Two-Level Sensorimotor Learning for Leader-Follower Consensus Control

J. Franco-Robles, J. Escareno, O. Labbani-Igbida

Two-Level Sensorimotor Learning for Leader-Follower Consensus Control

J. Franco-Robles*[1,2], J. Escareno[2] and O. Labbani-Igbida[2]

***Abstract*— The present work addresses the problem of leader-follower consensus based on a bioinspired sensorymotor approach. The herein presented learning scheme entails two levels: (*i*) a supervised offline *babbling* training, where babbling generates a preliminary inference of the unknown environment, and (*ii*) during the consensus the execution (online) stage the cortical map is re-training throughout agents, which simultaneously to the consensus lapse the learning model is refined. The purpose of the proposed strategy is to link the human sensorimotor postural model with the consensus problem to endow of natural plasticity to the MAS. In order to fullfil the leader-follower control objective a controller based Lyapunov stability theory is synthesized. A set of numerical simulations are conducted to evaluate the MAS performance while following the cortical-mapped leader trajectory.**

I. INTRODUCTION

In the last decade, it has been notorious for the surge of Unmanned Aircraft Systems (UAS) enhancing and upgrading various applications. Well-known UAS-based missions are staging part of civilian tasks, e.g., disaster rescue [1], object tracking [2], wildlife protection [3], remote delivery [4], among others. A significant subset of these developments has focused on multi-agent systems (MAS) under fascinating, and widely studied topics, such as formation and consensus characterized as control of agents to achieve prescribed restrictions in their states [5]. In [6], a remarkable analysis of formation control is presented, deriving into position-, displacement-, and distance-based concerning sensed and controlled variables. At the same time, the consensus is defined as a particular class of formation control. In particular, the quintessence of the consensus algorithm is that each unmanned aircraft vehicle (UAV) updates its information state based on the state of the local neighbor in such a way that information of each UAV converges towards the desired value [7]. The consensus problem is a highly challenging issue in distributed control and is approached as sampled-data-, quantized-, random networks-, leader-follower-, finite-time-, bipartite-, group cluster-, and scaled-consensus.

Leader-follower consensus is an engaging topic where the followers have an independent motion and receive instructions from a real or a virtual leader. To address the leader-follower consensus problem, various strategies have been studied, e.g, full state feedback control [8], observed-type protocol [9], prescribed performance control [10], multiple leaders communication protocol [11], and artificial potential fields [12]. As a complementary study to control models, intelligent approaches have been extensively utilized to simplify challenging consensus tasks from hybrid control systems to creating a completely end-to-end bio-inspired model [13]. Regarding intelligent leader-follower consensus, the authors in [14] explore the recurrent neural networks approach to learn the system uncertainties and consensus tracking of a MAS. In [15] the unknown control direction, unknown system dynamics, and unknown external disturbance are addressed with a radial basis neural network and an adaptive method in a MAS. Moreover, in [11], a reinforcement learning with a Q-learning approach concerning a value function based on state error for the direct graph minimizing the energy assumption is applied. The bio-inspired results encourage our research to construct a coherent brain-inspired map capable of associating the leader-follower consensus regarding the sensory-motor neurological system since the central nervous system shows an astounding ability to control and link perception with joints.

A. Related work

The brain's multiple association capacity and information acquisition create the bases for agents' collective control. Inside the brain cortex, the correlation between the activity of the efferent motor patterns and the afferent patterns perceived by the sensory cortex with the rest of the nervous system is structured in a sensory-motor acquisition. The authors in [16], introduce the process defined as body schema (BoS) based on the organizational hierarchy of the sensory-motor system regarding the posture of the body and its perception in a dynamic nature. Furthermore, the nervous system's extraordinary ability to transform coordinates between different frames of reference in the human body has been widely investigated by roboticists [17].

In [18], the authors present a path planning algorithm based on BoS capable of driving a micro aerial vehicle (MAV) in a controlled space in addition to calculating its inverse kinematics through potential fields as a function of the desired position of the MAV. In [19], the authors propose a new approach based on BoS capable of acquiring the position of different UAVs in different configuration spaces using a single sensor space. The authors create independent potential fields where desired positions and obstacles are established for each UAV, allowing it to maintain the desired formation and collision between agents. Despite the limited research in BoS related to MAS, preliminary results in formation control motivate this research toward an extension in MAS to enhance the consensus problem.

*Corresponding author email: francorobles@xlim.fr

[1]J. Franco-Robles are with 3iL Ingénieurs, Limoges, France

[2]J. Franco-Robles, J. Escareno, and O. Labanni-Igbida are with the Robotics and Mechatronics Department, XLIM Laboratory UMR CNRS 7252, University of Limoges, France

B. Main contributions

The main contribution of this manuscript is to develop a sensorimotor leader-follower consensus control. The architecture called *cortical map* is based on the human sensorimotor model to emulate the behavior of visual perception and limb manipulation. The main objective is to capture the leader trajectory in the sensor space while the MAS control lies within the configuration space. The perception of both spaces enriched the offline learning (babbling phase) provided by the position of the MAS given by sensors and the first-order dynamics control of the leader during an exploration of the workspace. Regarding navigation, it is handled based on online learning and consensus theory, while the leader's perception during the consensus navigation reinforces the babbling learning. Consequently, followers capture the leader's information in a directed graph. The Lyapunov function guarantees the stability of the whole system. In detail, the contribution of the current research is displayed:

- Inspired by the BoS postural representation, we propose a sensorimotor hybrid architecture based on a *cortical map* to capture and learn the leader trajectory on the sensor space, inferring the configuration space to control a MAS with a consensus scheme.
- The model resembles the earliest stages of the human babbling phase regarding exploring the workspace with an offline supervised learning technique that additionally reinforces the learned task utilizing consensus navigation through online learning.
- According to [20], equilibrium point theory describes the motion of the postural model, which is highly associated with potential fields. In the same way, we establish the control of the leader-follower consensus associated with the potential field of consensus theory.

