Proceedings of the ASME 2024
International Design Engineering Technical Conferences and
Computers and Information in Engineering Conference
IDETC/CIE2024
August 25–28, 2024, Washington, DC

IDETC/CIE2024: 143377

DISTRIBUTED MULTI-AGENT BAYESIAN OPTIMIZATION FOR UNKNOWN DESIGN SPACE EXPLORATION

Siyu Chen[1], Alparslan Emrah Bayrak [2], Zhenghui Sha[1,*],

[1]Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX
[2]Department of Mechanical Engineering and Mechanics, Lehigh University, Bethlehem, PA

ABSTRACT

In multi-agent Bayesian optimization for Design Space Exploration (DSE), identifying a communication network among agents to share useful design information for enhanced cooperation and performance, considering the trade-off between connectivity and cost, poses significant challenges. To address this challenge, we develop a distributed multi-agent Bayesian optimization (DMABO) framework and study how communication network structures/connectivity and the resulting cost would impact the performance of a team of agents when finding the global optimum. Specifically, we utilize Lloyd's algorithm to partition the design space to assign distinct regions to individual agents for exploration in the distributed multi-agent system (MAS). Based on this partitioning, we generate communication networks among agents using two models: 1) a range-limited model of communication constrained by neighborhood information; and 2) a range-free model without neighborhood constraints. We introduce network density as a metric to quantify communication costs. Then, we generate communication networks by gradually increasing the network density to assess the impact of communication costs on the performance of MAS in DSE. The experimental results show that the communication network based on the range-limited model can significantly improve performance without incurring high communication costs. This indicates that increasing the density of a communication network does not necessarily improve MAS performance in DSE. Furthermore, the results indicate that communication is only beneficial for team performance if it occurs between specific agents whose search regions are critically relevant to the location of the global optimum. The proposed DMABO framework and the insights obtained can help identify the best trade-off between communication structure and cost for MAS in unknown design space exploration.

Keywords: Distributed Multi-agent System, Bayesian Optimization, Design Space Exploration

*Corresponding author: zsha@austin.utexas.edu

1. INTRODUCTION

Unknown Design Space Exploration (DSE) [1] is a decision-making process involving systematically evaluating different designs to identify the optimal solution. Often modeled with a black-box objective function, the goal of unknown DSE is to locate the optimum of this objective function. Bayesian Optimization (BO) [2] is a widely used method for DSE due to its sequential approach to a rational decision-making [3]. By continuously updating the surrogate model with newly collected data, BO uses the Acquisition Function (AF) to select the next design point for evaluation. This process empowers designers with the insights needed to strategically choose where to sample next, leading them toward the most effective design solutions.

BO has gathered significant attention over the last century, yet the exploration into team decision-making through Multi-agent BO (MABO) remains relatively sparse. The objective of MABO is to develop an Multi-agent System (MAS) as a design team to find the best design solution for complex DSE. In our previous study [4], we introduced the MABO framework and developed a global evaluator that allows agents to share information. Therefore, our previous study operates on a centralized model, requiring a global evaluator to aggregate data from all agents, learn the GP model, and optimize the AF. However, the centralized nature of traditional optimization systems often leads to significant limitations in scalability and robustness [5]. In this paper, to address these limitations, we extend the centralized MABO to a distributed setting, and develop a distributed multi-agent Bayesian optimization (DMABO), where each agent maintains its own data and independently constructs surrogate models and AFs.

Distributed MASs have shown benefits in many applications, such as source search in robotics, where agents hold different beliefs about the objective function [6, 7]. Inspired by real-world MAS, such as robotics, various strategies have been explored in the literature to facilitate MABO in a distributed setting [8]. Thompson Sampling is highlighted as a decentralized strategy,

notable for its use of Monte Carlo methods to generate multiple sampling points and selection based on the highest posterior probability [9, 10]. Further advances in DMABO can be seen in the study by Yue et al. [5], which introduces the MABO framework with consensus and allows agents to agree on decisions. The study by Garcia et al. [11], introduces stochastic policies for BO to generate multiple sampling points for the next iteration. However, this approach can be considered semi-centralized, since it only updates one surrogate model and generates multiple sampling points for the MAS by stochastic policies in one iteration. In our study, we develop a fully distributed DMABO framework. We aim to further decentralize the process by allowing each agent within the distributed MAS to independently model its own sequence of decision-making and search activities.

Within the DMABO framework, two primary challenges emerge: *developing communication structures* and *managing communication costs*. The first challenge is to develop a communication structure based on segmentation of the design space into distinct regions for targeted sampling by agents in DMABO. This strategy can promote focused exploration within each agent's allocated region. For example, in the study [12, 13], they developed DMABO methods that segment the search space into local areas. However, these frameworks do not address the communication structure/network among agents, a critical element for collective decision-making. Our research addresses this by integrating a communication network tailored to the segmented design space with the aim of enhancing the performance of MAS.

The second challenge within DMABO revolves around managing communication costs, a metric quantifiable by network density [14, 15]. Developing a communication network involves high connectivity, which is crucial for effective information sharing and impacts collective decision-making and overall MAS performance. Therefore, maintaining high connectivity among agents is essential to improve performance. However, achieving high connectivity, or network density, often results in increased communication costs in practical applications, presenting a conflict between connectivity and cost (see Fig. 1). In response to these two challenges, our study introduces a novel strategy for communication structures that improves collective decision-making while efficiently managing communication costs. To the best of our knowledge, this is the first attempt to quantitatively study the trade-off between communication connectivity and cost through communication networks of MAS in order to achieve the best performance in DMABO.

