



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



UNIVERSITÀ DEGLI STUDI  
DI CAGLIARI



# Federated Learning in Open Multi-Agent Systems via Peaceman-Rachford Splitting

Diego Deplano<sup>†</sup>, Nicola Bastianello<sup>‡</sup>, Mauro Franceschelli<sup>†</sup>, Karl H. Johansson<sup>‡</sup>

<sup>†</sup>Department of Electrical and Electronic Engineering, University of Cagliari, Italy

<sup>‡</sup> School of Electrical Engineering and Computer Science, and Digital Futures, KTH Royal Institute of Technology, Stockholm, Sweden

*IFAC World Congress 2026 – August 23-28 – Busan, Republic of Korea*

# Outline

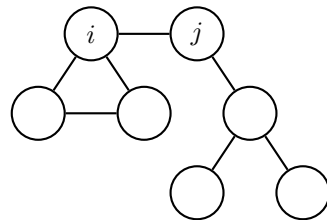
- 1 Proposed algorithm and main convergence result
- 2 Proof of the convergence result via open operator theory
- 3 Numerical simulations
- 4 Conclusions

# Outline

- 1 Proposed algorithm and main convergence result
- 2 Proof of the convergence result via open operator theory
- 3 Numerical simulations
- 4 Conclusions

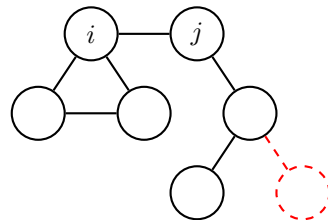
# Notation

- Undirected network  $\rightarrow \mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$
- Set of agents  $\rightarrow \mathcal{V}_k = \{1, \dots, n_k\}$
- Set of interactions  $\rightarrow \mathcal{E}_k \subseteq \mathcal{V}_k \times \mathcal{V}_k$
- State of agent  $i \rightarrow x_k^i \in \mathbb{R}$
- Neighbors of agent  $i \rightarrow \mathcal{N}_k^i = \{j \mid (i, j \in \mathcal{V}_k) \in \mathcal{E}_k\}$
- Framework  $\rightarrow$  Discrete-time  $k \in \mathbb{N}$



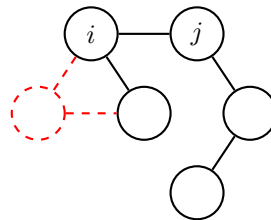
# Notation

- Undirected network  $\rightarrow \mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$
- Set of agents  $\rightarrow \mathcal{V}_k = \{1, \dots, n_k\}$
- Set of interactions  $\rightarrow \mathcal{E}_k \subseteq \mathcal{V}_k \times \mathcal{V}_k$
- State of agent  $i \rightarrow x_k^i \in \mathbb{R}$
- Neighbors of agent  $i \rightarrow \mathcal{N}_k^i = \{j \mid (i, j \in \mathcal{V}_k) \in \mathcal{E}_k\}$
- Framework  $\rightarrow$  Discrete-time  $k \in \mathbb{N}$



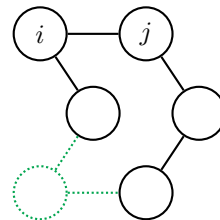
# Notation

- Undirected network  $\rightarrow \mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$
- Set of agents  $\rightarrow \mathcal{V}_k = \{1, \dots, n_k\}$
- Set of interactions  $\rightarrow \mathcal{E}_k \subseteq \mathcal{V}_k \times \mathcal{V}_k$
- State of agent  $i \rightarrow x_k^i \in \mathbb{R}$
- Neighbors of agent  $i \rightarrow \mathcal{N}_k^i = \{j \mid (i, j \in \mathcal{V}_k) \in \mathcal{E}_k\}$
- Framework  $\rightarrow$  Discrete-time  $k \in \mathbb{N}$



# Notation

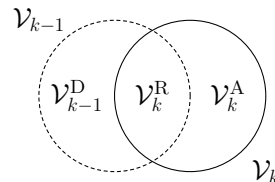
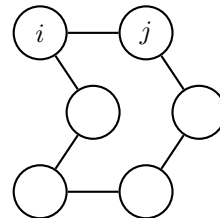
- Undirected network  $\rightarrow \mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$
- Set of agents  $\rightarrow \mathcal{V}_k = \{1, \dots, n_k\}$
- Set of interactions  $\rightarrow \mathcal{E}_k \subseteq \mathcal{V}_k \times \mathcal{V}_k$
- State of agent  $i \rightarrow x_k^i \in \mathbb{R}$
- Neighbors of agent  $i \rightarrow \mathcal{N}_k^i = \{j \mid (i, j \in \mathcal{V}_k) \in \mathcal{E}_k\}$
- Framework  $\rightarrow$  Discrete-time  $k \in \mathbb{N}$



# Notation

- Undirected network  $\rightarrow \mathcal{G}_k = (\mathcal{V}_k, \mathcal{E}_k)$
- Set of agents  $\rightarrow \mathcal{V}_k = \{1, \dots, n_k\}$
- Set of interactions  $\rightarrow \mathcal{E}_k \subseteq \mathcal{V}_k \times \mathcal{V}_k$
- State of agent  $i \rightarrow x_k^i \in \mathbb{R}$
- Neighbors of agent  $i \rightarrow \mathcal{N}_k^i = \{j \mid (i, j \in \mathcal{V}_k) \in \mathcal{E}_k\}$
- Framework  $\rightarrow$  Discrete-time  $k \in \mathbb{N}$
- Remaining agents  $\rightarrow \mathcal{V}_k^R = \mathcal{V}_k \cap \mathcal{V}_{k-1}$
- Arriving agents  $\rightarrow \mathcal{V}_k^A = \mathcal{V}_k \setminus \mathcal{V}_{k-1}$
- Departing agents  $\rightarrow \mathcal{V}_k^D = \mathcal{V}_k \cap \mathcal{V}_{k+1}$

$$x_k^i = \mathbf{T}_k^i(x_{k-1}) = \begin{cases} \mathbf{F}_k^i(x_{k-1}) & \text{if } i \in \mathcal{V}_k^R, \\ x_k^{A,i} & \text{if } i \in \mathcal{V}_k^A, \end{cases} \quad k \in \mathbb{N} \setminus \{0\},$$