The remainder of the paper is organized as follows. Section II presents the theoretical preliminaries and the problem statement. Section III, it is described the multi-agent sensorimotor control system structure. We develop numerical simulations for the proposed multi-agent configuration in section IV. Section V presents the comparison between the learning phases and the ablation. The perspectives and conclusions of the research are provided in section VI.

II. Theoretical Preliminaries

The actual section furnishes theoretical prerequisites concerning graph theory, to define the model of MAS systems, and sensorimotor model, detailing the basics brain-inspired cross-map between configuration $\mathcal{C}$ and sensor $\mathcal{S}$ that defines the cortical map.

A. Graph Theory

A graph is a collection of vertices (nodes) and edges that that are connected (relationships) regarding the information exchange (position, data, etc.). The nodes correspond to entities and edges are the connections between objects. The graph edges sometimes have weights indicating an attribute, as strength connection between nodes.

Graph theory allows to mathematically represent information flow, i.e. topology, described by

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}\ \mathcal{V} = \{1, ..., n\} \tag{1}$$

where $\mathcal{V}$ stands for the nodes set and $\mathcal{E}$ represents the edges of the graph. G is said to be *directed* graph (*digraph*) and without self-loops if

$$\mathcal{E} \subseteq \{(i, j) : i, j \in \mathcal{V}, j \neq i\} \tag{2}$$

On the other hand, $\mathcal{G}$ is said *undirected* if the node i can obtain information about the node j and viceversa, i.e.

$$(i, j) \in \mathcal{E} \Leftrightarrow (j, i) \in \mathcal{E} \tag{3}$$

Adjacency between two nodes, or agents, i and j, exists if there is an edge (i, j) that connects both nodes. In this sense, such nodes are said to be *adjacent* and the aforementioned relation is formally represented as

$$\mathcal{E} = (i, j) \in \mathcal{V} \times \mathcal{V} : i, j \tag{4}$$

The matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ is called the *adjacency matrix* and its elements a_{ij} describe the adjacency between nodes i and j such that

$$a_{ij} = \begin{cases} 0 & \text{if } i = j \\ 1 & \text{if } (i, j) \in \mathcal{E} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The degree matrix is given by, $\mathcal{D} = \text{diag}\left(\sum_{j=1}^{n} a_{ij}\right)$

The *Laplacian matrix* $\mathcal{L}$ of $\mathcal{G}$ is defined as $\mathcal{L} = \mathcal{D} - \mathcal{A}$. For undirected graphs, $\mathcal{L}$ is symmetric and positive semi-definite, i.e., $\mathcal{L} = \mathcal{L}^T \geq 0$. Moreover, the row sums of $\mathcal{L}$ are zero. For connected graphs, $\mathcal{L}$ has exactly one zero eigenvalue, and the eigenvalues can be listed in increasing order $0 = \lambda_1(\mathcal{G}) < \lambda_2(\mathcal{G}) \leq ... \leq \lambda_N(\mathcal{G})$. The second eigenvalue $\lambda_2(\mathcal{G})$ is called the *algebraic connectivity*.

In the actual case, let a directed topology defines the interaction among the leader and the MAS. Regarding the distributed case, $\mathcal{B} = \text{diag}\{b_1, b_2, \ldots, b_n\}$ is utilised to define the access to the virtual leader situation, i.e. $b_i = 1$ if the i^{th}-agent has access to leaders information, and $b_i = 0$ otherwise. Thus, let the digraph $\bar{\mathcal{G}}$ stand for the communication topology encompassing followers and leader.

Lemma II.1. *Consider a digraph $\mathcal{G}$ and its corresponding Laplacian $\mathcal{L}$. Within the set of eigenvalues of $\mathcal{L}$ at least one is zero, while others feature positive real part.*

Definition II.1. *A directed graph, or digraph, $\mathcal{G}$ is balanced if*

$$\sum_{j \neq i} a_{ij} = \sum_{j \neq i} a_{ji} \tag{6}$$

Likewise, balanced digraphs verify

$$\mathcal{L}\mathbf{1}_n = \mathbf{1}_n\mathcal{L} = \mathbf{0}_n \tag{7}$$

with $\mathbf{1}_n = [1, 1, \ldots, 1]^T \in \mathbb{R}^n$ and $\mathbf{0}_n = [0, 0, \ldots, 0]^T \in \mathbb{R}^n$.

Lemma II.2. *The matrix $\mathcal{L}+\mathcal{B}$ has full rank whether the augmented digraph $\bar{\mathcal{G}}$ has a spanning tree with leader as the root, which implies that $\mathcal{L}+\mathcal{B}$ is non-singular.*

B. Sensorimotor Learning

The human sensorimotor model relates the periphery of the external environment with the limbs' posture. The visual sensory inputs are captured by the central nervous system in two pathways involved in the goal-directed reaching: a) the perception and object recognition and b) the related localization of the object in the extra personal space and the action system for the object manipulation. The perception pathway starts from the visual cortex to the temporal cortex, while the localization and action go from the visual cortex to the parietal lobe. The posterior parietal cortex receives the perception data and creates a plan to act based on the target. The cerebellum sends each movement output to the motor cortex and brain stem. Consequently, the brain stem communicates the impulses to activate the spinal cord connections to stimulate the muscles through motor neurons (see fig.1).

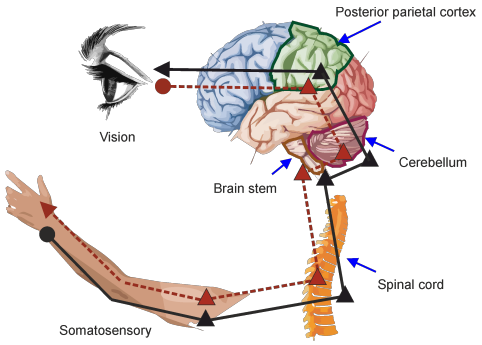

Fig. 1: The neuroanatomical structure of the human sensorimotor control is described in a hierarchical order of the brain, spinal order, and peripheral system [21].