This study is motivated by answering the following research question (RQ): *What are the impacts of communication structures and costs on the performance of a distributed multi-agent system in unknown design space exploration?* To answer this RQ, we contribute to the literature with the following: *First*, we introduce a DMABO for DSE. Specifically, in the distributed MAS, we leverage the Lloyd's algorithm to segment the design space, allocating specific regions to individual agents. *Second*, based on the segmented spaces derived from the Lloyd's algorithm, we establish the communication networks between agents by two models: 1) a range-limited model constrained by neighborhood information, which gives us a communication network that can be obtained from the adjacency matrix of the partitioned regions; 2) a range-free model without any neighborhood constraints, where we randomly connect a pair of nodes to establish a communication link. *Third*, we adopt network density as a metric to quantify communication costs, allowing us to systematically generate diverse networks with varying communication costs. By increasing network density, we explore the implications of both communication network models and the resulting cost on the performance of MAS in DSE.

The structure of this paper is organized as follows: Section 2 introduces the preliminary concepts essential for this study. Section 3 explores the technical aspects of the proposed DMABO framework, detailing the problem formulation, the communication structure, and the DMABO-based solution methodology. Section 4 describes the experimental settings and results, showing a comparative analysis of communication networks with and without neighborhood constraints under varying network densities. Section 5 discusses in detail the research findings and the insights. The paper concludes in Section 6 with a summary of the main results and limitations that suggest directions for future research.

2. PRELIMINARIES

2.1 Bayesian Optimization

Bayesian Optimization (BO) represents a global optimization approach designed to identify the optimal value of a black-box objective function, denoted as $f(\mathbf{x})$, where $\mathbf{x}$ represents a vector of design variables within the domain A, $A \subseteq \mathbb{R}^d$. The effectiveness of BO is attributed to two primary elements: first, the application of a statistical inference method, typically Gaussian Process (GP) regression, to predict the objective function based on sampled data; second, the acquisition function (AF) to determine the subsequent sampling point in the design space for the optimal solution [16]. For brevity, we omit the comprehensive technical overview of the BO used in this paper and refer the readers to Section 2 of our previous study [4].

2.2 Division of design space

In this study, we utilize Voronoi diagrams to partition the design space, integrating the Lloyd's algorithm to ensure an even segmentation and balance the workload of each agent. This section outlines the key principles of Voronoi diagrams and Lloyd's algorithm.

2.2.1 Voronoi Diagrams. Voronoi diagrams, are a way to divide a space into local regions based on a set of seed points, denoted as $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N\}$ [17]. Each local region, or Voronoi cell, $V(\mathbf{p}_i)$, associated with a seed point $\mathbf{p}_i$, consists of all points in the space closer to $\mathbf{p}_i$ than to any other seed point $\mathbf{p}_j$. Voronoi cells can be represented mathematically as:

$$V(\mathbf{p}_i) = \left\{ \mathbf{x} \in \mathbb{R}^n : \forall \mathbf{p}_j \in \mathbf{P}, i \neq j, d(\mathbf{x}, \mathbf{p}_i) < d\left(\mathbf{x}, \mathbf{p}_j\right) \right\} \quad (1)$$

where $d(\mathbf{x}, \mathbf{p}_i)$ is the Euclidean distance between any point $\mathbf{x}$ in the design space and the seed $\mathbf{p}_i$. Essentially, Voronoi diagrams provide a unique way of partitioning space into Voronoi cells, each corresponding to a specific seed point, ensuring that every location within a cell is closest to its corresponding seed compared to any other.

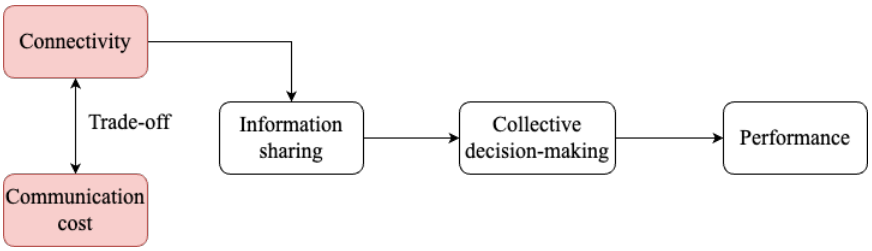


FIGURE 1: TRADE-OFF BETWEEN COMMUNICATION AND COST

2.2.2 Lloyd's Algorithm. Lloyd's algorithm iteratively refines a Voronoi diagram to achieve a more uniform partitioning. The algorithm involves computing the centroid $\mathbf{c}_i$ of each Voronoi cell $V(\mathbf{p}_i)$ and then moving the seed $\mathbf{p}_i$ to $\mathbf{c}_i$ [18]. The centroid $\mathbf{c}_i$ of a Voronoi cell is calculated as:

$$\mathbf{c}_i = \frac{1}{\left|V\left(\mathbf{p}_i\right)\right|} \sum_{\mathbf{x} \in V\left(\mathbf{p}_i\right)} \mathbf{x} \tag{2}$$

After relocating each seed, the Voronoi diagram is updated, and this process continues until the set number of iterations is completed.

