

Towards Optimal Information Exchange Instants in Multi-Agent Systems

Domagoj Tolić

University of Dubrovnik & RIT Croatia

Chair of Information-Oriented Control
Technische Universität München
June 13, 2017



Outline

- 1 Introduction
 - Motivation
 - Preliminaries
- 2 Problem Statement
 - Decentralized Control
 - Optimal Intermittent Feedback
- 3 Methodology
 - Stabilizing Broadcasting Intervals
 - Suboptimal Broadcasting Intervals
- 4 Validation
- 5 Summary

Decentralized Control

- **feedback** is instrumental for **control**
- data exchange among **neighbors** is instrumental in MAS coordination
- data are exchanged **intermittently**
- realistic communication channels **distort** data, introduce **delays** and packet **dropouts**
- however, sensing and broadcasting consume **energy**

Optimal Decentralized Control

- **quantify** the repercussions of intermittent feedback
- MAS control **performance** vs. MAS **lifetime**
- local Dynamic Programming (DP) problems are **coupled**
⇒ **nonautonomous** dynamics ⇒ **non-stationary**
cost-to-go
- the need for an **online model-free** Reinforcement Learning (RL) method
- Kalman Filtering (KF) for delayed, sampled and noisy data

Impulsive delayed systems

$$\Sigma \left\{ \begin{array}{ll} \dot{x}(t) = Ax(t) + A_d x(t-d) + B\omega(t), & t \notin \mathcal{T}, \\ y(t) = Cx(t) + C_d x(t-d) + D\omega(t), & t \geq t_0, \\ x(t^+) = Ex(t) + E_d x(t-d), & t \in \mathcal{T}, \end{array} \right.$$

Impulsive delayed systems

$$\Sigma \begin{cases} \dot{x}(t) = Ax(t) + A_d x(t-d) + B\omega(t), & t \notin \mathcal{T}, \\ y(t) = Cx(t) + C_d x(t-d) + D\omega(t), & t \geq t_0, \\ x(t^+) = Ex(t) + E_d x(t-d), & t \in \mathcal{T}, \end{cases}$$

where $x \in \mathbb{R}^{n_x}$ is the **state**, $\omega \in \mathbb{R}^{n_\omega}$ is the **input**, $y \in \mathbb{R}^{n_y}$ is the **output** and $d \geq 0$ is the time **delay**

$\mathcal{L}_p$ -stability w.r.t. Set and with Bias

- $\mathcal{L}_p$ -norm w.r.t. a set $\mathcal{B} \subset \mathbb{R}^n$:

$$\|f[a, b]\|_{p, \mathcal{B}} := \left(\int_{[a, b]} \|f(s)\|_{\mathcal{B}}^p ds \right)^{1/p}, \text{ where}$$
$$\|f(s)\|_{\mathcal{B}} := \inf_{b \in \mathcal{B}} \|f(s) - b\| \text{ and } p \in [1, \infty]$$

- output set:

$$\mathcal{B}_y := \left\{ y \in \mathbb{R}^{n_y} \mid \exists b \in \mathcal{B} \text{ such that } y = (C + C_d)b \right\}, \text{ where}$$
$$\mathcal{B} := \text{Ker}(A + A_d)$$

$\mathcal{L}_p$ -stability w.r.t. Set and with Bias

- $\mathcal{L}_p$ -norm w.r.t. a set $\mathcal{B} \subset \mathbb{R}^n$:

$$\|f[a, b]\|_{p, \mathcal{B}} := \left(\int_{[a, b]} \|f(s)\|_{\mathcal{B}}^p ds \right)^{1/p}, \text{ where}$$

$$\|f(s)\|_{\mathcal{B}} := \inf_{b \in \mathcal{B}} \|f(s) - b\| \text{ and } p \in [1, \infty]$$

- output set:

$$\mathcal{B}_y := \left\{ y \in \mathbb{R}^{n_y} \mid \exists b \in \mathcal{B} \text{ such that } y = (C + C_d)b \right\}, \text{ where}$$

$$\mathcal{B} := \text{Ker}(A + A_d)$$

Definition ($\mathcal{L}_p$ -Stability w.r.t. $\mathcal{B}$ with Bias b)