Inspired by the latter, the sensorimotor model is structured as a cortical map f with $\mathcal{P}$ neurons, or processing elements, that map perception to motor commands. The cortical map input vector $\boldsymbol{\beta} \in \mathbb{R}^s$ stands for the perception:

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_s)^T \in \mathcal{S}, \tag{8}$$

where $\boldsymbol{\beta}$ is an s-dimensional vector in the sensor space $\mathcal{S}$. Analogous to the activation and processing of the posterior parietal cortex neurons, a subset $\mathcal{Q} \subset \mathcal{P}$ measures the activation probability of a region of f regarding $\boldsymbol{\beta}$ by means of a k-th softmax activation function:

$$\Phi_k(\boldsymbol{\beta}) = \frac{\exp(-\left(\frac{||\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}_k||}{\sigma}\right)^2)}{\sum_{p=1}^{\mathcal{P}} \exp(-\left(\frac{||\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}_p||}{\sigma}\right)^2)} \tag{9}$$

related to a k-th neuron. The latter is associated with a s-dimensional *prototype* vector $\tilde{\boldsymbol{\beta}}$ with a limited receptive field concerning the mean value (center) and a degree of reception σ^2 (variance). Therefore, in case of a correspondence between the input vector $\boldsymbol{\beta}$ and the *prototype* vector $\tilde{\boldsymbol{\beta}}$, the softmax function is activated at a specific value as a function of the neighboring *prototype* vectors stimulated radially by σ^2. It is possible to recover $\boldsymbol{\beta}$ when considering the activated processing elements with a *reconstruction equation*:

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\beta}) = \sum_{k=1}^{\mathcal{P}} \tilde{\boldsymbol{\beta}}_k \Phi_k(\boldsymbol{\beta}) \tag{10}$$

Consequently, the activation pattern provided (eq. 9) stimulates a m-dimensional *prototype* vector $\tilde{\boldsymbol{\mu}}(\boldsymbol{\beta}) \in \mathbb{R}^m$ that constitute the degrees of freedom (DoF) of the human postural configuration space $\mathcal{C}$:

$$\hat{\boldsymbol{\mu}}(\boldsymbol{\beta}) = \sum_{k=1}^{\mathcal{P}} \tilde{\boldsymbol{\mu}}_k \Phi_k(\boldsymbol{\beta}) \tag{11}$$

where

$$\hat{\boldsymbol{\mu}}(\boldsymbol{\beta}) = (\mu_1, \mu_2, \ldots, \mu_m)^T \in \mathcal{C} \tag{12}$$

Likewise, the k-th $\tilde{\boldsymbol{\mu}}$ prototype vector is also related to the k-th neuron. Particularly, the sensor data and their corresponding domain in the motor space are represented in an ordered pair $(\tilde{\boldsymbol{\mu}}_k, \tilde{\boldsymbol{\beta}}_k)$ called *body-icon*. Each *prototype vector* is adjusted employing supervised learning with *gradient descent with momentum* as a function of a training set $\in \mathcal{C}, \mathcal{S}$.

From the neurobiological perspective, the interdependence between $\mathcal{S}$ and $\mathcal{C}$ yields to *circular reaction*, where the coordination between the motor commands and the proprioceptive and exteroceptive sensors are explained from an equilibrium point hypothesis. In other words, the initial and final joint positions are states of equilibrium represented on the cortical map as potential fields.

The body's behavior explained as motor commands, describes the *predictive/forward control* as potential fields to reach smooth and dexterous manipulation through two essential tasks in human sensorimotor learning: the babbling and skills development phase. The former is conceptualized in [22], where infants sample and learn their body schema/image by exploring the sensory world. The latter addresses a complex limb association in a highly tuned coordinate system during a repetition task development. The present manuscript proposes such learning in a multi-agent control system. In other words, the cornerstone is to describe the perception of agents related to the human visual system and the control system inspired by limb manipulation.

III. Main Result: MAS Sensorimotor Control Scheme (MAS-ScS)

The herein proposed sensorimotor model extends the cortical map architecture proposed in [20] to the multi-agent case (MAS-ScS) via consensus theory. The MAS-ScS is designed to acquire information corresponding to the leader kinematics to train the model. Unlike [20], we propose: a) a supervised offline learning phase for babbling and b) an online learning phase during consensus (see fig. 2).

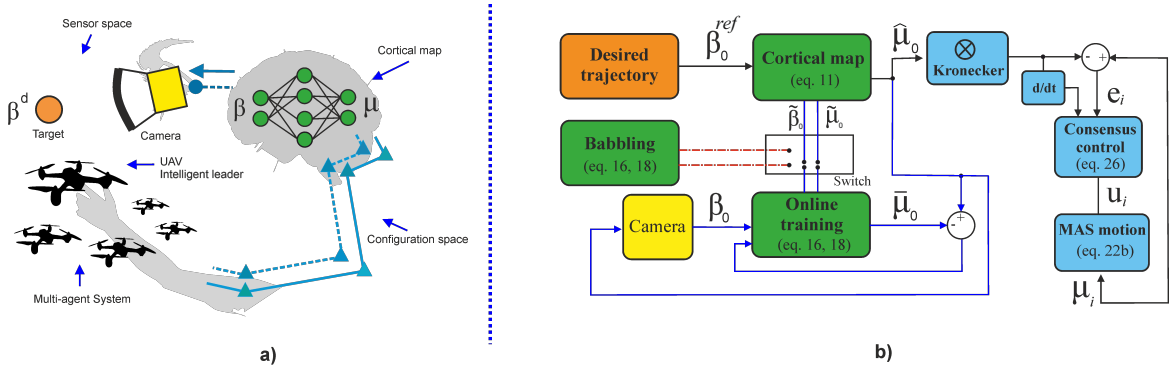

Fig. 2: MAS Sensorimotor Control Scheme: a) The associate model relates the sensorimotor structure in a cortical map to capture the leader's trajectory through visual perception and control the desired configuration of followers during navigation. b) The initialization resembles the babbling phase through exploring the workspace. Afterward, the learned body-icons are tuned with online learning during the consensus concerning the desired trajectory.