3. METHODS

In this section, we introduce the proposed DMABO framework. In our previous work [4], we extended a single agent DSE to a team context with multiple agents. However, our previous model was centralized, with each agent contributing information to a global evaluator based on BO. In this paper, we develop a new distributed MAS for DSE, focusing on empowering individual agents with autonomous decision-making capabilities to identify optimal solutions in their respective search space. This new distributed framework consists of three main components: Problem formulation, communication structure generation, and DMABO implementation.

3.1 Problem Formulation

In this section, we show a problem formulation to find the minimum black-box function in a d dimensional design space $A \subseteq \mathbb{R}^d$ with N distributed agents. The goal of MAS is to find the global minimum $\mathbf{x}^*$.

$$\mathbf{x}^* = \underset{\mathbf{x} \in A}{\operatorname{argmin}}\ f(\mathbf{x}), \tag{3}$$

where $f(\cdot)$ represents a black-box objective function. $\mathbf{x} = (x_1, x_2, ..., x_d) \in \mathbb{R}^d$ is defined as a vector of design variables representing a specific design decision in the d-dimensional space, and it is an element of the set A.

We assume that the design space, denoted as A, is segmented into N distinct local regions, represented as A_i. The specific approach employed for partitioning A is detailed in Section 3.2. In this framework, each agent is allocated a unique area A_i within A. This allocation facilitates the distribution of tasks among agents, ensuring a structured and efficient division of labor. In the context of the DMABO, agents are restricted to their designated local regions; they are not allowed to extend their search beyond the boundaries between local regions. However, agents can communicate and share information with their neighbors. The MAS performance is defined as the convergence speed for the global minimum, i.e., the least number of iterations to find the global minimum for any agent in the system.

3.2 Communication Structure

To examine the effects of communication network structures and associated costs on system performance, we introduce a method for generating communication networks among agents. This method starts with segmenting the design space using Lloyd's algorithm and proceeds to establish communication networks through two models: a range-limited model constrained by neighborhood information and a range-free model without such constraints. We also adopt network density as a metric to quantify communication costs. This indicates that the communication cost is proportional to the number of links established between agents. Fig. 2 and Algorithm 1 outline how to establish a communication network in an MAS by dividing the design space using Lloyd's algorithm, resulting in an adjacency matrix that defines the communication links between agents.

3.2.1 Division of design space. To divide the design space by Lloyd's algorithm, the initial step of the algorithm involves the random placement of seed points, represented as $\mathbf{P} = \left\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N\right\}$. These seeds serve as the foundational points for constructing Voronoi diagrams used for the initial division of the design space. The construction of Voronoi cells $V\left(\mathbf{p}_i\right)$, for each seed $\mathbf{p}_i$ is guided by Eq. (1). Subsequent to the creation of Voronoi cells, the algorithm progresses by computing the centroid $\mathbf{c}_i$ of each cell following Eq. (2). Each seed $\mathbf{p}_i$ is then relocated to its corresponding centroid $\mathbf{c}_i$. This iterative relocation is performed for a predefined number of iterations, denoted as T. After the iterative process is finalized, the established Voronoi cells $V\left(\mathbf{p}_i\right)$ are designated as specific local regions A_i for the MAS.

Lloyd's algorithm is essential because it guides the process toward a more equitable division of the design space. This equitable allocation ensures a balanced workload among all agents, preventing situations where some agents are overloaded due to sampling a larger region, which could affect overall convergence. Furthermore, it helps to control the effect of design space division on performance. Since the scope of this study focuses on assessing the impact of communication structure and cost, it is desired to isolate these factors from the influence of the partition of the design space.

3.2.2 Communication network. In order to analyze the impacts of communication network structures on MAS performance, we developed two models to generate the adjacency matrix $\mathbf{W}$ for the communication networks between agents, where $\mathbf{W} \subseteq \mathbb{R}^{N\times N}$, and N represents the total number of agents. In this matrix, W_{ij} defines if there is a connection between Agent i and Agent j. If $W_{ij} = 1$, there is a connection between the two agents and, otherwise, if $W_{ij} = 0$.

1) *Model 1: Range-limited model constrained by neighborhood information.* This model produces communication networks based on the adjacent matrix of the partitioned regions. From this division of the design space, an adjacency matrix, Φ, representing the adjacent relations between the partitioned regions is constructed, mapping the communication between these regions according to their spatial adjacency. The elements of Φ, ϕ_{ij} reflect the adjacency between Voronoi cells: $\phi_{ij} = 1$ indicates that the local regions A_i and A_j are directly adjacent, thus denoting direct connectivity between the corresponding agents, while $\phi_{ij} = 0$ implies no direct adjacency. Here, the communication network $\mathbf{W}$ is generated with the constraint that only agents who are neighbors can be connected. Therefore, the adjacent matrix $\mathbf{W}$ of the MAS communication network is a subnet of the spatial adjacent matrix Φ of the local regions.

2) *Model 2: Range-free model without neighborhood constraint.* In this model, we randomly connect a pair of nodes to establish a communication link regardless of whether two agents are neighbors or not. Thus, the resulting communication network is independent of the space division. In other words, the adjacent matrix $\mathbf{W}$ of the communication network has no relationship with the spatial adjacent matrix of the partitioned local regions Φ.