## Special case of Theorem 1 in our previous work [R1]

Consider the state evolution  $x_k$  of an Open MAS ruled by  $T_k : \mathbb{R}^{p|\mathcal{V}_{k-1}|} \rightarrow \mathbb{R}^{p|\mathcal{V}_k|}$ , and split the state  $x_k = [x_k^R; x_k^A]$  into two vectors of remaining  $x_k^R$  and arriving  $x_k^A$  components. Assume that

- (a) the **standard iteration**  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the **trajectory of points of interest (TPI)** has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1})/\sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the **arrival process** is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k)/\sqrt{p|\mathcal{V}_k|} \leq H$ ;
- (d) the **departure process** is bounded with  $\beta \in (\gamma, 1)$ , i.e.,  $\sqrt{|\mathcal{V}_k|}/\sqrt{|\mathcal{V}_{k-1}|} \geq \beta$ .

Then,  $x_k$  **converges linearly** to the TPI  $\hat{x}_k$  with rate  $\theta = \gamma/\beta \in (0, 1)$  within a **radius**

$$R = \frac{B + H}{1 - \theta}. \quad (1)$$

[R1] D. Deplano, N. Bastianello, M. Franceschelli, K.H. Johansson (2026). Optimization and learning in Open Multi-Agent Systems, IEEE Transactions on Automatic Control.

**The full general theorem involves:**

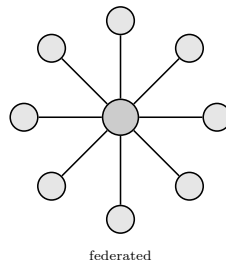
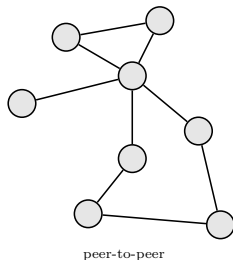
- Paracontractive operators
- Trajectory of sets of interest ( $\approx$  multiple fixed points)

## Problem of interest

We are interested in **federated** algorithms to be executed in open multi-agent systems (OMAS) to solve optimization and learning problems of the following kind

$$\min_{y \in \mathbb{R}^p} \sum_{i \in \mathcal{V}_k} f_i(y), \quad \forall k \in \mathbb{N},$$

where  $f_i : \mathbb{R}^p \rightarrow \mathbb{R}$  denotes the local cost of agent  $i$  present in the network at time  $k$ . Federated algorithms involve agents cooperate through a central coordinating entity, called the *coordinator*.



## Proposed algorithm

To solve the problem in (1) in a federated way over an open network of agents, we designed an open version of the Peaceman-Rachford Splitting (PRS), which we call *Open FedPRS*.

Open FedPRS requires each agent  $i \in \mathcal{V}_k$  to update/initialize two state variables  $x_k^i, y_k^i \in \mathbb{R}^p$  and requires the coordinator to update only a state variable  $y_k^c \in \mathbb{R}^p$ .

---

### Algorithm 1 Open FedPRS

---

**Input:** The relaxation  $\alpha \in (0, 1)$  and the penalty  $\rho > 0$ .

**Output:** Each agent returns  $y_k^i$  that is an (approximated) solution to the optimization problem in (1).

**for**  $k = 0, 1, 2, \dots$ :

**for each agent**  $i \in \mathcal{V}_k$ :

**if**  $i \in \mathcal{V}_k^A$  **is an arriving agent:**

      initializes the state variable at the local optimum

$$x_k^i = y^{i,*}$$

**else if**  $i \in \mathcal{V}_k^R$  **is a remaining agent:**

      receives  $y_{k-1}^c$  from the coordinator

      updates the state variable according to

$$y_k^i = \text{prox}_{f_i}^\rho (2y_{k-1}^c - x_{k-1}^i)$$

$$x_k^i = x_{k-1}^i + 2(y_k^i - y_{k-1}^c)$$

**end if**

      transmits  $x_k^i$  to the coordinator

**end for**

The coordinator averages the information  $x_k^i$  received from all agents  $i \in \mathcal{V}_k$  and sends it back to them:

$$y_k^c = \frac{1}{n_k} \sum_{i \in \mathcal{V}_k} x_k^i$$

**end for**

---

# Main result

## Theorem 4 of this work [R0]

Consider an OMAS executing Open FedPRS to solve a problem as in eq. (1) such that:

- (i) the local cost functions  $f_i$  are: proper, closed,  $L$ -smooth, and  $\mu$ -strongly convex;
- (ii) the distance between current  $y_k^*$  and previous  $y_{k-1}^*$  global solutions is upper bounded by  $\sigma \geq 0$ ;
- (iii) the distance between any local solution  $y^{i,*}$  and the global solution  $y_k^*$  is upper bounded by  $\omega \geq 0$ .

If  $|\mathcal{V}_k| \geq \beta |\mathcal{V}_{k-1}|$  with  $0 < \gamma \leq \beta < 1$ , then the agents' estimate  $y_k = [y_k^1, \dots, y_k^{n_k}] \in \mathbb{R}^{pn_k}$  converges with linear rate

$$\theta = \frac{\gamma}{\beta}, \quad \text{where} \quad \gamma := \max \left\{ \frac{\rho L - 1}{\rho L + 1}, \frac{1 - \rho \mu}{\rho \mu + 1} \right\}, \quad (2)$$

to consensus on the optimal solution within a distance bounded by  $R$  given next

$$R := \frac{(1 + \rho L)(\sigma + \omega)}{\sqrt{p}(1 - \theta)} \quad (3)$$

[R0] D. Deplano, N. Bastianello, M. Franceschelli, K.H. Johansson (2026). Federated Learning in Open Multi-Agent Systems via Peaceman-Rachford Splitting, IFAC WC.

