CO4 Workshop, Toulouse, France

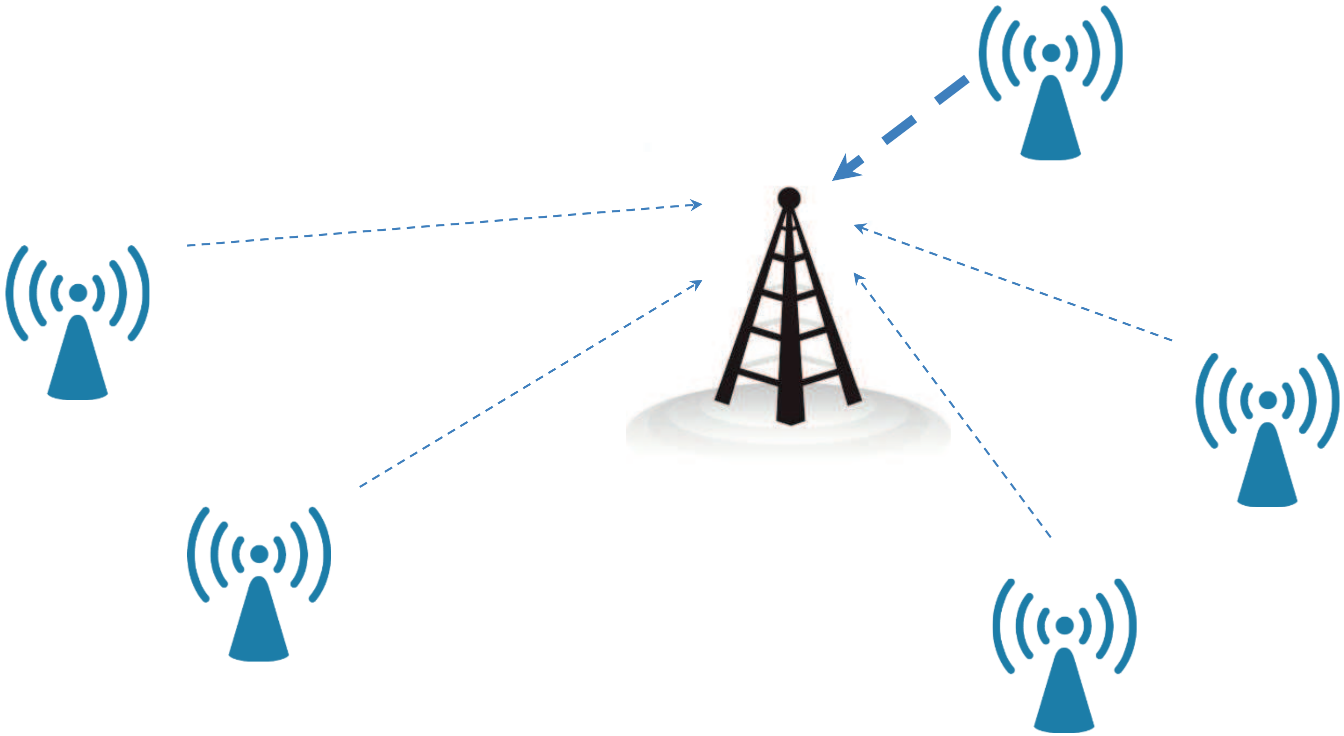
Information constraints in multiple agent problems with i.i.d. states

S. Lasaulce
L2S (CNRS–CentraleSupelec) – lasaulce@lss.supelec.fr
joint work with several co-authors

October 26, 2016

Motivation example: Power control

Power control in wireless networks



A typical distributed power control setup

- Set of transmitters: $\mathcal{K} = \{1, ..., K\}$
- What Transmitter k does:

$$\begin{array}{cccc} g_k(1) & g_k(2) & g_k(3) & \cdots \\ \hline p_k(1) & p_k(2) & p_k(3) & \cdots \end{array}$$

- Performance criteria: data rate, energy, etc.

$$u_k(\underbrace{g_1, ..., g_K}_{a_0}, p_1, ..., p_K)$$

- **Local decision, local observation $\rightarrow$ Ultimate average performance? (Cf Opt. Ctrl.)**