To demonstrate the process of generating communication networks based on Lloyd's algorithm with Algorithm 1 and the two models above, an example is presented in Fig. 4. Fig. 3a shows the division of the design space into five Voronoi cells, corresponding to the number of agents. The seeds of these cells are distributed randomly according to the number of agents. After completion of T iterations, segmentation for this design space is achieved following this algorithm, as shown in Fig. 3b. The corresponding adjacency matrix of the partitioned regions, denoted as Φ, is shown below:

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}. \tag{4}$$

Thus, in Model 1, we generate the adjacency matrix $\mathbf{W}$ of communication network constrained by the partitioned regions. To show the example representing Model 1, for the density 0.5 (introduced in Section 3.2.3), the corresponding $\mathbf{W}$ can be generated, and the communication network is shown in Fig. 4a:

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \tag{5}$$

where the adjacency matrix $\mathbf{W}$ is constrained by the partitioned regions.

In Model 2, the corresponding $\mathbf{W}$ is randomly generated and is not correlated with the adjacency relationships among the local regions in this design space. To show the example representing Model 2, the adjacency matrix $\mathbf{W}$ of communication network with density 0.5 is shown in Fig. 4b:

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}. \tag{6}$$

Note that the two models may generate multiple cases, and we are presenting just one example here.

Algorithm 1 COMMUNICATION NETWORK GENERATION BY LLOYD'S ALGORITHM

Randomly initialize the locations of seeds $\mathbf{P}$, where $\mathbf{P} = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N\}$.
for $t = 1$ to T **do**
  Generate Voronoi cells $V(\mathbf{p}_i)$ based on criterion Eq. (1), $i = 1$ to N
  Calculate the centroid $\mathbf{c}_i$ of a Voronoi cell as Eq. (2), $i = 1$ to N
  $\mathbf{p}_i \leftarrow \mathbf{c}_i$, $i = 1$ to N
end for
Generate the adjacency matrix Φ of partitioned regions.
Create the local regions A_i according each Voronoi cell $V(\mathbf{p}_i)$, $i = 1$ to N
Method 1: Generate the adjacency matrix $\mathbf{W}$ of MAS based on the partitioned design space;
Or Method 2: Generate the adjacency matrix $\mathbf{W}$ of MAS by randomly connect a pair of nodes
Output the adjacency matrix $\mathbf{W}$ as the communication network of MAS

3.2.3 Communication cost. To investigate the impact of communication costs on MAS performance, we use network density as a metric to quantify these costs. The density of a network is quantified by the formula $D = \frac{2E}{N(N-1)}$, where E represents the number of edges (connections) and N represents the number of agents (nodes) [19]. This metric, integral to network theory, measures the proportion of existing connections relative to the maximum potential connections among agents. The values range from 0 to 1, where 0 means a network without connections and 1 denotes a fully connected network. Lower density values indicate fewer connections, which can lead to reduced communication costs. However, this reduction in connections could affect the agents' capacity for effective information exchange and consequently the overall performance of an MAS.

In this study, we create various networks by progressively increasing the network density according to the two network generation models. For the resulting networks, we also examine their average clustering coefficients as an altenative measure of network connectivity. The *clustering coefficient* (CC) assesses

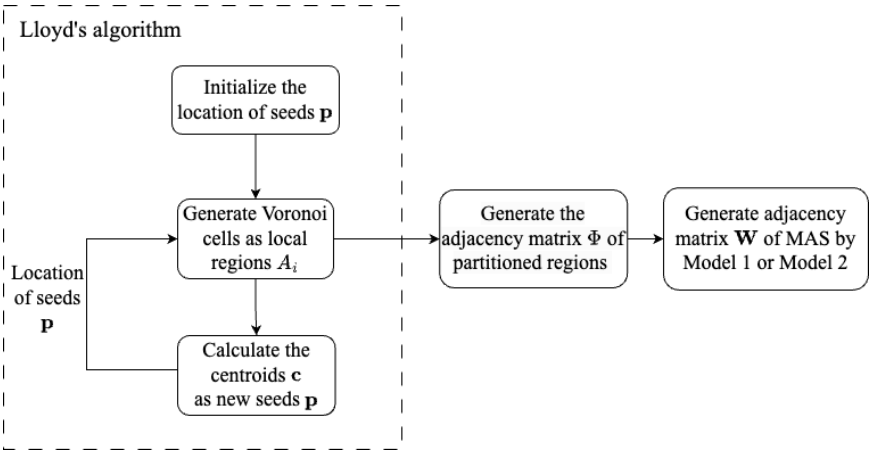
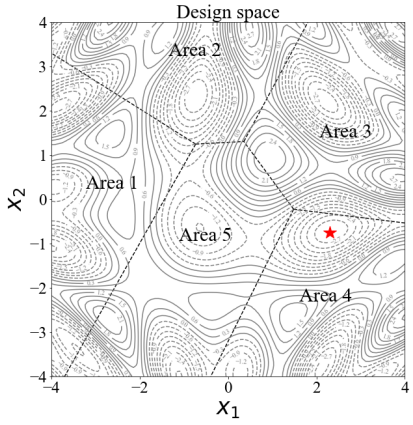
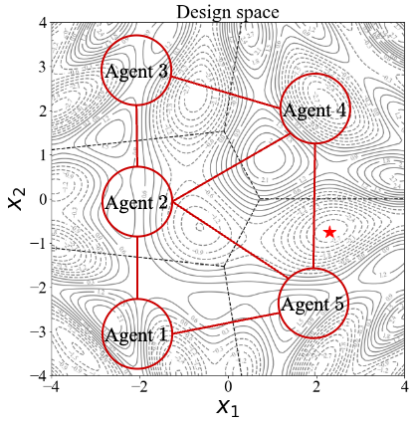