# Outline

- 1 Proposed algorithm and main convergence result
- 2 Proof of the convergence result via open operator theory**
- 3 Numerical simulations
- 4 Conclusions

Conditions to be fulfilled to exploit the main convergence theorem:

- (a) the standard iteration  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the trajectory of points of interest (TPI) has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1}) / \sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k) / \sqrt{p|\mathcal{V}_k|} \leq H$ ;

#### Definition. Nonexpansiveness and contractivity

An operator  $F$  such that  $\|F(x) - F(y)\| \leq \gamma \|x - y\|$  for all  $x, y$ , is called:

- “ $\gamma$ -contractive” if  $\gamma \in (0, 1)$ ;
- “nonexpansive” if  $\gamma = 1$ .

Conditions to be fulfilled to exploit the main convergence theorem:

- (a) the standard iteration  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the trajectory of points of interest (TPI) has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1}) / \sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k) / \sqrt{p|\mathcal{V}_k|} \leq H$ ;

### Definition. Nonexpansiveness and contractivity

An operator  $F$  such that  $\|F(x) - F(y)\| \leq \gamma \|x - y\|$  for all  $x, y$ , is called:

- “ $\gamma$ -contractive” if  $\gamma \in (0, 1)$ ;
- “nonexpansive” if  $\gamma = 1$ .

The federated PRS in open multi-agent systems (called **Open FedPRS**) can be written as follows:

$$y_k^i = \text{prox}_{f_i}^{\rho} (2y_{k-1}^c - x_{k-1}^i) \quad \text{if } i \in \mathcal{V}_{k+1}^R$$

$$x_k^i = T_k^i(x_{k-1}) = \begin{cases} x_{k-1}^i + 2(y_k^i - y_{k-1}^c) & \text{if } i \in \mathcal{V}_{k+1}^R, \\ y^{i,*} & \text{if } i \in \mathcal{V}_{k+1}^A, \end{cases}$$

where:

- the local coordinator computes and transmit at each step the following aggregate information

$$y_k^c = \frac{1}{|\mathcal{V}_k|} \sum_{i \in \mathcal{V}_k} x_k^i ;$$

- the proximal operator and reflective operators are defined by

$$\text{prox}_{f_i}^{\rho}(w) := \underset{z}{\operatorname{argmin}} \left\{ f_i(z) + \frac{1}{2} \|z - w\|_2^2 \right\}, \quad \text{refl}_{f_i}^{\rho}(w) := 2 \text{prox}_{f_i}^{\rho}(w) - w;$$

- $T_k = [\dots, T_k^i, \dots]^T$  is an *open operator* because it produces sequences  $\{x_k\}_{k \in \mathbb{N}}$  of points of different dimension and/or composition.

**Definition.** Open operators and standard iterations in open multi-agent systems

In an OMAS, the iteration of the open operator  $T_k$  can be written component-wise for each  $i \in \mathcal{V}_k$ , as follows:

$$x_k^i = T_k^i(x_{k-1}) = \begin{cases} F_k^i(x_{k-1}) & \text{if } i \in \mathcal{V}_k^R \\ x_k^{A,i} & \text{if } i \in \mathcal{V}_k^A. \end{cases}$$

If the set of agents remain unchanged, that is,  $\mathcal{V}_k = \mathcal{V}_{k-1}$ , then the operator  $F_k = [\dots, F_k^i, \dots]^\top$  rules the so-called *standard iteration*:

$$x_k = F_k(x_{k-1}), \quad \text{when } \mathcal{V}_k = \mathcal{V}_{k-1}.$$

Let  $J_k = |\mathcal{V}_k|^{-1}(\mathbf{1}_{|\mathcal{V}_k|} \mathbf{1}_{|\mathcal{V}_k|}^\top \otimes I_p)$ ,  $\tilde{y}_k^c = \mathbf{1}_{n_k} \otimes y_k^c$ ,  $f_k(y) = \sum_{i \in \mathcal{V}_k} f_i(y)$ .

Then, the standard iteration of Open FedPRS is given by

$$\begin{aligned} F_k(x_{k-1}) &= x_{k-1} + 2(y_k - \tilde{y}_{k-1}^c) = 2y_k - (2\tilde{y}_{k-1}^c - x_{k-1}) \\ &= 2 \operatorname{prox}_{f_i}^\rho(2y_{k-1}^c - x_{k-1}) - (2\tilde{y}_{k-1}^c - x_{k-1}) \\ &= \operatorname{reft}_{f_k}^\rho(2\tilde{y}_{k-1}^c - x_{k-1}) = \operatorname{reft}_{f_k}^\rho(2J_{k-1}x_{k-1} - x_{k-1}) \\ &= \operatorname{reft}_{f_k}^\rho((2J_{k-1} - I)x_{k-1}) \end{aligned}$$

**Definition.** Open operators and standard iterations in open multi-agent systems

In an OMAS, the iteration of the open operator  $T_k$  can be written component-wise for each  $i \in \mathcal{V}_k$ , as follows:

$$x_k^i = T_k^i(x_{k-1}) = \begin{cases} F_k^i(x_{k-1}) & \text{if } i \in \mathcal{V}_k^R \\ x_k^{A,i} & \text{if } i \in \mathcal{V}_k^A. \end{cases}$$

If the set of agents remain unchanged, that is,  $\mathcal{V}_k = \mathcal{V}_{k-1}$ , then the operator  $F_k = [\dots, F_k^i, \dots]^\top$  rules the so-called *standard iteration*:

$$x_k = F_k(x_{k-1}), \quad \text{when } \mathcal{V}_k = \mathcal{V}_{k-1}.$$

Let  $J_k = |\mathcal{V}_k|^{-1}(\mathbf{1}_{|\mathcal{V}_k|} \mathbf{1}_{|\mathcal{V}_k|}^\top \otimes I_p)$ ,  $\tilde{y}_k^c = \mathbf{1}_{n_k} \otimes y_k^c$ ,  $f_k(y) = \sum_{i \in \mathcal{V}_k} f_i(y)$ .