General problem statement

- Set of agents: $\mathcal{K} = \{1, ..., K\}$

- Stage/instantaneous perf. criteria (utilities):

$$u_k(a_0, a_1, ..., a_K)$$

- Action sets: $\mathcal{A}_k$ with $|\mathcal{A}_k| < \infty$

- T iterations/samples/stages: $t \in \{1, ..., T\}$

- Observation: $(a_0(t), a_1(t), ..., a_K(t)) \to o_k(t)$

Ultimate Goal: characterize the feasible long-term utility region

- Long-term utilities:

$$v_k^\infty = \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} u_k(a_0(t), a_1(t), ..., a_K(t))$$

Ultimate goal: characterize the feasible long-term utility region

- Long-term utilities:

$$v_k^{\infty} = \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[u_k(a_0(t), a_1(t), ..., a_K(t))\right]$$

Ultimate goal: characterize the feasible long-term utility region

- Long-term utilities:

$$v_k^{\infty}(f_1, ..., f_K) = \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[u_k(a_0(t), a_1(t), ..., a_K(t))\right]$$

Ultimate goal: characterize the feasible long-term utility region

- Long-term utilities:

$$v_k^{\infty}(f_1, ..., f_K) = \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[u_k(a_0(t), a_1(t), ..., a_K(t))\right]$$

- $f_1, ..., f_K =$???

Strategies

- Strategies informal definition: history $\mapsto$ action

- Observation structure: $o_k = (s_k, y_k)$ where s_k is given by $\daleth_k(s_k|a_0)$ and y_k by $\Gamma_k(y_k|a_0, a_1, ..., a_K)$

- Example: $s_k = \widehat{g}_k$ and $y_k = \widehat{u}_k$

Strategies. Continued

- Definition (causal scenario):

$$a_k(t) = f_{k,t}(s_k(1), ..., s_k(t), y_k(1), ..., y_k(t-1))$$

Strategies. Continued

- Definition (causal scenario):

$$a_k(t) = f_{k,t}(s_k(1), ..., s_k(t), y_k(1), ..., y_k(t-1))$$

- Definition (noncausal scenario):

$$a_k(t) = f_{k,t}(s_k(1), ..., s_k(T), y_k(1), ..., y_k(t-1))$$

Problem statement: recap

Find the feasible region $(v_1^{\infty}, ..., v_K^{\infty})$

- for the causal/**noncausal** scenario

- when $a_0(t)$ is i.i.d. and $\sim \rho_0$

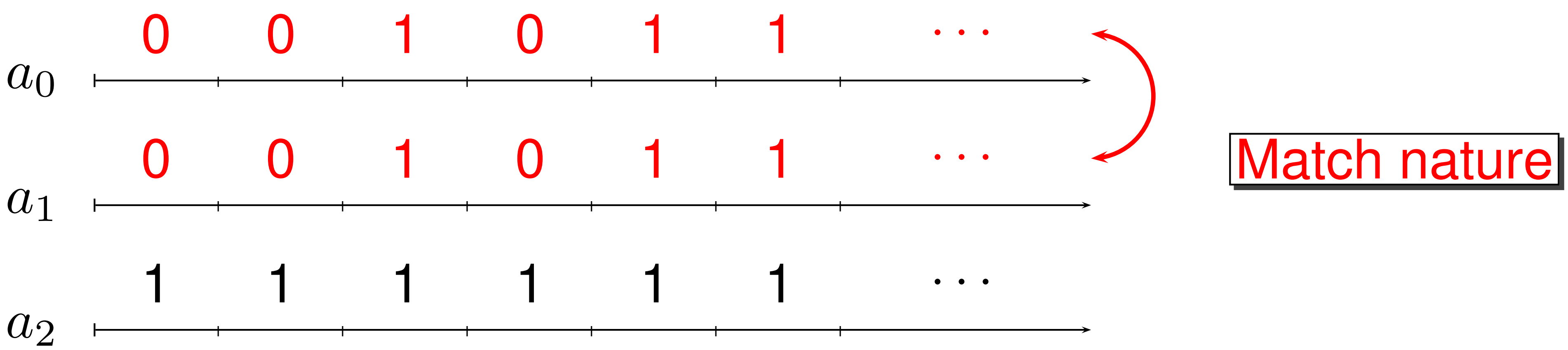
- with the memoryless observation structure given by $\daleth_k$ and Γ_k