Let $p \in [1, \infty]$. The system Σ is $\mathcal{L}_p$ -stable w.r.t. a set $\mathcal{B}$ and with bias $b(t) \equiv b \geq 0$ from ω to y with gain $\gamma \geq 0$ if there exists $K \geq 0$ such that, for each $t_0 \in \mathbb{R}$ and each $\psi_x \in PC([t_0 - d, t_0], \mathbb{R}^{n_x})$, each solution to Σ from ψ_x at $t = t_0$ satisfies

$$\|y[t_0, t]\|_{p, \mathcal{B}_y} \leq K \|\psi_x\|_{d, \mathcal{B}} + \gamma \|\omega[t_0, t]\|_p + \|b[t_0, t]\|_p \text{ for each } t \geq t_0.$$

Agent Dynamics

- consider N **heterogeneous** linear agents given by

$$\begin{aligned}\dot{\xi}_i &= A_i \xi_i + B_i u_i + \omega_i, \\ \zeta_i &= C_i \xi_i,\end{aligned}\tag{1}$$

where $\xi_i \in \mathbb{R}^{n_{\xi_i}}$ is the **state**, $u_i \in \mathbb{R}^{n_{u_i}}$ is the **input**, $\zeta_i \in \mathbb{R}^{n_{\zeta}}$ is the **output** of the i^{th} agent, $i \in \{1, 2, \dots, N\}$, and $\omega_i \in \mathbb{R}^{n_{\xi_i}}$ reflects exogenous **disturbances** and/or modeling **uncertainties**

Agent Dynamics

- consider N **heterogeneous** linear agents given by

$$\begin{aligned}\dot{\xi}_i &= A_i \xi_i + B_i u_i + \omega_i, \\ \zeta_i &= C_i \xi_i,\end{aligned}\tag{1}$$

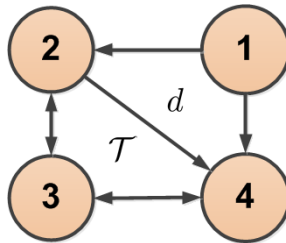
where $\xi_i \in \mathbb{R}^{n_{\xi_i}}$ is the **state**, $u_i \in \mathbb{R}^{n_{u_i}}$ is the **input**, $\zeta_i \in \mathbb{R}^{n_{\zeta}}$ is the **output** of the i^{th} agent, $i \in \{1, 2, \dots, N\}$, and $\omega_i \in \mathbb{R}^{n_{\xi_i}}$ reflects exogenous **disturbances** and/or modeling **uncertainties**

- a common decentralized policy is

$$u_i(t) = -K_i \sum_{j \in \mathcal{N}_i} (\zeta_i(t) - \zeta_j(t)),\tag{2}$$

where K_i is an $n_{u_i} \times n_{\zeta}$ gain matrix

Agent Interconnections



Closed-Loop Dynamics

- define $\xi := (\xi_1, \dots, \xi_N)$, $\zeta := (\zeta_1, \dots, \zeta_N)$ and $\omega := (\omega_1, \dots, \omega_N)$
- utilizing the **Laplacian matrix** L of the **communication graph** $\mathcal{G}$, we reach

$$\begin{aligned}\dot{\xi}(t) &= A^{\text{cl}}\xi(t) + A^{\text{cld}}\xi(t-d) + \omega(t), \\ \zeta &= C^{\text{cl}}\xi,\end{aligned}$$

with

$$\begin{aligned}A^{\text{cl}} &= \text{diag}(A_1, \dots, A_N), & A^{\text{cld}} &= [A_{ij}^{\text{cld}}], \\ A_{ij}^{\text{cld}} &= -l_{ij}B_iK_iC_j, & C^{\text{cl}} &= \text{diag}(C_1, \dots, C_N),\end{aligned}$$