Then, the standard iteration of Open FedPRS is given by

$$\begin{aligned} F_k(x_{k-1}) &= x_{k-1} + 2(y_k - \tilde{y}_{k-1}^c) = 2y_k - (2\tilde{y}_{k-1}^c - x_{k-1}) \\ &= 2 \operatorname{prox}_{f_i}^\rho(2y_{k-1}^c - x_{k-1}) - (2\tilde{y}_{k-1}^c - x_{k-1}) \\ &= \operatorname{refl}_{f_k}^\rho(2\tilde{y}_{k-1}^c - x_{k-1}) = \operatorname{refl}_{f_k}^\rho(2J_{k-1}x_{k-1} - x_{k-1}) \\ &= \operatorname{refl}_{f_k}^\rho((2J_{k-1} - I)x_{k-1}) \end{aligned}$$

**Proof of condition (a)** (the standard iteration is contractive with  $\gamma \in (0, 1)$ ).

To satisfy this condition is sufficient to show that the operator  $F_k$  is contractive.

- ①  $F_k$  is the composition of the reflective operator  $\text{refl}_f^\rho$  with a linear operator  $\mathcal{J}_{k-1} : x \mapsto (2J_{k-1} - I)x$ ;
- ② The operator  $\mathcal{J}_{k-1}$  is nonexpansive because all eigenvalues of  $2J_{k-1} - I$  are within the unit circle (see Lemma 3 in [R3]): the matrix  $|\mathcal{V}_k|^{-1}(\mathbf{1}_{|\mathcal{V}_k|}\mathbf{1}_{|\mathcal{V}_k|}^\top \otimes I_p)$  has all eigenvalues equal to 0 but one that is equal to 1 and the Kronecker product with the identity matrix  $I_p$  leaves the eigenvalues unchanged;
- ③ The reflected operator  $\text{refl}_f^\rho$  is  $\gamma$ -contractive when  $f$  is  $\mu$ -strongly convex and  $L$ -smooth, and the contraction rate  $\gamma := \max\left\{\frac{\rho L - 1}{\rho L + 1}, \frac{1 - \rho m}{1 + \rho m}\right\}$  holds by Theorem 1 in [R2];
- ④ Since  $F_k$  is the composition of a nonexpansive operator with a  $\gamma$ -contractive operator, then  $F_k$  is  $\gamma$ -contractive.

**Remark.** The constant  $\gamma$  already holds at any time  $k \in \mathbb{N}$  and it does not depend on the network topology due to the presence of the coordinator.

**Conclusion:** The contraction factor  $\gamma$  holds uniformly.

---

[R2] Giselsson P and Boyd S (2017). Linear Convergence and Metric Selection for Douglas-Rachford Splitting and ADMM, IEEE Transactions on Automatic Control.

[R3] Belgioioso G, Fabiani F, Blanchini F, and Grammatico S (2018). On the Convergence of Discrete-Time Linear Systems: A Linear Time-Varying Mann Iteration Converges IFF Its Operator Is Strictly Pseudocontractive, IEEE Control Systems Letters.

Conditions to be fulfilled to exploit the main convergence theorem:

- (a) the standard iteration  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the trajectory of points of interest (TPI) has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1}) / \sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k) / \sqrt{p|\mathcal{V}_k|} \leq H$ ;

**Definition.** Trajectory of points of interest

Consider an OMAS and assume that the standard dynamics has a unique solution  $\hat{x}_k$  at each time  $k$ ,

$$\hat{x}_{k+1} = F_{k+1}(\hat{x}_k).$$

The open sequence  $\{\hat{x}_k\}_{k \in \mathbb{N}}$  is called the “trajectory of points of interest” (TPI) of the OMAS.

**Remark:** For autonomous and closed system, this definition boils down to that of an equilibrium point.

Conditions to be fulfilled to exploit the main convergence theorem:

- (a) the standard iteration  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the trajectory of points of interest (TPI) has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1}) / \sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k) / \sqrt{p|\mathcal{V}_k|} \leq H$ ;

### Definition. Trajectory of points of interest

Consider an OMAS and assume that the standard dynamics has a unique solution  $\hat{x}_k$  at each time  $k$ ,

$$\hat{x}_k = F_{k+1}(\hat{x}_k).$$

The open sequence  $\{\hat{x}_k\}_{k \in \mathbb{N}}$  is called the “trajectory of points of interest” (TPI) of the OMAS.

**Remark:** For autonomous and closed system, this definition boils down to that of an equilibrium point.

**Definition.** Open distance between points of interest

In an OMAS, the open distance between points of interest at different times (with possibly different dimension and composition) is given by

$$d(\hat{x}_k, \hat{x}_{k-1}) := \sqrt{\sum_{i \in \mathcal{V}_k^R} \|\hat{x}_k^i - \hat{x}_{k-1}^i\|^2},$$

i.e., only the state of remaining agents contribute to the total distance.

At a fixed point  $\hat{x}_k$  of the standard iteration, the coordinator's output satisfies  $J_k \hat{x}_k = \mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^*$  and, by the update rule, it holds  $J_k \hat{x}_k = \text{prox}_f^\rho((2J_k - I)\hat{x}_k)$ . Since the proximal operator is characterized by  $a = \text{prox}_{f_k}^\rho(b) \Leftrightarrow b - a = \nabla f_k(a)$ , then:

$$(2J_k - I)\hat{x}_k - J_k \hat{x}_k = \rho \nabla f_k(J_k \hat{x}_k)$$

$$\hat{x}_k = J_k \hat{x}_k - \rho \nabla f_k(J_k \hat{x}_k)$$

$$\hat{x}_k = \mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^* - \rho \nabla f_k(\mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^*).$$

Component-wise:  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*)$ ,  $\forall i \in \mathcal{V}_k$ .