FIGURE 2: COMMUNICATION NETWORK GENERATION BY LLOYD'S ALGORITHM

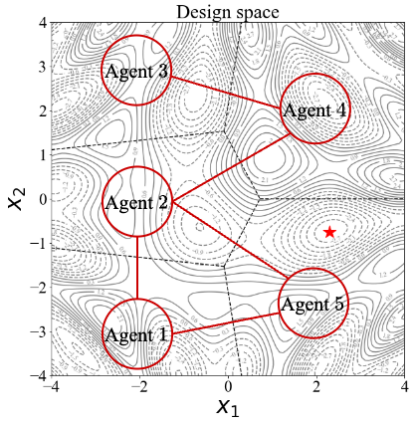


(a) Voronoi diagram

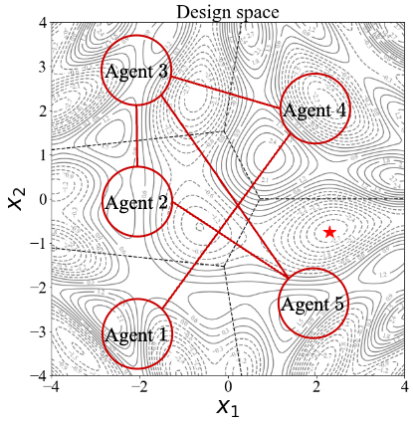


(b) Lloyd's diagram with its communication network of partitioned regions

FIGURE 3: DIVISION OF DESIGN SPACE



(a) Model 1



(b) Model 2

FIGURE 4: EXAMPLES COMMUNICATION NETWORKS BY TWO MODELS, WHERE THE NETWORK DENSITY IS 0.5.

the probability that two neighbors of a specific node are also neighbors, shedding light on the cohesiveness of local groups within the network [20]. The formula for calculating the clustering coefficient for a single node is given by $C_i = \frac{2e_i}{k_i(k_i-1)}$, where C_i is the clustering coefficient of agent i, e_i is the number of connections between the neighbors of agent i, and k_i represents the number of neighbors that agent i has. The average clustering coefficient (ACC) of a network is simply the average clustering coefficient of each node. Since networks with the same density (i.e., the same number of links) may have different CCs, we are interested in knowing the potential confounding effects between density and CCs on an MAS's performance in DSE.

3.3 Distributed Multi-Agent Bayesian Optimization (DMABO)

The proposed DMABO, illustrated in Fig. 5 and detailed in Algorithm 2, presents the decision-making processes of agents for collaborative optimization within a distributed MAS. The fundamental principle of this DMABO framework is to ensure that each agent operates autonomously with its own dataset, maintaining data privacy. Consequently, every agent updates its individual GP model and AF to guide its sampling decisions for the next iterations.

The process begins by initializing a local region (A_i) for each agent within the network and establishing the communication network ($\mathbf{W}$) using Algorithm 1. In this framework, each agent is assigned a GP prior in their respective local region A_i. The algorithm unfolds in iterative steps K for each agent. Agent i starts by updating the GP posterior using all available historical data, encompassing both its local data and the data exchanged with its neighbors according to the communication network $\mathbf{W}$. Subsequently, the agent computes the acquisition function a_i to decide where to sample next. In this study, we chose expected improvement (EI) as the acquisition function. Therefore, the agent's decision is guided by selecting the next sampling point $\mathbf{x}_i^*$ in its local region that maximizes the expected improvement. It is imperative to note that the agent is restricted to sampling exclusively within its own local region and that sampling in regions assigned to others is expressly prohibited. Once the sampling point is determined, the agent evaluates the function f at $\mathbf{x}_i^*$, thus obtaining a new observation $(\mathbf{x}_i^*, f(\mathbf{x}_i^*))$. The agent then shares this new observation with its neighbors from their communication network $\mathbf{W}$, facilitating the collaborative aspect of the search strategy. At the same time, agent i collects data $\mathcal{D}_{ij}^k$ from these neighboring agents.

Using the design space partition and the communication network shown in Fig. 3b as an example, the GP posteriors and AFs shown in Fig. 6 demonstrate one decision step of the DMABO process. Each agent uses its collection of points to construct its own GP model, which estimates the posterior means and variances for the entire design space, as illustrated in the first and second columns of Fig. 6. Following the computation of the GP model, each agent updates the value of its own AF, displayed in the third column of Fig. 6. The agents then identify the point that maximizes their AFs within their local region, marked with red points, and select this point as the next sampling point. The agents proceed to exchange the new point's information with their immediate neighbors, collecting all sampled points from those adjacent to them. Fig. 7 shows this process with points in different colors and numbers that annotate the sequence of the sampling trajectory in ten iterations.

Algorithm 2 DISTRIBUTED MULTI-AGENT BAYESIAN OPTIMIZATION (DMABO)

initialize local region A_i for each agent and generate the team structure $\mathbf{W}$ with Algorithm 1.
Place a Gaussian process prior $\mathcal{D}_i^0$ in A_i for each agent.
for $k = 1$ to K **do**
  for each agent i, $i = 1$ to N **do**
    Update the posterior $\mathcal{GP}_i^k$ using all available data $\mathcal{D}_i^{k-1}$.
    Compute the acquisition function a_i.
    Identify the next sampling point $\mathbf{x}_i^k$ within A_i by optimizing a_i.
    Evaluate the function $f(\mathbf{x}_i^k)$ at $\mathbf{x}_i^k$.
    Share the new evaluations $(\mathbf{x}_i^k, f(\mathbf{x}_i^k))$ with neighboring agents as dictated by the adjacency matrix $\mathbf{W}$.
    Collect data $\mathcal{D}_{ij}^k$ from neighbors.
    $\mathcal{D}_i^k \leftarrow \mathcal{D}_i^{k-1} \cup \mathcal{D}_{ij}^k \cup (\mathbf{x}_i^k, f(\mathbf{x}_i^k))$
  end for
end for
Return a solution: the point $\mathbf{x}^*$ evaluated with the global optimum f^*.