Optimal Intermittent Feedback

- $t_i^j \in \mathcal{T}$, $i \in \mathbb{N}$ – broadcasting instants of the j^{th} agent
- **asynchronous** communication
- $x_i := (\dots, \zeta_i - \zeta_j, \dots)$, where $i \in \{1, \dots, N\}$ and $j \in \mathcal{N}_i$

Optimal Intermittent Feedback

- $t_i^j \in \mathcal{T}$, $i \in \mathbb{N}$ – broadcasting instants of the j^{th} agent
- **asynchronous** communication
- $x_i := (\dots, \zeta_i - \zeta_j, \dots)$, where $i \in \{1, \dots, N\}$ and $j \in \mathcal{N}_i$

Problem

For each $j \in \{1, \dots, N\}$, **minimize** the following cost function that captures **performance vs. energy** trade-offs

$$\mathbb{E}_{\omega} \left\{ \sum_{i=1}^{\infty} (\gamma_j)^i \left[\underbrace{\int_{t_{i-1}^j}^{t_i^j} (x_j^{\top} P_j x_j + u_j^{\top} R_j u_j) dt}_{r_j(x_j, u_j, \tau_i^j)} + S_j \right] \right\} \quad (3)$$

for the j^{th} agent of MAS (1)-(2) over all sampling policies τ_i^j and for all initial conditions $x_j(t_0) \in \mathbb{R}^{n_{xj}}$.

Interconnecting Nominal and Error System

- introduce

$$e(t) = (e_1(t), \dots, e_N(t)) := \hat{\zeta}(t) - \zeta(t-d)$$

- closed-loop dynamics become

$$\dot{\xi}(t) = A^{\text{cl}}\xi(t) + A^{\text{cl}d}\xi(t-d) + A^{\text{cle}}e(t) + \omega(t),$$

$$\zeta = C^{\text{cl}}\xi,$$

with $A^{\text{cle}} = [A_{ij}^{\text{cle}}]$, $A_{ij}^{\text{cle}} = -l_{ij}B_iK_i$

Interconnecting Nominal and Error System

- introduce

$$e(t) = (e_1(t), \dots, e_N(t)) := \hat{\zeta}(t) - \zeta(t-d)$$

- closed-loop dynamics become

$$\begin{aligned}\dot{\xi}(t) &= A^{\text{cl}}\xi(t) + A^{\text{cl}d}\xi(t-d) + A^{\text{cle}}e(t) + \omega(t), \\ \zeta &= C^{\text{cl}}\xi,\end{aligned}$$

with $A^{\text{cle}} = [A_{ij}^{\text{cle}}]$, $A_{ij}^{\text{cle}} = -l_{ij}B_iK_i$

- ZOH sampling yields

$$\dot{e}(t) = -\dot{\zeta}(t-d) = -C^{\text{cl}}\dot{\xi}(t-d),$$

- for each $t_i^j + d \in (\mathcal{T} + d)$ we have

$$\begin{aligned}e_k((t_i^j + d)^+) &= e_k(t_i^j + d), & k \in \{1, \dots, N\}, k \neq j, \\ e_j((t_i^j + d)^+) &= \nu_j(t_i^j + d)\end{aligned}$$

Small-Gain Theorem

- select

$$\tilde{\zeta} := -C^{\text{cl}}[A^{\text{cl}}\xi(t-d) + A^{\text{cl}d}\xi(t-2d) + \omega(t-d)]$$

to be the output of the **nominal system** for which

$$\|\tilde{\zeta}[t_0, t]\|_{p, \mathcal{B}_{\tilde{\zeta}}} \leq K_n \|\psi_{\xi}\|_{d, \mathcal{B}} + \gamma_n \|(e, \omega)[t_0, t]\|_p \quad (4)$$