**Definition.** Open distance between points of interest

In an OMAS, the open distance between points of interest at different times (with possibly different dimension and composition) is given by

$$d(\hat{x}_k, \hat{x}_{k-1}) := \sqrt{\sum_{i \in \mathcal{V}_k^R} \|\hat{x}_k^i - \hat{x}_{k-1}^i\|^2},$$

i.e., only the state of remaining agents contribute to the total distance.

At a fixed point  $\hat{x}_k$  of the standard iteration, the coordinator's output satisfies  $J_k \hat{x}_k = \mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^*$  and, by the update rule, it holds  $J_k \hat{x}_k = \text{prox}_f^\rho((2J_k - I)\hat{x}_k)$ . Since the proximal operator is characterized by  $a = \text{prox}_{f_k}^\rho(b) \Leftrightarrow b - a = \nabla f_k(a)$ , then:

$$\begin{aligned} (2J_k - I)\hat{x}_k - J_k \hat{x}_k &= \rho \nabla f_k(J_k \hat{x}_k) \\ \hat{x}_k &= J_k \hat{x}_k - \rho \nabla f_k(J_k \hat{x}_k) \\ \hat{x}_k &= \mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^* - \rho \nabla f_k(\mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^*). \end{aligned}$$

Component-wise:  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*)$ ,  $\forall i \in \mathcal{V}_k$ .

**Proof of condition (b)** (the TPI has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1})/\sqrt{p|\mathcal{V}_k|} \leq B$ ).

**Step 1: characterize the points of interest.** Component-wise it holds  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*) \quad \forall i \in \mathcal{V}_k$ .

**Step 2: bound for remaining agents.** For each remaining agent  $i \in \mathcal{V}_k^R$ , it holds:

$$\begin{aligned} \|\hat{x}_k^i - \hat{x}_{k-1}^i\| &= \|y_k^* - y_{k-1}^* - \rho(\nabla f_i(y_k^*) - \nabla f_i(y_{k-1}^*))\| \\ &\stackrel{(i)}{\leq} \|y_k^* - y_{k-1}^*\| + \rho \|\nabla f_i(y_k^*) - \nabla f_i(y_{k-1}^*)\| \\ &\stackrel{(ii)}{\leq} (1 + \rho L) \|y_k^* - y_{k-1}^*\| \stackrel{(iii)}{\leq} (1 + \rho L) \sigma, \end{aligned}$$

where (i) triangle inequality; (ii)  $L$ -smoothness of  $f_i$ ; (iii) assumption (ii).

**Step 3: normalize.** Since  $|\mathcal{V}_k^R| \leq |\mathcal{V}_k|$ :

$$\frac{d(\hat{x}_k, \hat{x}_{k-1})}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{1}{\sqrt{p|\mathcal{V}_k|}} \sqrt{\sum_{i \in \mathcal{V}_k} \|\hat{x}_k^i - \hat{x}_{k-1}^i\|^2} \leq \frac{(1 + \rho L)\sigma}{\sqrt{p}} =: B.$$

**Proof of condition (b)** (the TPI has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1})/\sqrt{p|\mathcal{V}_k|} \leq B$ ).

**Step 1: characterize the points of interest.** Component-wise it holds  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*) \quad \forall i \in \mathcal{V}_k$ .

**Step 2: bound for remaining agents.** For each remaining agent  $i \in \mathcal{V}_k^R$ , it holds:

$$\begin{aligned} \|\hat{x}_k^i - \hat{x}_{k-1}^i\| &= \|y_k^* - y_{k-1}^* - \rho(\nabla f_i(y_k^*) - \nabla f_i(y_{k-1}^*))\| \\ &\stackrel{(i)}{\leq} \|y_k^* - y_{k-1}^*\| + \rho \|\nabla f_i(y_k^*) - \nabla f_i(y_{k-1}^*)\| \\ &\stackrel{(ii)}{\leq} (1 + \rho L) \|y_k^* - y_{k-1}^*\| \stackrel{(iii)}{\leq} (1 + \rho L) \sigma, \end{aligned}$$

where (i) triangle inequality; (ii)  $L$ -smoothness of  $f_i$ ; (iii) assumption (ii).

**Step 3: normalize.** Since  $|\mathcal{V}_k^R| \leq |\mathcal{V}_k|$ :

$$\frac{d(\hat{x}_k, \hat{x}_{k-1})}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{1}{\sqrt{p|\mathcal{V}_k|}} \sqrt{\sum_{i \in \mathcal{V}_k} \|\hat{x}_k^i - \hat{x}_{k-1}^i\|^2} \leq \frac{(1 + \rho L) \sigma}{\sqrt{p}} =: B.$$

**Proof of condition (b)** (the TPI has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1})/\sqrt{p|\mathcal{V}_k|} \leq B$ ).

**Step 1: characterize the points of interest.** Component-wise it holds  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*) \quad \forall i \in \mathcal{V}_k$ .

**Step 2: bound for remaining agents.** For each remaining agent  $i \in \mathcal{V}_k^R$ , it holds:

$$\begin{aligned} \|\hat{x}_k^i - \hat{x}_{k-1}^i\| &= \|y_k^* - y_{k-1}^* - \rho(\nabla f_i(y_k^*) - \nabla f_i(y_{k-1}^*))\| \\ &\stackrel{(i)}{\leq} \|y_k^* - y_{k-1}^*\| + \rho \|\nabla f_i(y_k^*) - \nabla f_i(y_{k-1}^*)\| \\ &\stackrel{(ii)}{\leq} (1 + \rho L) \|y_k^* - y_{k-1}^*\| \stackrel{(iii)}{\leq} (1 + \rho L) \sigma, \end{aligned}$$

where (i) triangle inequality; (ii)  $L$ -smoothness of  $f_i$ ; (iii) assumption (ii).