4. EXPERIMENTS

4.1 Experimental settings

To answer the research question, we conducted an experiment using the Cosine function (Fig. 8) as the objective function to test the proposed DMABO framework. Since the investigation on the impact of design complexity on MAS performance is beyond the scope of this study, we use only one objective function and keep it unchanged in this experiment. The MAS team comprises five agents, allowing for the division of the design space into five separate local regions. We have constructed three distributed MAS scenarios (Figs. 9-11) by partitioning the design space and assigning each agent to one local region using Algorithm 1. These partitions and the corresponding adjacent matrix of the partitioned regions are shown in Figs 9a, 10a, 11a.

To assess the impact of communication network structures on MAS performance, we evaluate the convergence speed in three scenarios, using communication networks constructed by the two models described in Section 3.2.2. In Scenario 1, depicted in Fig. 9, the agent, located in the region where the global optimum resides, lacks comprehensive information from the communication network (e.g., Agent 4 is only connected to Agents 2, 3, and 5). Scenario 2, shown in Fig. 10, demonstrates that the agent, who has the ability to access information from all the other agents (i.e., Agent 1 is connected to every agent), is located in the region where the global optimum is. Meanwhile, Scenario 3, presented in Fig. 11, describes a case in which the agent sits in the region where the global optimum exists (Agent 3), but cannot have comprehensive information similar to Scenario 1; however, the design space in Scenario 3 is partitioned differently, producing a different network structure of the partitioned regions.

Furthermore, to assess how network density (i.e., communication cost) impacts system performance, MAS communication networks were generated across a spectrum of densities from 0 to 1, in steps of 0.1. For each incremental step in density, we evaluated the convergence speed. To account for the probabilistic nature due to the random generation process of communication networks, we conducted ten trials at each density level and then reported the mean convergence speed of these ten trials.

4.2 Experimental results

Figures 9 - 11 display the mean convergence speed for each network density in Scenarios 1-3. Each network has a unique

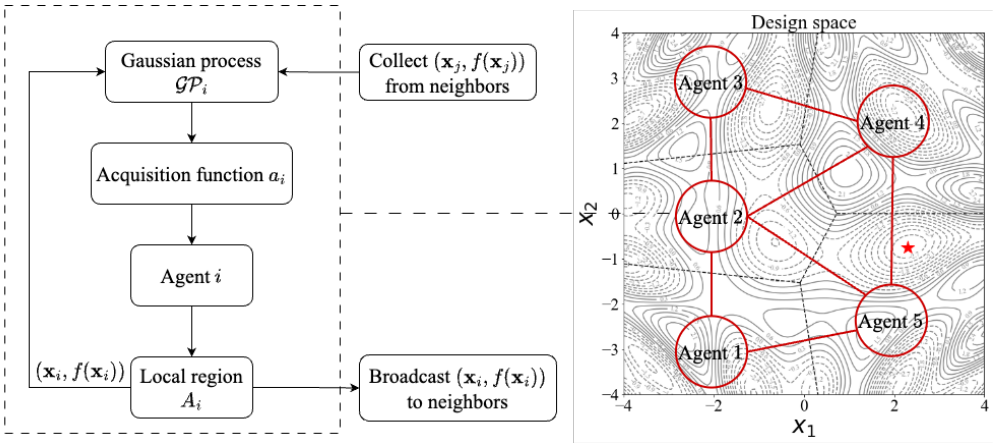


FIGURE 5: DISTRIBUTED MULTI-AGENT BAYESIAN OPTIMIZATION (DMABO)

density and clustering coefficient, and both metrics vary from 0 to 1. Networks with the same density can have different clustering coefficients, indicating that a single level of network density can correspond to multiple clustering coefficients. One point (red color for Model 1 of network generation, as shown in Fig. 9b - 11b and blue color for Model 2 as shown in Fig. 9c - 11c) indicates one particular network's density and its CC. The size of a point is proportional to the convergence speed, and larger points denote faster convergence speeds. Specifically, point sizes of large, medium, and small represent the mean convergence speeds of 19, 12, and 3, respectively. The means and corresponding standard deviations for Scenarios 1-3 are shown in Appendix A. Detailed observations for the three scenarios are as follows:

- Scenario 1: As shown in Fig. 9, a comparison between Model 1 and Model 2 reveals that the incorporation of neighborhood constraints significantly improves the convergence speed, as indicated by the size of the points. Using Model 1 to generate communication networks, the fastest convergence speeds are achieved when the network density is 0.7. Moreover, Model 1 suggests that adding too many links could have an adverse effect, slowing the convergence, especially when the density exceeds 0.7. In contrast, in the communication networks generated using Model 2, the results show that the more links added to the network, the faster the speed could be achieved. However, such an improvement is not as significant as appears in the Model 1 situation, while increasing the network density from 0.0 to 1.0. Moreover, the MAS performance in Model 2 is on average worse than that in Model 1. The best performance in Model 2 is achieved when the network is fully connected, that is, when both the density and clustering coefficients reach 1.0.