A. *Offline Babbling Stage*

The k-th *body-icon* builds a correlation in the cortical map between the leader configuration space $\mathcal{C}$ and the sensor space $\mathcal{S}$. Initially, a set of $\mathcal{M} = \{(\tilde{\boldsymbol{\mu}}_{0_1}, \tilde{\boldsymbol{\beta}}_{0_1}), ..., (\tilde{\boldsymbol{\mu}}_{0_k}, \tilde{\boldsymbol{\beta}}_{0_k}), ..., (\tilde{\boldsymbol{\mu}}_{0_\mathcal{P}}, \tilde{\boldsymbol{\beta}}_{0_\mathcal{P}})\}$ *body-icons* with $\mathcal{P}$ neurons is defined. Alternatively to [20], we establish that the learning comes from perception to the configuration space to control the leader. Therefore, the k-th element is adjusted through supervised learning based on an input vector $\boldsymbol{\beta}_0$ associated with the corresponding configuration $\boldsymbol{\mu}_0$ generated during the workspace exploration.

The mainstay of babbling is the measurement activation degree of the adjacent and the k-th *body-icon* as a function of the association degree of the training input vector $\boldsymbol{\beta}_0$ concerning the k-th prototype vector $\tilde{\boldsymbol{\beta}}_{0_k}$. Accordingly, the Euclidean distance measures the association degree. In contrast, activation degree is expressed employing the Softmax function (eq. 9) that emulates *population coding* [23] of the cerebral cortex.

Significant activation is present in $\tilde{\boldsymbol{\beta}}_{0_k}$ regarding the minimum Euclidean distance and decreasing activation of the adjacent prototype vectors regarding increasing distances. The variance σ^2 controls the degree of activation, which is gradually diminished according to the current iterations it. Therefore, the influence of the input $\boldsymbol{\beta}_0$ on the adjacent prototype vectors decreases until the maximum number of iterations it_{max} is attained with an σ_{int} initial and σ_{fin} final degree:

$$\sigma = \sigma_{int}(\sigma_{fin}/\sigma_{int})^{(\text{it}/\text{it}_{max})} \tag{13}$$

Learning is supervised since the training vector is acquired from the leader kinematics $\boldsymbol{\mu}_0 \in \mathcal{C}$ regarding $\boldsymbol{\beta}_0 \in \mathcal{S}$ to adjust the body-icons. Therefore, the loss function:

$$E = \frac{1}{2}\|\hat{\boldsymbol{\mu}}_0(\boldsymbol{\beta}_0) - \boldsymbol{\mu}_0^{ref}\|^2 \tag{14}$$

where $\hat{\boldsymbol{\mu}}_0(\beta_0)$ stands for function relating $\mathcal{S}$ to $\mathcal{C}$, whereas $\boldsymbol{\beta_0}$ represents the position in $\mathcal{S}$ and $\boldsymbol{\mu}_0^{ref}$ is a reference sensor (ground truth) used for supervised training/learning purposes. The learning rule updates the k-th prototype vector $\tilde{\boldsymbol{\mu}}_{0_k}$ with the gradient descent with momentum optimization at each iteration:

$$\Delta\tilde{\boldsymbol{\mu}}_{0_k} = -\eta\frac{\partial E}{\partial\tilde{\mu}_0} = -\eta(\hat{\boldsymbol{\mu}}_0(\boldsymbol{\beta}_0) - \boldsymbol{\mu}_0^{ref})\Phi_k(\boldsymbol{\beta}_0) \tag{15}$$

$$\tilde{\boldsymbol{\mu}}_{0_k} = \tilde{\boldsymbol{\mu}}_{0_k}(\text{it}-1) + \Delta\tilde{\boldsymbol{\mu}}_{0_k}(\text{it}) + \alpha(\tilde{\boldsymbol{\mu}}_{0_k}(\text{it}-1) - \tilde{\boldsymbol{\mu}}_{0_k}(\text{it}-2)) \tag{16}$$

Likewise, $\tilde{\boldsymbol{\beta}}_{0_k}$ is trained with the same approach:

$$\Delta\tilde{\boldsymbol{\beta}}_{0_k} = \eta\frac{\partial E}{\partial\tilde{\mu}_0}\frac{\partial\tilde{\mu}_0}{\partial\tilde{\beta}_0}\boldsymbol{\mu}_0\Phi_k(\boldsymbol{\beta}_0)(\delta_{kl} - \Phi_l(\boldsymbol{\beta}_0))\frac{\boldsymbol{\beta}_0 - \tilde{\boldsymbol{\beta}}_{0_k}}{\sigma^2}$$
$$\delta_{kl} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \tag{17}$$

$$\tilde{\boldsymbol{\beta}}_{0_k} = \tilde{\boldsymbol{\beta}}_{0_k}(\text{it}-1) + \Delta\tilde{\boldsymbol{\beta}}_{0_k}(\text{it}) + \alpha(\tilde{\boldsymbol{\beta}}_{0_k}(\text{it}-1) - \tilde{\boldsymbol{\beta}}_{0_k}(\text{it}-2)) \tag{18}$$

with the condensed formulation (Kronecker delta function δ) of the softmax derivative regarding the l-th input of the k-th softmax. The learning rate α is constant and η decreases regarding an η_{int} initial rate and a η_{fin} final rate:

$$\eta = \eta_{int}(\eta_{fin}/\eta_{int})^{(\text{it}/\text{it}_{max})} \tag{19}$$

B. *MAS-ScS Navigation Strategy*

The motricity skills development phase relies upon effective and efficient collection and processing of the sensorimotor information associated to each action [24]. From a postural point of view, the muscles represent an elastic force field determining specific postural position while rejecting/compensating disturbances. Specific muscles configuration produce trajectories from an initial state towards the

target (equilibrium points). In this regard, potential fields correspond to set equilibrium points; the latter suggests having target/obstacle attraction/repulsion fields. Each posture carries an intention that implicitly creates an internal representation of desired states. Repetitive target-based movements reinforce neural connections of the corresponding body-icons. Hence, our model links *ideomotor* theory since online learning associates the effects of the generated movements with sensory information and the motor patterns when the intention to achieve the goal is developed. Each movement enriches the information of the k-th body-icon resembling the skill development task.