**Step 3: normalize.** Since  $|\mathcal{V}_k^R| \leq |\mathcal{V}_k|$ :

$$\frac{d(\hat{x}_k, \hat{x}_{k-1})}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{1}{\sqrt{p|\mathcal{V}_k|}} \sqrt{\sum_{i \in \mathcal{V}_k} \|\hat{x}_k^i - \hat{x}_{k-1}^i\|^2} \leq \frac{(1 + \rho L)\sigma}{\sqrt{p}} =: B.$$

Conditions to be fulfilled to exploit the main convergence theorem:

- (a) the standard iteration  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the trajectory of points of interest (TPI) has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1}) / \sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k) / \sqrt{p|\mathcal{V}_k|} \leq H$ ;

**Definition.** Open distance between arrival states and the point of interest

In an OMAS, the open distance between the vector  $x_k^A$  stacking the state variables of arriving agents and the point of interest  $\hat{x}_k$  (with possibly different dimension and composition) is given by

$$d(x_k^A, \hat{x}_k) := \sqrt{\sum_{i \in \mathcal{V}_k^A} \|x_k^{A,i} - \hat{x}_k^i\|^2},$$

i.e., only the state of arriving agents contribute to the total distance.

Conditions to be fulfilled to exploit the main convergence theorem:

- (a) the standard iteration  $F_k$  is contractive with  $\gamma \in (0, 1)$ ;
- (b) the trajectory of points of interest (TPI) has bounded variation  $B \geq 0$ , i.e.,  $d(\hat{x}_k, \hat{x}_{k-1}) / \sqrt{p|\mathcal{V}_k|} \leq B$ ;
- (c) the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k) / \sqrt{p|\mathcal{V}_k|} \leq H$ ;

**Definition.** Open distance between arrival states and the point of interest

In an OMAS, the open distance between the vector  $x_k^A$  stacking the state variables of arriving agents and the point of interest  $\hat{x}_k$  (with possibly different dimension and composition) is given by

$$d(x_k^A, \hat{x}_k) := \sqrt{\sum_{i \in \mathcal{V}_k^A} \|x_k^{A,i} - \hat{x}_k^i\|^2},$$

i.e., only the state of arriving agents contribute to the total distance.

**Proof of condition (c)** (the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k)/\sqrt{p|\mathcal{V}_k|} \leq H$ ).

Open FedPRS initializes  $x_k^i = y^{i,*}$  for arriving agents  $i \in \mathcal{V}_k^A$ , while from the fixed-point characterization:  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*)$ .

**Step 1: bound per-agent distance.** Since  $\nabla f_i(y^{i,*}) = 0$  (local minimizer), for  $i \in \mathcal{V}_k^A$ :

$$\begin{aligned} \|x_k^i - \hat{x}_k^i\| &= \|y^{i,*} - y_k^* + \rho \nabla f_i(y_k^*)\| \\ &\stackrel{(i)}{=} \|y^{i,*} - y_k^* - \rho(\nabla f_i(y^{i,*}) - \nabla f_i(y_k^*))\| \\ &\stackrel{(ii)}{\leq} \|y^{i,*} - y_k^*\| + \rho \|\nabla f_i(y^{i,*}) - \nabla f_i(y_k^*)\| \\ &\stackrel{(iii)}{\leq} (1 + \rho L) \|y^{i,*} - y_k^*\| \stackrel{(iv)}{\leq} (1 + \rho L) \omega, \end{aligned}$$

where (i) holds by  $\nabla f_i(y^{i,*}) = 0$ ; (ii) by triangle inequality; (iii) by  $L$ -smoothness of  $f_i$ ; (iv) by assumption (iii).

**Step 2: normalize.** Since  $|\mathcal{V}_k^A| \leq |\mathcal{V}_k|$ :

$$\frac{d(x_k^A, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{1}{\sqrt{p|\mathcal{V}_k|}} \sqrt{\sum_{i \in \mathcal{V}_k} \|x_k^i - \hat{x}_k^i\|^2} \leq \frac{(1 + \rho L)\omega}{\sqrt{p}} =: H.$$

**Proof of condition (c)** (the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k)/\sqrt{p|\mathcal{V}_k|} \leq H$ ).

Open FedPRS initializes  $x_k^i = y^{i,*}$  for arriving agents  $i \in \mathcal{V}_k^A$ , while from the fixed-point characterization:  $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*)$ .

**Step 1: bound per-agent distance.** Since  $\nabla f_i(y^{i,*}) = 0$  (local minimizer), for  $i \in \mathcal{V}_k^A$ :

$$\begin{aligned} \|x_k^i - \hat{x}_k^i\| &= \|y^{i,*} - y_k^* + \rho \nabla f_i(y_k^*)\| \\ &\stackrel{(i)}{=} \|y^{i,*} - y_k^* - \rho(\nabla f_i(y^{i,*}) - \nabla f_i(y_k^*))\| \\ &\stackrel{(ii)}{\leq} \|y^{i,*} - y_k^*\| + \rho \|\nabla f_i(y^{i,*}) - \nabla f_i(y_k^*)\| \\ &\stackrel{(iii)}{\leq} (1 + \rho L) \|y^{i,*} - y_k^*\| \stackrel{(iv)}{\leq} (1 + \rho L) \omega, \end{aligned}$$

where (i) holds by  $\nabla f_i(y^{i,*}) = 0$ ; (ii) by triangle inequality; (iii) by  $L$ -smoothness of  $f_i$ ; (iv) by assumption (iii).