Example (noncausal+asymmetrical scenario)

- Agents: $\{1, 2\}$; $0 \equiv$ nature.
- Action sets: $\mathcal{A}_0 = \mathcal{A}_1 = \mathcal{A}_2 = \{0, 1\}$ (think of PC).
- Stage utility function:

$$u(a_0, a_1, a_2) = \left| \begin{array}{ll} 1 & \text{if } a_0 = a_1 = a_2 \\ 0 & \text{otherwise} \end{array} \right. .$$

- Scheme 1:

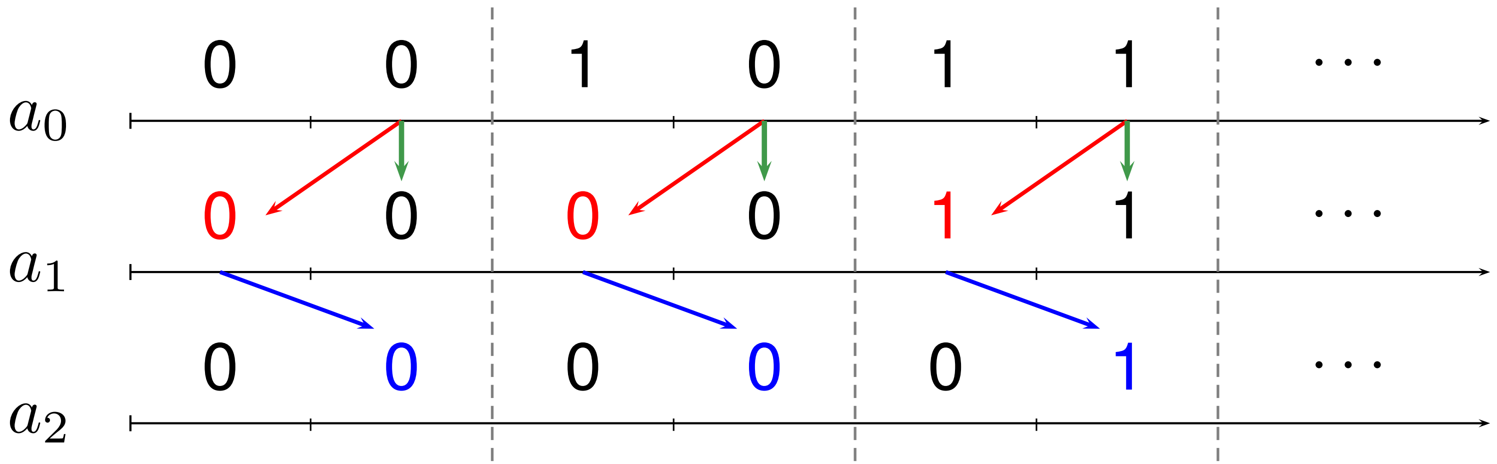


- Long-term utility

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} u(a(t))\right] \to \frac{1}{2} = 0.5. \qquad \text{for } A_0 \sim \mathcal{B}\left(\frac{1}{2}\right)$$

Long-term utility

- Scheme 2:



- Long term utility:

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} u(a(t))\right] \to \frac{5}{8} = 0.625.$$

Maximal long-term utility

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} u(a(t))\right] \to x^\star \simeq 0.81$$

where

$$x^\star \quad \text{is the solution of} \quad \frac{h(x)-1}{x-1} = \log_2 3$$

and $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$.