As a result, after the babbling stage, the leader's position $\boldsymbol{\beta}_0(\boldsymbol{\mu}_0) \in \mathcal{S}$ within sensors space is mapped to the configuration space via the trained cortical map (eq. 21) to construct the potential field ε between the leader $\boldsymbol{\mu}_0 = f(\boldsymbol{\beta}_0) \in \mathcal{C}$ and the followers $\boldsymbol{\mu}_i$

$$\boldsymbol{\beta}_0(\boldsymbol{\mu}_0) \in \mathcal{S} \quad \mapsto \boldsymbol{\mu}_0 = f(\boldsymbol{\beta}_0) \in \mathcal{C} \tag{20}$$

Depending on the number of neurons and Babbling density, the inference resulting from the offline stage learning reads

$$\hat{\boldsymbol{\mu}}_0(\boldsymbol{\beta}_0) = \sum_{k=1}^{\mathcal{P}} \tilde{\boldsymbol{\mu}}_{0_k} \Phi_k(\boldsymbol{\beta}_0) \tag{21}$$

Since this estimation encompasses a residual error, a complementary online learning is necessary to refine the neurons weights resulting from the offline Babbling stage. In this regard, the learning rules (16) and (18) update the body-icons in a shallow architecture (see fig. 3) during the consensus stabilization process.

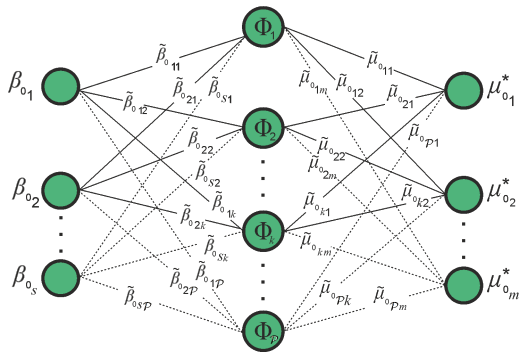

Fig. 3: Cortical map architecture

1) Consensus Control Scheme: During the consensus stabilization process, the sensory data coming from the leader's navigation reinforce the cortical map learning (see fig. 2). To this end, a controller scheme is synthesized via the Lyapunov theory to drive the MAS trajectories to achieve the leader-follower obejctive.

Let us consider a first-order MAS dynamics featuring a leader-follower topology and evolving within within a m-dimensional space.

$$\begin{aligned} \dot{\boldsymbol{\mu}}_0^{\star} &= f(\boldsymbol{\mu}_0^{\star}) \\ \dot{\boldsymbol{\mu}}_i &= \boldsymbol{u}_i \end{aligned} \tag{22}$$

with $i = 1, \ldots, n$. In terms of states , $\boldsymbol{\mu}_0^{\star} \in \mathbb{R}^m$ stands for two dimensional leader's position and $\boldsymbol{\mu}_i \in \mathbb{R}^{nm}$ represents the i^{th}-agent positions and both in the configuration space $\mathcal{C}$. The controller of leader is denoted by $\boldsymbol{u}_0 \in \mathbb{R}^m$ and $\boldsymbol{u}_i \in \mathbb{R}^{nm}$, respectively. Let us the define the disagreement variable as

$$\boldsymbol{e}_i = \boldsymbol{\mu}_i - [\mathbf{1}_n \otimes \hat{\boldsymbol{\mu}}_0(\boldsymbol{\beta}_0)] \tag{23}$$

Likewise, let us define the distributed lumped error as

$$\bar{\boldsymbol{e}}_i = [\mathcal{H} \otimes I_m][\boldsymbol{\mu}_i - (\mathbf{1}_n \otimes \hat{\boldsymbol{\mu}}_0(\boldsymbol{\beta}_0))] \tag{24}$$

with $\dot{\boldsymbol{\mu}}_i = \boldsymbol{u}_i$ being the control input, and $\mathcal{H} = \mathcal{L} + \mathcal{B}$. The time derivative of (24) defining the error dynamics reads

$$\dot{\bar{\boldsymbol{e}}}_i = [\mathcal{H} \otimes I_m][(\boldsymbol{u}_i - (\mathbf{1}_n \otimes \dot{\hat{\boldsymbol{\mu}}}_0(\boldsymbol{\beta}_0)))] \tag{25}$$

Since $\mathcal{H}$ is full-rank, thus invertible, fulfilling the tracking problem of (22) is equivalent to stabilize the lumped error dynamics expressed by (25).

Theorem III.1. *Given the system (22) and the lumped error dynamics (25). The controller fulfilling leader-follower tracking problem is given as*

$$\boldsymbol{u}_i = (\mathcal{H} \otimes I_m)^{-1}(-\Lambda \bar{\boldsymbol{e}} + (\mathcal{B} \otimes I_m)(\mathbf{1}_n \otimes \dot{\hat{\boldsymbol{\mu}}}_0)) \tag{26}$$

where $\Lambda > 0 \in \mathbb{R}^{n \times m}$

Proof: Let us consider the extended Lyapunov disagreement function,

$$\varepsilon = \frac{1}{2}\|\bar{\boldsymbol{e}}_i\|^2 \tag{27}$$

The corresponding time-derivative along the trajectories of the configuration space states yields

$$\dot{\varepsilon} = \bar{\boldsymbol{e}}_i^T \dot{\bar{\boldsymbol{e}}}_i = \bar{\boldsymbol{e}}_i^T [\mathcal{H} \otimes I_m][(\boldsymbol{u}_i - (\mathbf{1}_n \otimes \dot{\hat{\boldsymbol{\mu}}}_0)] \tag{28}$$

Then, using (26) leads

$$\dot{\varepsilon} = -\bar{\boldsymbol{e}}_i^T \Lambda \bar{\boldsymbol{e}}_i + [-(\mathcal{L}+\mathcal{B}) \otimes I_m + (\mathcal{B}) \otimes I_m](\mathbf{1}_n \otimes \dot{\hat{\boldsymbol{\mu}}}_0)) \tag{29}$$

recalling that $\mathcal{L}\mathbf{1}_n = \mathbf{1}_n\mathcal{L} = \mathbf{0}_n$, the prior equation reduces to

$$\dot{\varepsilon} = -\bar{\boldsymbol{e}}_i^T \Lambda \bar{\boldsymbol{e}}_i \leq 0 \tag{30}$$

since $\Lambda > 0$, such result guarantees that the equilibrium is asymptotically stable, that is to say, the trajectories of $\bar{\boldsymbol{e}}_i \to 0$ as $t \to \infty$.