**Step 2: normalize.** Since  $|\mathcal{V}_k^A| \leq |\mathcal{V}_k|$ :

$$\frac{d(x_k^A, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{1}{\sqrt{p|\mathcal{V}_k|}} \sqrt{\sum_{i \in \mathcal{V}_k} \|x_k^i - \hat{x}_k^i\|^2} \leq \frac{(1 + \rho L)\omega}{\sqrt{p}} =: H.$$

**Proof of condition (c)** (the arrival process is bounded with  $H \geq 0$ , i.e.,  $d(x_k^A, \hat{x}_k)/\sqrt{p|\mathcal{V}_k|} \leq H$ ).

Open FedPRS initializes  $x_k^i = y^{i,*}$  for arriving agents  $i \in \mathcal{V}_k^A$ , while from the fixed-point characterization:  
 $\hat{x}_k^i = y_k^* - \rho \nabla f_i(y_k^*)$ .

**Step 1: bound per-agent distance.** Since  $\nabla f_i(y^{i,*}) = 0$  (local minimizer), for  $i \in \mathcal{V}_k^A$ :

$$\begin{aligned} \|x_k^i - \hat{x}_k^i\| &= \|y^{i,*} - y_k^* + \rho \nabla f_i(y_k^*)\| \\ &\stackrel{(i)}{=} \|y^{i,*} - y_k^* - \rho(\nabla f_i(y^{i,*}) - \nabla f_i(y_k^*))\| \\ &\stackrel{(ii)}{\leq} \|y^{i,*} - y_k^*\| + \rho \|\nabla f_i(y^{i,*}) - \nabla f_i(y_k^*)\| \\ &\stackrel{(iii)}{\leq} (1 + \rho L) \|y^{i,*} - y_k^*\| \stackrel{(iv)}{\leq} (1 + \rho L) \omega, \end{aligned}$$

where (i) holds by  $\nabla f_i(y^{i,*}) = 0$ ; (ii) by triangle inequality; (iii) by  $L$ -smoothness of  $f_i$ ; (iv) by assumption (iii).

**Step 2: normalize.** Since  $|\mathcal{V}_k^A| \leq |\mathcal{V}_k|$ :

$$\frac{d(x_k^A, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{1}{\sqrt{p|\mathcal{V}_k|}} \sqrt{\sum_{i \in \mathcal{V}_k} \|x_k^i - \hat{x}_k^i\|^2} \leq \frac{(1 + \rho L)\omega}{\sqrt{p}} =: H.$$

## Conclusion of the proof.

Having verified conditions (a), (b), (c), and since (d) holds by assumption, then by Theorem 1 in [R1] it holds:

$$\limsup_{k \rightarrow \infty} \frac{d(x_k, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{(1 + \rho L)(\sigma + \omega)}{\sqrt{p}(1 - \theta)} =: R$$

We now show that the distance of the estimation vector  $y_k = [\dots, y_k^{i^\top}, \dots]^\top$  of the whole network from the consensus state on the solution  $c_k^* = \mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^*$  is upper bounded by  $d(x_k, \hat{x}_k)$  as follows:

$$\begin{aligned} d(y_k, c_k^*) &= \|y_k - (\mathbf{1}_{n_k} \otimes y_k^*)\| \\ &\stackrel{(i)}{=} \left\| \text{prox}_{f_k}^\rho((2J_k - I)x_{k-1}) - \text{prox}_{f_k}^\rho((2J_k - I)\hat{x}_{k-1}) \right\| \\ &\stackrel{(ii)}{\leq} \|(2J_k - I)(x_k - \hat{x}_k)\| \stackrel{(iii)}{\leq} \|(2J_k - I)\| d(x_k, \hat{x}_k) \stackrel{(iv)}{=} d(x_k, \hat{x}_k) \end{aligned}$$

where (i) holds by the specific update rule of Open FedPRS and by the relation between the fixed point and the optimal solution; (ii) follows by the non-expansiveness of the proximal; (iii) holds by the sum-multiplicative property; (iv) holds because matrix  $2J_k - I$  is doubly stochastic. The proof is then completed by:

$$\limsup_{k \rightarrow \infty} \frac{d(y_k, c_k^*)}{\sqrt{p|\mathcal{V}_k|}} \leq \limsup_{k \rightarrow \infty} \frac{d(x_k, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq R.$$

[R1] D. Deplano, N. Bastianello, M. Franceschelli, K.H. Johansson (2026). Optimization and learning in Open Multi-Agent Systems, IEEE Transactions on Automatic Control.

## Conclusion of the proof.

Having verified conditions (a), (b), (c), and since (d) holds by assumption, then by Theorem 1 in [R1] it holds:

$$\limsup_{k \rightarrow \infty} \frac{d(x_k, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq \frac{(1 + \rho L)(\sigma + \omega)}{\sqrt{p}(1 - \theta)} =: R$$

We now show that the distance of the estimation vector  $y_k = [\dots, y_k^{i^\top}, \dots]^\top$  of the whole network from the consensus state on the solution  $c_k^* = \mathbf{1}_{|\mathcal{V}_k|} \otimes y_k^*$  is upper bounded by  $d(x_k, \hat{x}_k)$  as follows:

$$\begin{aligned} d(y_k, c_k^*) &= \|y_k - (\mathbf{1}_{n_k} \otimes y_k^*)\| \\ &\stackrel{(i)}{=} \|\text{prox}_{f_k}^\rho((2J_k - I)x_{k-1}) - \text{prox}_{f_k}^\rho((2J_k - I)\hat{x}_{k-1})\| \\ &\stackrel{(ii)}{\leq} \|(2J_k - I)(x_k - \hat{x}_k)\| \stackrel{(iii)}{\leq} \|(2J_k - I)\| d(x_k, \hat{x}_k) \stackrel{(iv)}{=} d(x_k, \hat{x}_k) \end{aligned}$$

where (i) holds by the specific update rule of Open FedPRS and by the relation between the fixed point and the optimal solution; (ii) follows by the non-expansiveness of the proximal; (iii) holds by the sum-multiplicative property; (iv) holds because matrix  $2J_k - I$  is doubly stochastic. The proof is then completed by:

$$\limsup_{k \rightarrow \infty} \frac{d(y_k, c_k^*)}{\sqrt{p|\mathcal{V}_k|}} \leq \limsup_{k \rightarrow \infty} \frac{d(x_k, \hat{x}_k)}{\sqrt{p|\mathcal{V}_k|}} \leq R.$$

---

[R1] D. Deplano, N. Bastianello, M. Franceschelli, K.H. Johansson (2026). Optimization and learning in Open Multi-Agent Systems, IEEE Transactions on Automatic Control.