Performance characterization of stochastic games with i.i.d. states

Noncausal scenario

Key observation. Say $\mathcal{A} = \{0, 1\}$, $K = 1$

$$\frac{1}{T}\sum_{t=1}^{T} u(a(t))$$

$$= \frac{1}{T}\left[u(0) + u(1) + u(1) + u(0) + \ldots + u(1)\right]$$

$$= \frac{N_0}{T} u(0) + \frac{N_1}{T} u(1)$$

$$= q_0 u(0) + q_1 u(1) \qquad (q_0 \geq 0, q_1 \geq 0, q_0 + q_1 = 1)$$

$$= \sum_{a \in \mathcal{A}} q_a u(a) = \mathbb{E}[u(a)]$$

More formally

$$
\begin{aligned}
& v_i^{\infty}(f_1, ..., f_K) \\
&= \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[u_k(a_0(t), a_1(t), ..., a_K(t))\right] \\
&= \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{a_0, ..., a_K} P_t(a_0, ..., a_K) u_k(a_0, ... a_K) \\
&= \sum_{a_0, ..., a_K} u_k(a_0, ..., a_K) \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} P_t(a_0, ..., a_K)
\end{aligned}
$$

Implementable coordination

Definition $Q(a_0, a_1, ..., a_K)$ is implementable if $\exists$ $(f_1, ..., f_K)$ s.t.

$$\frac{1}{T}\sum_{t=1}^{T} P_t(a_0, ..., a_K) \to Q(a_0, a_1, ..., a_K)$$

Reminder:

$$a_k(t) = f_{k,t}(s_k(1), ..., s_k(T), y_k(1), ..., y_k(t-1))$$

Extreme cases of distribution

- Team control $u_k = u$

- Zero correlation (lower bound):

$$Q(a_0, a_1, ..., a_K) = \rho_0(a_0) \prod_{k=1}^{K} P_{A_k}(a_k)$$

- Full correlation (upper bound):

$$(a_1(t), ..., a_K(t)) \in \text{arg} \max_{a_1,...,a_K} u(a_0(t), a_1, ..., a_K)$$

$$\rightarrow Q(a_0, \underbrace{a_1, ..., a_K}_{a}) = \rho_0(a_0)\delta(a - m(a_0))$$

Implementable distribution set characterization

Proposition 1 Solving the problem is at least as hard as solving the two-way channel.

Theorem 2

$- K = 2$

$- a_1(t) = f_{1,t}(s_1(1), ..., s_1(T))$

$- a_2(t) = f_{2,t}(s_2(1), ..., s_2(T), y_2(1), ..., y_2(t-1))$

► Then $Q(a_0, a_1, a_2)$ is implementable iff ... see [Larrousse et al ITW 2015].

Review of the entropy function

For a vector in the unit simplex $x = (x_1, ..., x_N)$**,** $x_n \geq 0$**:** $\sum_n x_n = 1$**.**

$$H(x) = -\sum_n x_n \log x_n.$$

For a random variable $A \sim P$**:**

$$H(A) = -\sum_{a \in \mathcal{A}} P(a) \log P(a).$$

Implementable distribution set characterization

Corollary 3

– $K = 2$

– $a_1(t) = f_{1,t}(a_0(1), ..., a_0(T))$

– $a_2(t) = f_{2,t}(a_1(1), ..., a_1(t-1))$

► Then $Q(a_0, a_1, a_2)$ is implementable iff its marginal w.r.t (a_1, a_2) is ρ_0 and

$$H_Q(A_0, A_1, A_2) \geq H_Q(A_0) + H_Q(A_2).$$

Utility region characterization

Pareto frontier: use $w_\alpha = \alpha u_1 + (1-\alpha) u_2$

$$
\begin{array}{ll}
\text{minimize} & -\sum_{a_0,a_1,a_2} Q(a_0,a_1,a_2) w_\alpha(a_0,a_1,a_2) \\
\text{subject to} & \begin{aligned} H_Q(A_0) + H_Q(A_2) - H_Q(A_0,A_1,A_2) &\le 0 \\ -Q(a_0,a_1,a_2) &\le 0 \\ -1 + \sum_{a_0,a_1,a_2} Q(a_0,a_1,a_2) &= 0 \\ -\rho_0(a_0) + \sum_{a_1,a_2} Q(a_0,a_1,a_2) &= 0 \end{aligned}
\end{array}
$$