- Scenario 2: Upon analyzing Fig. 10, a comparison of Model 1 and Model 2 again demonstrates that the construction of communication networks among agents using the range-limited model with neighborhood constraints significantly enhances the convergence speed (as evidenced by the point sizes). Furthermore, in both models (Fig. 10b and Fig. 10c), there is a tendency for larger points to occur more frequently at large values of density and clustering coefficient, hinting at a possible correlation between higher density and faster convergence speed. In Model 1, the increase in network density from 0.7 to 1.0 results in little improvement in convergence speed. This finding suggests that the best MAS performance is achieved by a communication network with a density of 0.7, and it is unnecessary to spend additional costs to increase connections. Thus, an optimal trade-off between connectivity and cost is realized. Moreover, when comparing Scenarios 1 and 2, it is interestingly observed that even if the two scenarios have the same topology of the spatial network of the partitioned regions, the MAS in Scenario 2 exhibits faster convergence speeds overall. This highlights the importance of the global optimum's location. Specifically, when an agent is located in a region where it connects to all other agents and, by happen, the global optimum also reside in that region, the convergence speed can be significantly accelerated.

- Scenario 3: In Fig. 11, observations similar to Scenario 1 emerge, with the agent positioned in the region where the global optimum exists being not fully connected to all other agents. A good trade-off between connectivity and cost is found at a network density of 0.8. The primary difference between this scenario and Scenario 1 is that the partition of the design space is different, thus producing a different network structure of the partitioned regions. Upon comparing the convergence speeds, as denoted by the point sizes, Scenario 3 demonstrates faster convergence on average than that in Scenario 1.

To conclude, we have three key observations by comparing the three scenarios:

- Comparing Scenarios 1 and 2, it is observed that regardless of the location of the global optimum, a good trade-off is needed between density and communication cost to find the best communication network for the fastest convergence speed.

- Comparing Scenarios 1 and 2, if an agent is located in the area where the global optimum exists and this agent can

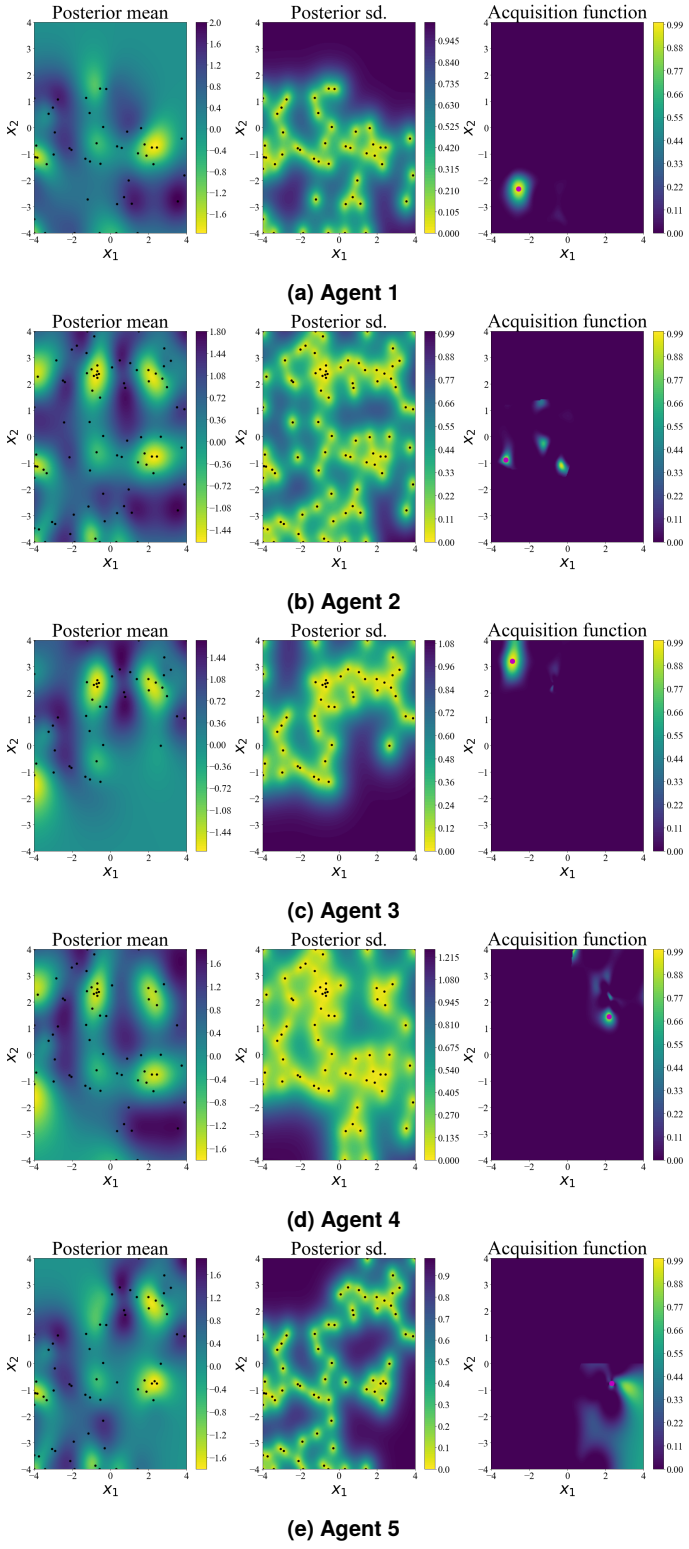


FIGURE 6: THE GP MODELS AND ACQUISITION FUNCTIONS FOR AGENTS 1, 2, 3, 4 AND 5.

exchange information with all other agents, the convergence speed would become faster on average in all communicate networks tested regardless of their density.