# Outline

- ① Proposed algorithm and main convergence result
- ② Proof of the convergence result via open operator theory
- ③ Numerical simulations**
- ④ Conclusions

We have tested Open FedPRS in classification problems characterized by the local costs

$$f_i(y) = \frac{1}{m_i} \sum_{j=1}^{m_i} \log \left( 1 + \exp \left( -b_{i,j} a_{i,j}^\top y \right) \right) + \frac{\epsilon}{2} \|y\|^2$$

where  $a_{i,j} \in \mathbb{R}^p$  and  $b_{i,j} \in \{-1, 1\}$  are pairs of feature vector and label, randomly generated with `scikit-learn`. In particular,  $p = 5$ ,  $m_i = 50$ , and  $\epsilon = 0.05$ .

The results will be presented in terms of the normalized distance  $e_k$  between the local models and the consensus set on the optimal model as in (3) and in terms of the normalized distance  $\varepsilon_k$  between the coordinator model and the optimal model:

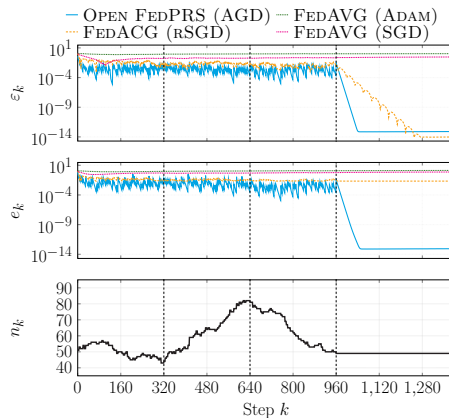
$$e_k := \frac{d(y_k, c_k^*)}{\sqrt{pn_k}}, \quad \varepsilon_k = \frac{d(y_k^c, y_k^*)}{\sqrt{p}}.$$

We benchmark Open FedPRS against two state-of-the-art algorithms in federated learning over open networks:

- FedAVG [R4] (one of the pioneering methods that introduced the federated learning paradigm);
- FedACG [R5] (arguably one of the most competitive approach currently available in the literature)

Since these algorithms were originally designed for closed (static) networks, we adapt them to the open-network setting by incorporating the mechanisms proposed in [R6] and [R7], respectively.

The chosen local optimizers are accelerated gradient descent (AGD) for Open FedPRS, stochastic gradient descent (SGD) and Adam for FedAVG, and regularized SGD (rSGD) for FedACG.



- [R4] B. McMahan, E. Moore, D. Ramage, S. B.A. y Arcas (2017). Communication-efficient learning of deep networks from decentralized data, Artificial intelligence and statistics.
- [R5] G. Kim, J. Kim, B. Han (2024). Communication-efficient federated learning with accelerated client gradient, IEEE/CVF Conference on Computer Vision and Pattern Recognition.
- [R6] Y. Sun, Y. Mao, J. Zhang (2024). MimiC: Combating Client Dropouts in Federated Learning by Mimicking Central Updates, IEEE Transactions on Mobile Computing.
- [R7] Z.L. Chang, D.J. Han, R. Parasnis, S. Hosseinalipour, C.G. Brinton (2024). Federated learning with dynamic client arrival and departure: Convergence and rapid adaptation via initial model construction.

# Outline

- 1 Proposed algorithm and main convergence result
- 2 Proof of the convergence result via open operator theory
- 3 Numerical simulations
- 4 Conclusions**

General framework for the results in this talk: [Deplano, D., Bastianello, N., Franceschelli, M., & Johansson, K.H. \(2026\). Optimization and learning in open multi-agent systems. IEEE Transactions on Automatic Control.](#)

Ongoing applications within this framework and its extensions:

- **Estimation and tracking of time-varying graph quantities**

Deplano, D., Franceschelli, M., & Giua, A. (2025). Distributed Tracking of Network Size, Diameter, Radius, and Node Eccentricities in Open Multi-Agent Systems. IEEE Transactions on Automatic Control.

- **Composite optimization under additive errors**

Deplano, D., Bastianello, N., Franceschelli, M., & Johansson, K. H. (2026). Composite Optimization in Open Multi-Agent Systems under Additive Errors. IEEE Control Systems Letters.

- **Potential-based coordination**

Miele, A., Lippi, M., Deplano, D., Franceschelli, M., & Gasparri, A. (2026). Distributed Potential-Based Coordination within Open Multi-Agent Systems. IEEE Control Systems Letters.

- **Centroid estimation in multi-robot systems**

Miele, A., Franceschelli, M., & Gasparri, A. (2025). Gossip-based estimation of centroid and common reference frame in open multi-agent systems. IEEE Control Systems Letters.



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



UNIVERSITÀ DEGLI STUDI  
DI CAGLIARI



# Federated Learning in Open Multi-Agent Systems via Peaceman-Rachford Splitting

Thank you for your attention!

**Diego Deplano**

**Email:** [diego.deplano@unica.it](mailto:diego.deplano@unica.it)

**Webpage:** <https://diegodeplano.github.io/>



**SCAN ME**