Let us try to interpret by specializing further

- $H_Q(a_0) =$ constant $= -\sum_{a_0} \rho_0(a_0) \log \rho_0(a_0)$

- $H_Q(a_2) \sim$ constant

- Boltzmann-Gibbs is optimal:

$$Q^\star(a_0, a_1, a_2) = \frac{e^{\lambda u(a_0,a_1,a_2)}}{\sum_{a_0,a_1,a_2} e^{\lambda u(a_0,a_1,a_2)}}$$

Performance characterization of stochastic games with i.i.d. states

Causal scenario: omitted

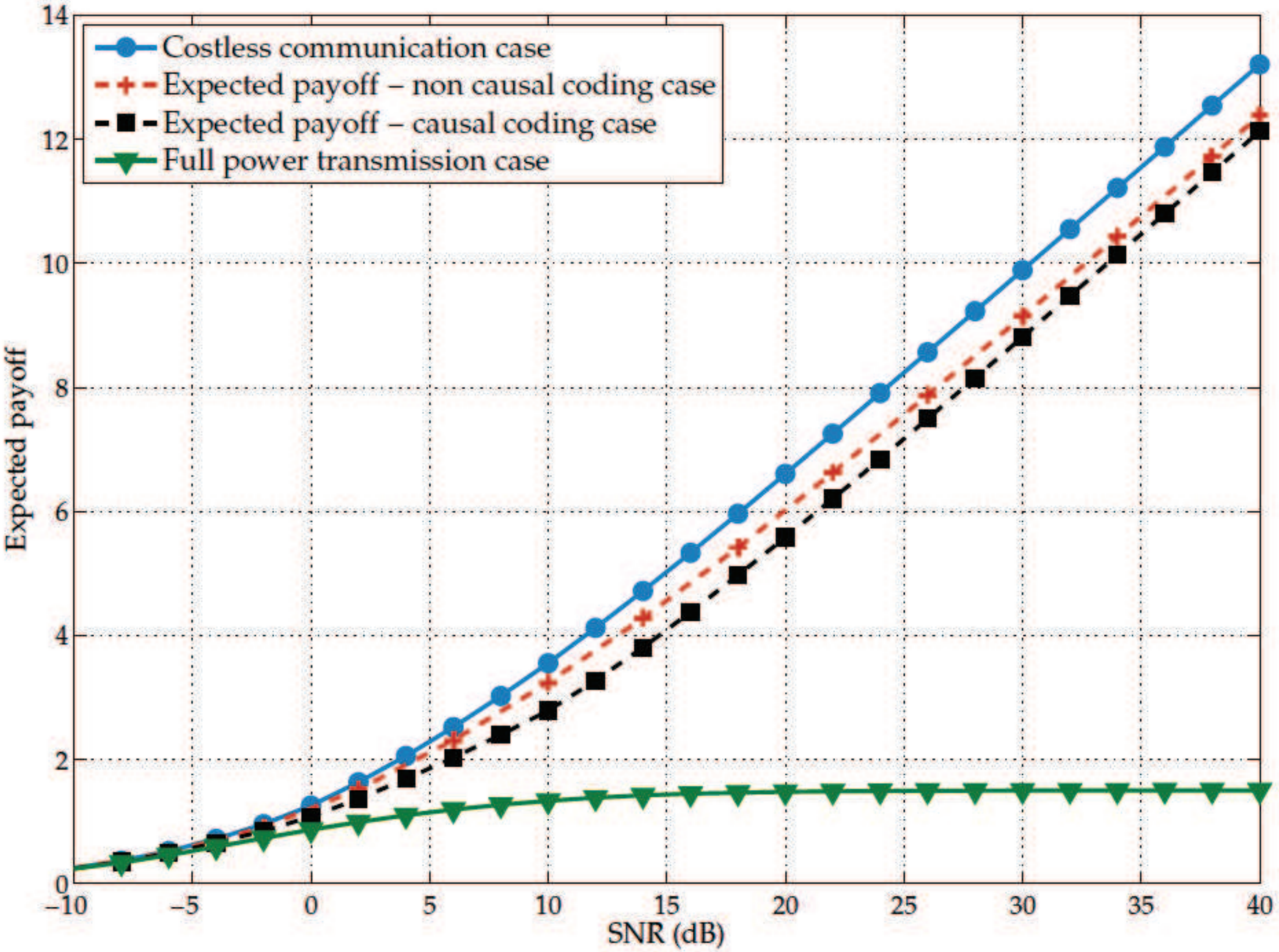
Application to power control over the interference channel