IV. Numerical Simulation

A a set of numerical experiments are conducted to evaluate the effectiveness of the proposed MAS-ScS method. The position of the i-th agent is described in the $\mathcal{C}$-space, i.e $\boldsymbol{\mu}_i$ for the i-th agent, i.e.

$$\boldsymbol{\mu}_i = (\mu_{1_i}, \mu_{2_i})^T = (x_i, y_i)^T \in \mathcal{C} \tag{31}$$

The perception (sensor space $\mathcal{S}$) of the real-world space is assumed that a visual sensor is covering a $20 \times 20[m^2]$ workspace and featuring a (1920x1080 pixels) resolution. Hence, the image plane corresponds to $\mathcal{S}$-plane.

$$\boldsymbol{\beta} = (\beta_1, \beta_2)^T = (x^{cam}, y^{cam})^T \in \mathcal{S} \tag{32}$$

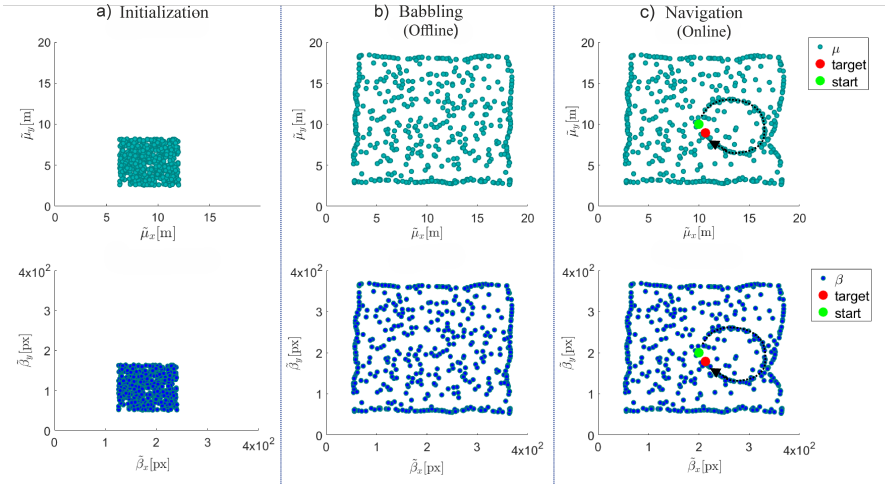

Fig. 4: The Multi-Agent Sensorimotor Control system is summarized in an a) initialization phase to establish a random association between body-icons to train during b) babbling (offline learning) and the c) navigation (online learning). The babbling phase resembles the exploration of the peripersonal space of an infant. At the same time, the navigation relates to the repetition of a human goal-directed task, showing that the involving body-icons strongly correlate with the navigation path, resulting in an improved generalization.

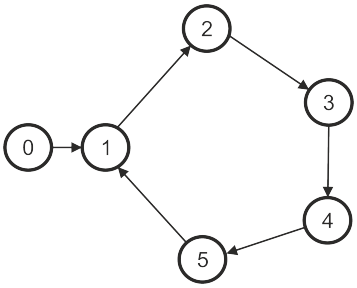

Fig. 5: MAS topology

The actual scenario herein addressed considers a collection of five omnidirectional agents ($n = 5$) evolving within a bi-dimensional space ($m = 2$).

Remark IV.1. *It is noteworthy that the proposed sensory-motor strategy is applicable to holonomic and nonholonomic mobile systems:*

- *For the aerial and nonholonomic terrestrial systems the actual approach correspond to high-level commands to drive horizontal translational ($\boldsymbol{u}_x$ and $\boldsymbol{u}_y$ as references) for roll/pitch and yaw, respectively.*
- *In the case of the holonomic aerial/terrestrial systems, the command is applied directly.*

The MAS corresponding information policy is defined by a directed topology (see Fig. 5).

Whose adjacency and degree matrices are written as

$$\mathcal{A} = \begin{pmatrix} 0 & a_{12} & 0 & 0 & 0 \\ 0 & 0 & a_{23} & 0 & 0 \\ 0 & 0 & 0 & a_{34} & 0 \\ 0 & 0 & 0 & 0 & a_{45} \\ a_{51} & 0 & 0 & 0 & 0 \end{pmatrix}, \mathcal{D} = \begin{pmatrix} a_{12} \\ a_{23} \\ a_{34} \\ a_{45} \\ a_{51} \end{pmatrix} \tag{33}$$

Likewise, the matrix defining the agents acquiring the virtual leader, A_0, which is A_1 in the actual case, is defined as

$$\mathcal{B} = \text{diag}(1, 0, 0, 0, 0) \tag{34}$$

And assuming a effective communication, i.e. $a_{ij} = 1$, the Laplacian is written as

$$\mathcal{L} = \mathcal{D} - \mathcal{A} \tag{35}$$

$$\mathcal{L} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{36}$$

The matrix $\mathcal{H}$ driving the collective behavior of leader-follower MAS system is thus written as

$$\mathcal{H} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{37}$$

A. Babbling

Throughout babbling (see fig. 4b), random exploration paths based on waypoints of the leader $\boldsymbol{\mu}_0$ within the image plane are applied. Remarkably, the body-icons form a multi-dimensional *grid* where each has a coordinate determined by learning in each space, allowing a geometric representation. Then, it is desirable to reproduce a uniform exploration paths within and over the *grid* boundaries to improve the generalization error. Hence, 100 random paths are generated for $\boldsymbol{\mu}_0$.

