system state where the amount of the data exchange between the two agents is minimal, i.e., $(a_1^*, a_2^*) = (\overline{D}_2, \overline{D}_1)$ as discussed in [4].

In this work, we are interested in the repeated games framework that might incentivize the agents to cooperate by exchanging data beyond the minimum imposed level of cooperation, naturally, without any pricing incentives [4]. We assume that the agents interact more than once under the same conditions, i.e. they play the same game repeatedly. Two scenarios are considered: i) finite-horizon repeated games in which the agents know in advance when their repeated interaction ends; and ii) infinite-horizon repeated games in which the agents do not know for sure when the game will end. In both cases, we study the possibility of sustaining cooperation by allowing the agents to make only credible commitments, i.e., commitments on which they have incentive to follow through. The equilibrium concept we investigate here is the subgame perfect equilibrium.

We assume that the game $\mathcal{G}$ is played several times, i.e., in $T > 1$ rounds where T can be either finite or infinite. In this scenario, players wish to maximize their averaged utilities over time. The particularity of repeated games as compared with the one-shot games is that players can 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 strategy profile at stage τ. The set of all possible histories at the end of stage t is denoted by $\mathcal{H}^{(t+1)}$ and considering $\mathcal{H}^{(1)}$ as the void set.

A pure strategy for a player is a contingent plan of how to play in each stage t for any history $h^{(t)}$.

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

The set of strategies for agent j is denoted by $\mathcal{S}_j$ and the set of strategy profiles is denoted by $\mathcal{S}$.

We assume that agents discount future payoffs. This means that present payoffs are more important to an agent than future ones.

Definition 2: The discounted payoff for player j for a joint strategy $s = (s_1, s_2)$ is given by

$$v_j(s) = (1-\rho)\sum_{t=1}^{T} \rho^{t-1} u_j(a^{(t)}), \tag{6}$$

where $a^{(t)}$ is the action profile induced by the joint strategy s, $u_j(\cdot)$ is the payoff function in (5), $\rho \in (0,1)$ is the discount factor which is known to all players.

The Nash equilibrium concept is defined in a way similar to that in one-shot games. However, some of the Nash equilibria in repeated games may rely on *empty threats* [5] of suboptimal play at histories that are not expected to occur. Thus, we focus our attention on a subset of Nash equilibria that are predictable in the sense that they allow players to make commitments only on which they have incentive to follow through: the subgame perfect equilibria. Before defining this concept, we first discuss subgames. Given any history $h^{(t)} \in \mathcal{H}^{(t)}$, known by the agents, the game from stage t onwards, with history $h^{(t)}$, is a subgame denoted by $\mathcal{G}_R(h^{(t)})$. The final history is, for this subgame, $h^{(T+1)} = (h^{(t)}, a^t, \ldots, a^T)$ and the strategies and payoffs are functions of the possible histories that are 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 3: A subgame perfect equilibrium, s^*, is a strategy profile (in a multi-stage game with observed actions) such that for all $h^{(t)} \in \mathcal{H}^{(t)}$, the restriction $s^*|h^{(t)}$ is a Nash equilibrium for the subgame $\mathcal{G}_R(h^{(t)})$.

We notice that this equilibrium concept is a refinement of the Nash equilibrium since it is required to be a Nash equilibrium not only in the entire history game but in every possible subgame. The main difference is that this concept allows the agents to make only credible commitments on which they have incentives to follow through.

A. Finite-horizon repeated game

Here, we assume that T is finite and that the agents know in advance its value and when the game ends precisely. Under these conditions, we show that cooperation beyond the minimum requirement cannot be enabled.

Theorem 2: In the finite-horizon repeated game $\mathcal{G}_R^{(T)} = (\mathcal{P}, \{\mathcal{S}_j\}_{j\in\mathcal{P}}, \{v_j\}_{j\in\mathcal{P}}, T)$, the unique subgame perfect equilibrium is "not to share any information above the minimum requirement" at each stage of the game and for both agents:

$$s_j^{(t),*} = \overline{D}_i, \ \forall t \in \{1, \ldots, T\}, \forall j \in \mathcal{P}. \tag{7}$$

The proof follows using tools similar to those used for the repeated prisoners' dilemma, by using an extension of the backward induction principle to dominance solvable games [5].

Being based on strict dominance, this result remains valid by replacing the discounted payoffs with simple averaged payoffs such as $\tilde{v}_j = \frac{1}{T}\sum_{t=1}^{T} u_j(a^{(t)})$.

B. Infinite-horizon repeated game

Here, we consider the situation in which the agents are unsure of the ending of the game or in which the game has infinite horizon. In this case, the set of subgame perfect equilibria cannot be determined by backward induction from the final stage since there is no precise information about the end of the game.