- $K = 2$ transmitter-receiver pairs

- Single-band case

- Utility $u = \log(1 + \mathrm{SINR}_1) + \log(1 + \mathrm{SINR}_2)$ with $\mathrm{SINR}_1 = \frac{g_{11}a_1}{\sigma^2 + g_{21}a_2}$, $\mathrm{SINR}_2 = \frac{g_{22}a_2}{\sigma^2 + g_{12}a_1}$.

Application to power control over the interference channel



Technical challenges

- ► Construct codes (see [Larrousse and Lasaulce ISIT 2013][Larrousse et al TIT 2016]). Joint control-communication problem.

- ► Controlled states.

- ► Nash equilibrium points.

Thank you for your attention!

Main references (1/2)

C. Zhang, N. Khalfet, S. Lasaulce, and S. Tarbouriech, ”Utility-oriented quantization and application to power control”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017.

B. Larrousse, S. Lasaulce, and M. Bloch, ”Coordination in distributed networks via coded actions with application to power control”, IEEE Transactions on Information Theory. Available on arXiv.

S. Berri, S. Lasaulce, and M. S. Radjef, ”Power control with partial observation in wireless ad hoc networks”, IEEE Proc. of the EUSIPCO conference, Budapest, Hungary, Aug. 2016.

A. Agrawal, S. Lasaulce, O. Beaude, and R. Visoz, ”A framework for decentralized power control with partial channel state information”, IEEE Proc. of the Fifth International Conference on Communications and Networking (ComNet2015), Hammamet, Tunisia, November 4-7, 2015.

O. Beaude, A. Agrawal, and S. Lasaulce, ”A framework for computing power consumption scheduling functions under uncertainty”, 6th IEEE International Conference on Smart Grid Communications (SmartGridComm 2015), Miami, Florida, USA, Nov. 2015.

Main references (2/2)

A. Agrawal, F. Danard, B. Larrousse, and S. Lasaulce, ”Implicit coordination in two-agent team problems”, European Conference on Control (ECC), Linz, Austria, July 2015.

B. Larrousse, S. Lasaulce, and M. Wigger, ”Coordination in state-dependent distributed networks: The two-agent case”, IEEE International Symposium on Information Theory (ISIT), Hong Kong, June 2015.

B. Larrousse, S. Lasaulce, and M. Wigger, ”Coordination in State-Dependent Distributed Networks”, IEEE Proc. of the Information Theory Workshop (ITW), Jerusalem, Israel, Apr.-May 2015.

B. Larrousse, A. Agrawal, and S. Lasaulce, ”Implicit coordination in two-agent team problems. Application to distributed power allocation”, IEEE 12th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Hammamet, Tunisia, May 2014.

Performance characterization of stochastic games with i.i.d. states

Causal scenario

Implementable distribution set characterization

Theorem 1 [Larrousse et al ITW 2015] Let $K \geq 2$. $Q(a_0, ..., a_K)$ is implementable if and only if it factorizes as

$$Q(a_0, ..., a_K) =$$

$$\sum_{s_1,...,s_k,w} \rho_0(a_0) \daleth(s_1, ..., s_k | a_0) P(w) \prod_{k=1}^{K} P_{A_k|S_k,W}(a_k | s_k, w)$$

Finding particular stationary strategies [Lasaulce and Visoz 2015]

- Set $u_k = u$

- Set $a_k(t) = \overline{f}_k(s_k(t), w(t))$

- Apply the sequential best response dynamics on

$$\mathbb{E}[u] = \sum_{a_0,a} Q(a_0, a) u(a_0, a) = U(P_{A_1|S_1,W}, ..., P_{A_K|S_K,W}, P(w))$$

$\rightarrow \overline{f}_1, ..., \overline{f}_K$