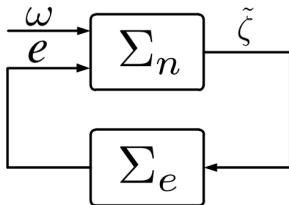
Small-Gain Theorem

- select

$$\tilde{\zeta} := -C^{\text{cl}}[A^{\text{cl}}\xi(t-d) + A^{\text{cl}d}\xi(t-2d) + \omega(t-d)]$$

to be the output of the **nominal system** for which

$$\|\tilde{\zeta}[t_0, t]\|_{p, \mathcal{B}_{\tilde{\zeta}}} \leq K_n \|\psi_{\xi}\|_{d, \mathcal{B}} + \gamma_n \|(e, \omega)[t_0, t]\|_p \quad (4)$$



Stabilizing Broadcasting Intervals

Theorem

Suppose the communication link delay d for the MAS (1)-(2) yields (4) for some $p \in [1, \infty]$. If the broadcasting intervals τ_i^j , $i \in \mathbb{N}$, $j \in \{1, \dots, N\}$, satisfy (I) and (II) for some $\lambda > 0$ and $M > 1$ such that $\frac{2}{\lambda} \sqrt{M} \gamma_n < 1$, then the MAS (1)-(2) is $\mathcal{L}_p$ -stable from ω to $(\tilde{\zeta}, e)$ w.r.t. $(\mathcal{B}, \mathbf{0}_{n_e})$ and with bias.

Stabilizing Broadcasting Intervals

Theorem

Suppose the communication link delay d for the MAS (1)-(2) yields (4) for some $p \in [1, \infty]$. If the broadcasting intervals τ_i^j , $i \in \mathbb{N}, j \in \{1, \dots, N\}$, satisfy (I) and (II) for some $\lambda > 0$ and $M > 1$ such that $\frac{2}{\lambda} \sqrt{M} \gamma_n < 1$, then the MAS (1)-(2) is $\mathcal{L}_p$ -stable from ω to $(\tilde{\zeta}, e)$ w.r.t. $(\mathcal{B}, \mathbf{0}_{n_e})$ and with bias.

- we can always choose τ_i^j 's such that

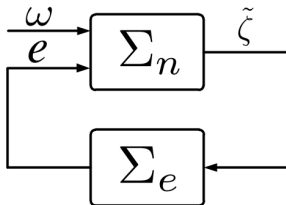
- (I) $\tau_i^j (\lambda + r + \lambda_1 M e^{-\lambda \tau_i^j}) < \ln M$, and
- (II) $\tau_i^j (\lambda + r + \frac{\lambda_1}{\lambda_2} e^{\lambda d}) < -\ln \lambda_2$,

with $r > 0$ being an arbitrary constant, $\lambda_1 := \frac{N \|C^{\text{cl}} A^{\text{cle}}\|^2}{r}$ and $\lambda_2 := \frac{N-1}{N}$.

Stabilizing Broadcasting Intervals

Corollary

Suppose the conditions of the theorem hold and ξ is $\mathcal{L}_p$ -detectable from $(e, \omega, \tilde{\zeta})$ w.r.t. $\mathcal{B}$. Then the MAS (1)-(2) is $\mathcal{L}_p$ -stable with bias w.r.t. $(\mathcal{B}, \mathbf{0}_{n_e})$ from ω to (ξ, e) .



Experimental Results

Experimental Results

Least Square Policy Iteration (LSPI) I

- LSPI **state-action approximate value function** is

$$\hat{Q}(x(t_i), \tau(t_i)) = \Phi^\top(x(t_i), \tau(t_i)) \alpha_\kappa, \quad (5)$$

where

$$\Phi(x(t_i), \tau(t_i)) = \psi(\tau(t_i)) \otimes \phi(x(t_i))$$

is the Kronecker product of the basis function vectors $\psi(\tau(t_i))$ and $\phi(x(t_i))$ formed with **Chebyshev polynomials** while α_κ is **being learned**

Least Square Policy Iteration (LSPI) II

- define $\tau(t_i) := t_{i+1} - t_i$
- **decision** $\tau(t_i) \in \mathcal{A}$ is given by

$$\tau(t_i) = h_{\kappa}(x(t_i)),$$

where

$$h_{\kappa}(x(t_i)) = \begin{cases} \text{u.r.a.} \in \mathcal{A} & \text{every } \varepsilon \text{ iterations,} \\ h_{\kappa}(x(t_i)) & \text{otherwise,} \end{cases}$$