Theorem 3: In the infinite-horizon repeated game $\mathcal{G}_R^{(\infty)} = (\mathcal{P}, \{\mathcal{S}_j\}_{j\in\mathcal{P}}, \{v_j\}_{j\in\mathcal{P}})$, if $D_{\min,j} > 0$ for all $j \in \mathcal{P}$, then the strategy "not to share any information above the minimum requirement" at each stage of the game and for both agents is a subgame perfect equilibrium:

$$s_j^{(t),*} = \overline{D}_i, \ \forall t \geq 1, \forall j \in \mathcal{P}. \tag{8}$$

The proof follows by applying the *one-stage-deviation principle* for infinite-horizon games with discounted payoffs that are uniformly bounded in each stage [5]. As opposed to the finite-horizon repeated game, 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. As discussed in [6] and references therein, the discount factor can be seen as the probability that the game does not stop, at each stage of the game. Thus, the probability that the game stops at stage t is $(1-\rho)\rho^{t-1}$ and the discounted payoff corresponds to an expected utility.

Inspired by the infinite-horizon repeated prisoners' dilemma, our objective is to prove that, depending on the discount factor and system parameters, non-trivial exchange of information can be sustainable in the long term and that the strategy given in Theorem 3 is not the only subgame perfect equilibrium. We focus on the sub-set of action profiles $(D_2^*, D_1^*) \in (D_{\min,2}, \overline{D}_2) \times (D_{\min,1}, \overline{D}_1)$ that satisfy the following conditions:

$$\begin{cases} u_1(D_2^*, D_1^*) & > & u_1(\overline{D}_2, \overline{D}_1) \\ u_2(D_1^*, D_2^*) & > & u_2(\overline{D}_1, \overline{D}_2). \end{cases} \tag{9}$$

The intuition is that, distortion pairs that are better than the one-shot Nash equilibrium $D_j^* < \overline{D}_j$ and which achieve strictly better utilities than this *threat point* for both agents (9) may be expected to be sustainable agreements between the two agents in the long term. These profiles are not sustainable

in the one-shot or the finite-horizon repeated game scenario, because each agent cannot rely on or trust on the other agent to play a strictly dominant strategy at one stage in the finite interaction framework. However, we will see that, when the agents are unsure of the end of the game, the commitment of playing any pair (D_2^*, D_1^*) is sustainable under sufficient conditions on the discount factor. In other words, if the probability of the game stopping is small enough, then the commitment of playing (D_2^*, D_1^*) is a credible commitment in the infinite-horizon repeated game.

Theorem 4: In the infinite-horizon repeated game $\mathcal{G}_R^{(\infty)} = (\mathcal{P}, \{\mathcal{S}_j\}_{j\in\mathcal{P}}, \{v_j\}_{j\in\mathcal{P}})$ *and for any agreement profile* $(D_2^*, D_1^*) \in (D_{\min,2}, \overline{D}_2) \times (D_{\min,2}, \overline{D}_2)$ *satisfying the condition in* (9)*, if the discount factor is bounded as follows:*

$$1 > \rho > \max_{j\in\mathcal{P},\ \hat{D}_i\in(D_i^*,\overline{D}_i]} \left\{ \frac{u_j(\hat{D}_i, D_j^*) - u_j(D_i^*, D_j^*)}{u_j(\hat{D}_i, D_j^*) - u_j(\overline{D}_i, \overline{D}_j)} \right\} \tag{10}$$

and $D_{\min,j} > 0$ *for all* $j \in \mathcal{P}$*, then the following strategy is a subgame perfect equilibrium: "Any agent* j*: shares data at the agreement point* D_i^* *in the first stage and continues to share data at this agreement point so 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".*

The proof follows using the one-stage deviation principle similarly to the proof of Theorem 3. This result shows that the agents can achieve better distortion levels in a natural way in a long-term interaction process and without the intervention of a central authority that uses pricing techniques or enforces an unfair sum-utility objective upon the agents. All the distortion levels (D_i^*, D_j^*) that satisfy the conditions in (10) are achievable. Any other distortion pair (D_i, D_j) for which $u_j(D_i, D_j) < u_j(\overline{D}_i, \overline{D}_j)$ for some j, cannot be sustainable in the long term repeated game because agent j will have an incentive to deviate at any stage. In other words, such an agreement point is not credible in the long term.