1) Ablation: Regarding the architecture of the cortical map, a range of 50 to 1000 neurons is established to infer the generalization/interpolation capacity, considering the training data size, the neurons, and the variance σ as significant factors. The activation range of the adjacent body-icons is described by $\sigma_{int} = 0.8$ to $\sigma_{fin} = 0.2$ during 360 epochs (100 training vectors/epoch) while the learning rate maintains $\eta = 0.6$. The mean squared error (MSE):

$$\tilde{\boldsymbol{\beta}}_{\text{MSE}} = \frac{1}{100}\sum_{\mathsf{b}=1}^{100}\left[\boldsymbol{\beta}_{0_{\mathsf{b}}}^{ref} - (\sum_{k=1}^{\mathcal{P}}\tilde{\boldsymbol{\beta}}_{0_k}\Phi_k(\boldsymbol{\beta}_0))_{\mathsf{b}}\right]^2 \tag{38}$$

$$\tilde{\boldsymbol{\mu}}_{\text{MSE}} = \frac{1}{100}\sum_{\mathsf{b}=1}^{100}\left[\boldsymbol{\mu}_{0_{\mathsf{b}}}^{ref} - (\sum_{k=1}^{\mathcal{P}}\tilde{\boldsymbol{\mu}}_{0_k}\Phi_k(\boldsymbol{\beta}_0))_{\mathsf{b}}\right]^2 \tag{39}$$

reveals a quasi-linear correlation regarding the number of neurons, where 800 neurons perform better, resulting in 0.1698 for $\hat{\mu}$ and 0.6791 for $\hat{\beta}$ (see fig. 6).

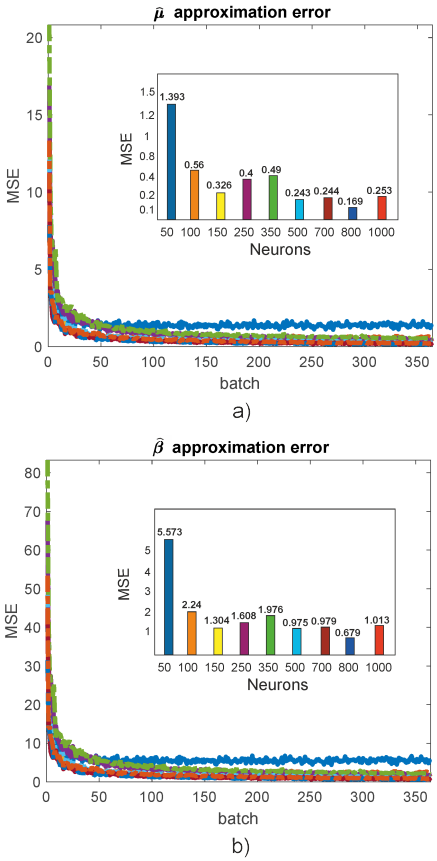

Fig. 6: Approximation error of a) $\hat{\mu}$ and b) $\hat{\beta}$ measured by the mean squared error using 360 epochs during the babbling phase setting 50-1000 neurons.

B. Consensus

The dynamics of the leader $\boldsymbol{\beta}_0$ on sensor space is described on function of time t:

$$\boldsymbol{\beta}_0 = \begin{bmatrix} \sin(0.6\pi)t + 0.01\sin(\mu_{0_x}) \\ \cos(0.6\pi)t + 0.01\sin(\mu_{0_y}) \end{bmatrix} \tag{40}$$

Based on the ablation performance, two prominent simulation cases are established: consensus with babbling learning and consensus with online learning. The former reaches a desirable MSE performance with 800 neurons. However, the inference from the cortical map at each iteration during consensus exhibits a predictable precision error inherent to interpolation. In other words, the absence of a body-icon corresponding to an inference value creates the interpolation error (see fig. 4b). However, fig. 7b shows that the followers reach a consensus in a finite time. After that, the leaders start to make a circular displacement with perturbations due to the interpolation error. Consequently, the followers capture the leader's information and replicate the same error (see fig. 7a).

Intuitively, the babbling stage improves the interpolation inference if the number of body-icons increases. As an analogy, an infant's complete exploration task enriches their workspace's knowledge. Nevertheless, the training time increases proportionally to the size of the exploration task. It is perceptible in fig. 7c that online training reduces the interpolation error and reinforces the connections of the body-icons involved on the path (see fig. 4c). The consensus is reached in a finite time fig. 7d. Unlike babbling, the dynamics of leader $\boldsymbol{\beta}_0$ is well-predicted since the trajectory presents negligible interpolation errors. Consequently, the followers successfully reach a consensus during the leader dynamics.

V. Conclusions

In this study we have presented an alternative two-level learning architecture to address the leader-follower consensus problem. Such scheme is inspired on the human sensorimotor system and developed sequentially within the babbling and collective navigation (consensus) levels. The initial stage relates to offline learning concerning exploring the workspace with the leader, resembling the earliest stage of the human sensorimotor perception and motion in unknown environments. The ablation study shows that 800 neurons perform a MSE of 0.1698 and 0.6791 for inferring the configuration space and the sensor space, respectively. Over the succeeding stage, the sensorimotor model develops a well-adjusted generalization in function of the consensus that reinforces the neurons connections via online learning. The numerical simulation shows that the consensus reaches a circular trajectory without significant disturbances. Forthcoming perspectives consist in extending the approach to the robust case and heterogeneous MAS.

References

[1] A. Birk, B. Wiggerich, H. Buelow, M. Pfingsthorn, and S. Schwertfeger, "Safety, security, and rescue missions with an unmanned aerial vehicle (uav)," *Journal of Intelligent Robotic Systems*, vol. 64, pp. 57–76, 10 2011.