Least Square Policy Iteration (LSPI) II

- define $\tau(t_i) := t_{i+1} - t_i$
- **decision** $\tau(t_i) \in \mathcal{A}$ is given by

$$\tau(t_i) = h_{\kappa}(x(t_i)),$$

where

$$h_{\kappa}(x(t_i)) = \begin{cases} \text{u.r.a.} \in \mathcal{A} & \text{every } \varepsilon \text{ iterations,} \\ h_{\kappa}(x(t_i)) & \text{otherwise,} \end{cases}$$

where “u.r.a.” stands for “uniformly chosen random action” and yields **exploration** every ε steps while $h_{\kappa}(x(t_i))$ is the **policy** obtained according to

$$h_{\kappa}(x(t_i)) \in \arg \min_u \hat{Q}(x(t_i), \tau(t_i)) \quad (6)$$

Least Square Policy Iteration (LSPI) III

- α_κ is updated every $\kappa \geq 1$ steps from the **projected Bellman equation** for **model-free policy iteration**

$$\Gamma_i \alpha_\kappa = \gamma \Lambda_i \alpha_\kappa + z_i,$$

where γ is from (3) and

$$\Gamma_0 = \beta_\Gamma I, \quad \Lambda_0 = \mathbf{0}, \quad z_0 = \mathbf{0},$$

$$\Gamma_i = \Gamma_{i-1} + \phi(x(t_i), \tau(t_i)) \phi(x(t_{i-1}), \tau(t_{i-1}))^\top,$$

$$\Lambda_i = \Lambda_{i-1} + \phi(x(t_i), \tau(t_i)) \phi(x(t_i), h(x(t_{i+1})))^\top,$$

$$z_i = z_{i-1} + \phi(x(t_i), \tau(t_i)) r(t_i),$$

where Γ_i , Λ_i and z_i are updated at every iteration step i

Least Square Policy Iteration (LSPI) III

- α_κ is updated every $\kappa \geq 1$ steps from the **projected Bellman equation** for **model-free policy iteration**

$$\Gamma_i \alpha_\kappa = \gamma \Lambda_i \alpha_\kappa + z_i,$$

where γ is from (3) and

$$\Gamma_0 = \beta_\Gamma I, \quad \Lambda_0 = \mathbf{0}, \quad z_0 = \mathbf{0},$$

$$\Gamma_i = \Gamma_{i-1} + \phi(x(t_i), \tau(t_i)) \phi(x(t_{i-1}), \tau(t_{i-1}))^\top,$$

$$\Lambda_i = \Lambda_{i-1} + \phi(x(t_i), \tau(t_i)) \phi(x(t_i), h(x(t_{i+1})))^\top,$$

$$z_i = z_{i-1} + \phi(x(t_i), \tau(t_i)) r(t_i),$$

where Γ_i , Λ_i and z_i are updated at every iteration step i

- new α_κ improves the Q -function (5)
- improved policies (in the sense of Problem) are obtained from (6)

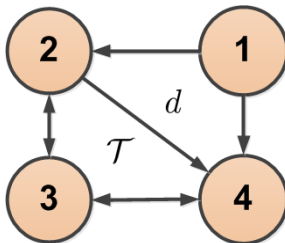
AR.Drone Parrot Quadcopter Identification

- a group of four agents with identical dynamics

$$\dot{\xi}_i = \begin{bmatrix} 0 & 1 \\ 0 & -T_p \end{bmatrix} \xi_i + \begin{bmatrix} 0 \\ K_p \end{bmatrix} u_i + \omega_i,$$
$$\zeta_i = [0.05 \quad 0.025] \xi_i,$$