IV. Numerical results

Let $r_1 \triangleq \hat{w}_1/\tilde{w}_1$ and $r_2 \triangleq \hat{w}_2/\tilde{w}_2$. We consider the following scenario: $\alpha = 0.9$, $\beta = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.1$, $\overline{D}_j = D_{\min,j} + 0.5(D_{\max,j} - D_{\min,j})$ for $j \in \mathcal{P}$, $D_{\min,1} = 0.3088$, $\overline{D}_1 = 0.3926$, $D_{\min,2} = 0.2183$ and $\overline{D}_2 = 0.2388$. For the cases in which the leakage term is important (e.g., $r_1 = r_2 = 1$ or $r_1 = 1, r_2 = 2$) there is no distortion pair (D_2^*, D_1^*) that strictly improves both players' payoffs w.r.t $(\overline{D}_2, \overline{D}_1)$. This means that the improvement in an agent's estimation fidelity resulting from the other player cooperating above the minimum requirement is not good enough to justify the loss in privacy incurred by the agreement point. However, if these ratios increase, then the sustainable region of pairs (D_2^*, D_1^*) that satisfy jointly the conditions in (9) and (10) is not trivial. This fact is illustrated in Fig. 1 for the cases: i) $r_1 = 1$, $r_2 = 5$; and ii) $r_1 = 5$, $r_2 = 5$. The white region represents all the possible agreements that are credible commitments, whereas

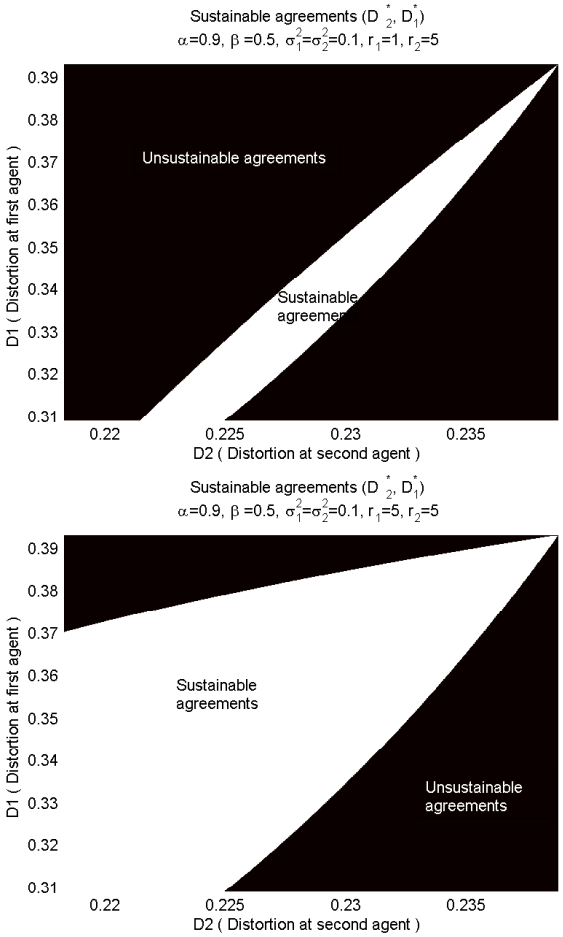


Fig. 1. The subset of all possible (D_2^*, D_1^*) that are credible commitments in the infinite-horizon repeated game when $\alpha = 0.9$, $\beta = 0.5$, $\sigma_1^2 = \sigma_2^2 = 0.1$: i) $r_1 = 1, r_2 = 5$; ii) $r_1 = r_2 = 5$.

the dark region represent the distortion points that cannot be achieved. The upper-right corner represents the minimum cooperation requirement $(\overline{D}_2, \overline{D}_1)$. We observe that the very asymmetric points are not generally achievable. This is simply because one player will not share a lot of information while expecting little in return. We observe also that the higher the impact of the fidelity w.r.t the leakage term, the larger the sustainable region is and the lower the distortion levels that are achievable.

V. Conclusions

A repeated non-cooperative game has been investigated as a possibility to enable data sharing among autonomous interconnected agents to improve the fidelity of their own state estimates. In the finite horizon case, it is not possible to sustain data sharing beyond the minimum requirement. If the interaction of the agents is repeated infinitely many times, non-trivial solutions exist for which the impact of estimation fidelity is larger than that of the leakage of information. As opposed to the situation with pricing techniques [4], not all

the distortion pairs are achievable: the data-sharing must be mutual to be credible.

REFERENCES

[1] T. V. Cutsem, J. L. Howard, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Trans. Power App. Syst.*, vol. 100, no. 8, pp. 3722–3732, Aug. 1981.

[2] A. Gómez-Expósito, A. Abur, A. de la Villa Jaén, and C. Gómez-Quiles, "A multilevel state estimation paradigm for smart grids," *Proceedings of the IEEE*, vol. 99, no. 6, pp. 952–976, Jun. 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. 2nd IEEE Intl. Conf. on Smart Grid Commun.*, Brussels, Belgium, Oct. 2011.

[4] E. V. Belmega, L. Sankar, H. V. Poor, and M. Debbah, "Pricing mechanisms for cooperative state estimation," in *Proc. 5th Intl. Symp. on Commun., Control and Signal Process. (ISCCSP)*, Roma, Italy, May 2012, pp. 1–4.

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

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