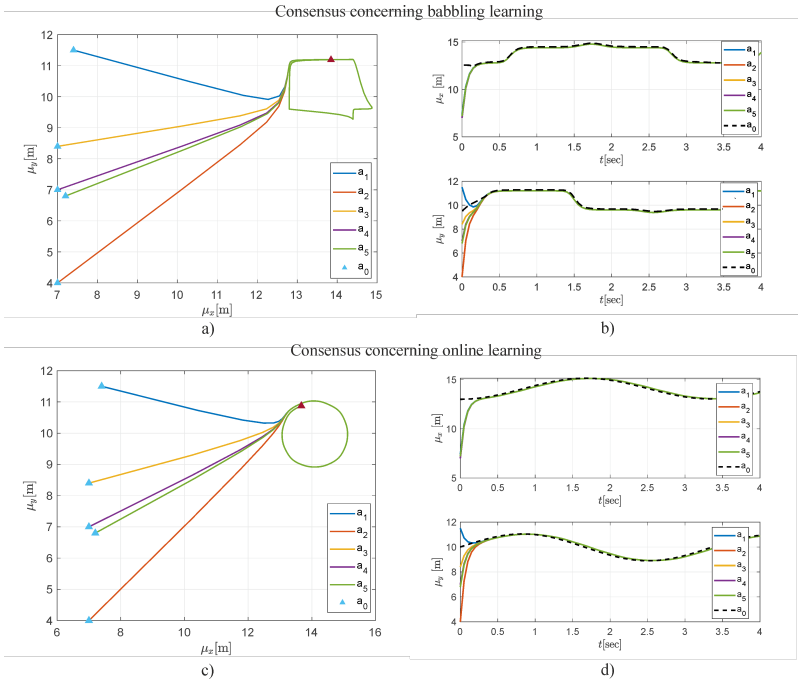

Fig. 7: Simulations have been performed for two cases: a-b) The consensus utilizing the body-icons learned from babbling and c-d) the consensus considering online training of the body-icons. The former shows an inadequate inference of the leader's circular trajectory to the configuration space. While the latter performs the desired trajectory with negligible perturbation.

[2] C. Fu, A. Carrio, M. A. Olivares-Mendez, R. Suarez-Fernandez, and P. Campoy, "Robust real-time vision-based aircraft tracking from unmanned aerial vehicles," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5441–5446, 2014.

[3] M. A. Olivares-Mendez, C. Fu, P. Ludivig, T. F. Bissyandé, S. Kannan, M. Zurad, A. Annaiyan, H. Voos, and P. Campoy, "Towards an autonomous vision-based unmanned aerial system against wildlife poachers," *Sensors*, vol. 15, no. 12, pp. 31362–31391, 2015.

[4] C. Wang, J. Wang, X. Zhang, and X. Zhang, "Autonomous navigation of uav in large-scale unknown complex environment with deep reinforcement learning," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 858–862, 2017.

[5] X. Liu, S. S. Ge, and C.-H. Goh, "Formation potential field for trajectory tracking control of multi-agents in constrained space," *International Journal of Control*, vol. 90, no. 10, pp. 2137–2151, 2017.

[6] K.-K. Oh, M.-C. Park, and H.-S. Ahn, "A survey of multi-agent formation control," *Automatica*, vol. 53, pp. 424–440, 2015.

[7] W. Ren, "Consensus strategies for cooperative control of vehicle formations," *IET Control Theory Applications*, vol. 1, pp. 505–512(7), March 2007.

[8] H. Zhang, F. L. Lewis, and A. Das, "Optimal design for synchronization of cooperative systems: State feedback, observer and output feedback," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1948–1952, 2011.

[9] Z. Li, Z. Duan, and L. Huang, "Leader-follower consensus of multi-agent systems," in *2009 American Control Conference*, pp. 3256–3261, 2009.

[10] F. Chen and D. V. Dimarogonas, "Consensus control for leader-follower multi-agent systems under prescribed performance guarantees," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 4785–4790, 2019.

[11] Y. Shang and Y. Ye, "Fixed-time group tracking control with unknown inherent nonlinear dynamics," *IEEE Access*, vol. 5, pp. 12833–, 07 2017.

[12] Y. Huang, J. Tang, and S. Lao, "Uav group formation collision avoidance method based on second-order consensus algorithm and improved artificial potential field," *Symmetry*, vol. 11, no. 9, 2019.

[13] P. Shi and B. Yan, "A survey on intelligent control for multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 161–175, 2021.

[14] C.-W. Kuo, C.-C. Tsai, and C.-T. Lee, "Intelligent leader-following consensus formation control using recurrent neural networks for small-size unmanned helicopters," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 2, pp. 1288–1301, 2021.

[15] J. Ni and P. Shi, "Adaptive neural network fixed-time leader–follower consensus for multiagent systems with constraints and disturbances," *IEEE Transactions on Cybernetics*, vol. 51, no. 4, pp. 1835–1848, 2021.

[16] H. Head and Holmes in *Studies in Neurology*, vol. 2, 1846-1922.

[17] J. Sturm, C. Plagemann, and W. Burgard, "Body schema learning for robotic manipulators from visual self-perception," *Journal of Physiology-Paris*, vol. 103, no. 3, pp. 220–231, 2009. Neurorobotics.

[18] E. T. Enikov and J. Escareno, "Application of sensory body schemas to path planning for micro air vehicles (mavs)," in *2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, vol. 01, pp. 25–31, 2015.

[19] J. Franco-Robles, J. Escareno, D. Soto-Guerrero, and O. Labbani-Igbida, "Feedforward formation control based on self-organized body-schema," in *2021 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1299–1307, 2021.

[20] P. Morasso and V. Sanguineti, "Self-organizing body schema for motor planning," *Journal of Motor Behavior - J MOTOR BEHAV*, vol. 27, pp. 52–66, 03 1995.

[21] A. Shumway-Cook and M. Woollacott, *Motor Control: Translating Research Into Clinical Practice*. Lippincott Williams & Wilkins, 2007.

[22] R. Saegusa, G. Metta, G. Sandini, and S. Sakka, "Active motor babbling for sensorimotor learning," in *2008 IEEE International Conference on Robotics and Biomimetics*, pp. 794–799, 2009.

[23] B. B. Averbeck, P. E. Latham, and A. Pouget, "Neural correlations, population coding and computation," *Nature Reviews Neuroscience*, vol. 7, pp. 358–366, May 2006.

[24] D. Wolpert, J. Diedrichsen, and J. Flanagan, "Principles of sensorimotor learning," *Nature reviews. Neuroscience*, vol. 12, pp. 739–51, 12 2011.