where $K_p = 5.2$ and $T_p = 0.38$

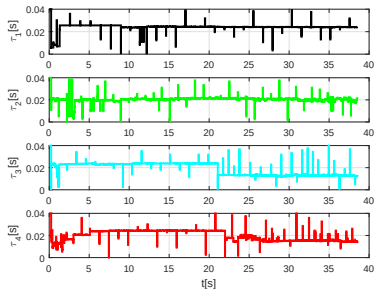
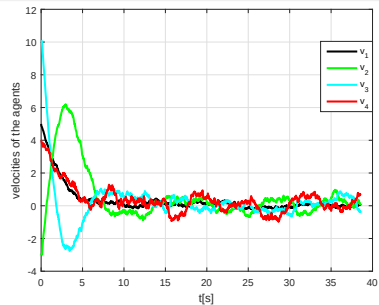
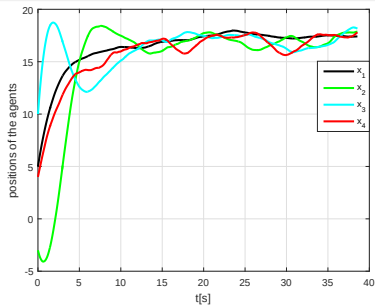
- communication delay is $d = 0.104$ s



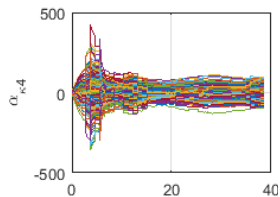
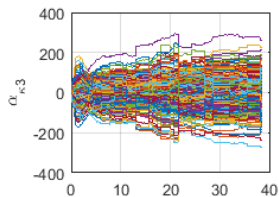
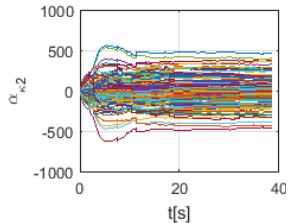
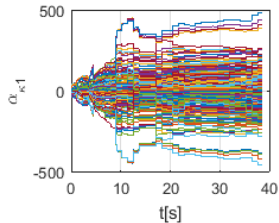
Simulation Parameters

- select $K_1 = \dots = K_4 = 0.5$ in (2)
- $\tau_i^j \in \mathcal{A} := [\underline{\tau}, \bar{\tau}]$
- the theorem yields $\bar{\tau} = 0.04$ s, while we choose $\underline{\tau} = 10^{-5}$ s
- tuning parameters for LSPI are: $\kappa = 2$ and $\varepsilon = 50$
- we choose $\mathcal{X} = [-30, 30]$
- cost function parameters: $\gamma_1 = \dots = \gamma_4 = 0.99$,
 $P_2 = P_3 = 5I_2$, $P_4 = 5I_3$, $R_1 = \dots = R_4 = 5$ and
 $S_1 = \dots = S_4 = 20$

Numeric Results I



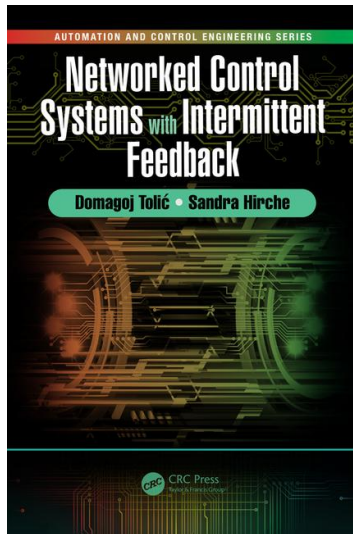
Numeric Results II



Concluding Remarks

- optimal intermittent feedback problem in MASs
- a goal function that captures local MAS performance vs. agent lifetime trade-offs
- **first**, compute provably stabilizing upper-bounds on agents' broadcasting intervals
- **second**, bring together estimation (KF) and an online model-free LSPI method to tackle coupled partially observable DP problems
- directed and unbalanced communication topologies
- large delays

References I



Thanks

This work has been supported in part by Croatian Science Foundation under the project IP-2016-06-2468 “ConDyS” and by NATO under the grant SfP - 984807 - MORUS - “Unmanned System for Maritime Security and Environmental Monitoring”.



**Questions? Comments?
Suggestions?**

**Questions? Comments?
Suggestions?**

Thank You for Your attention!!