- Comparing Scenarios 1 and 3, it is observed that no matter how the design space is partitioned (i.e., regardless of the network structure of the partitioned regions), a good trade-off between density and communication cost can be achived.

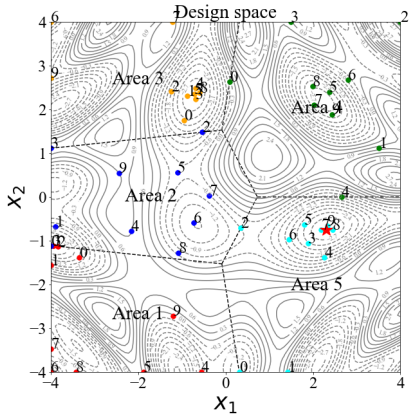


FIGURE 7: SAMPLING PROCESS FOR FIVE AGENTS. THE RED STAR IS THE TRUE OPTIMUM. THE NUMBERS LABELED ON THE POINTS ARE THE SEARCH IN EACH ITERATION STEP.

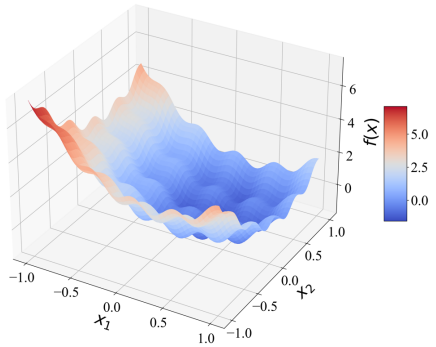


FIGURE 8: COSINES FUNCTION

5. DISCUSSION

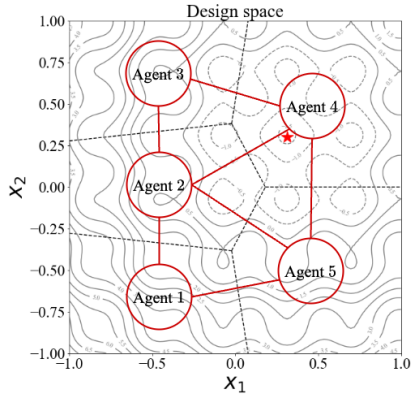
To explore the influence of communication networks and costs on distributed MAS performance, we investigated two methods across various network densities. Analyzing the experimental results enables us to identify the optimal communication structure and its associated costs for best performance. The experimental findings generated by the DMABO framework yield three key insights:

First, if possible, agents should communicate with neighbors who share the correlated information in a design space. The results imply that an improvement of the MAS performance can be achieved by encouraging agents to communicate primarily with their neighboring agents in the adjacent regions. Method 1 consistently outperforms Method 2 in terms of convergence speed as both density and clustering coefficient increase. This is likely attributed to the fact that the information shared between adjacent agents is closely correlated from the design space they are searching, and the continuity of the space between adjacent regions allows agents to better construct the surrogate model and AF.

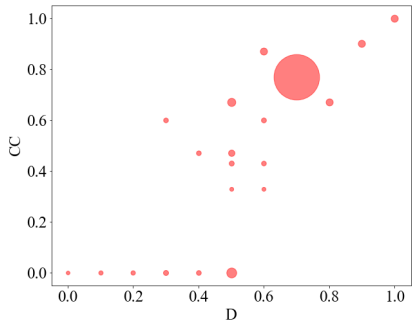
Second, increasing the density of a communication network does not necessarily improve MAS performance in DSE. When

examining the impact of communication costs on performance, quantified by network density, we observe a different trend in Method 1 and 2 with increasing density. In Method 1, we can identify a balance between communication costs and connectivity. Specifically, Method 1 demonstrates that, adhering to neighborhood constraints, a network with a density of 0.7-0.8 helps achieve the best MAS performance. Increasing the density beyond this point could hamper convergence. For example, adopting a fully connected network does not offer additional benefits. The potential reason is explained in the first insight above. In fact, sharing uncorrelated information can adversely affect system performance. However, in Method 2, when agents are free to communicate with all the other agents without neighborhood constraints, enabling full communication across the network helps to achieve the best performance while incurred the highest communication cost. This is because any networks with lower density may have links connects agents who are likely not neighbors, and the information exchange between these agents can even negatively influence the system performance.

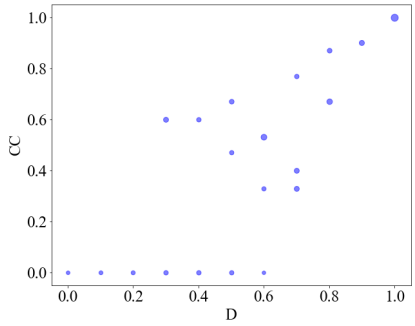
Third, the convergence speed to the global optimum is dependent on its location within the design space. For example, by comparing Scenarios 1 and 2, we can see if an agent is in the region where the global optimum exists and meanwhile can exchange information with all other agents, the system can have a great improvement in convergence speed, across all cases of network densities and clustering coefficients (by comparing the point sizes using the same method in Scenarios 1 and 2). This is because an agent located at or near the global optimum can access the most useful information that helps the exploitation of the design space.



(a) Segmentation 1

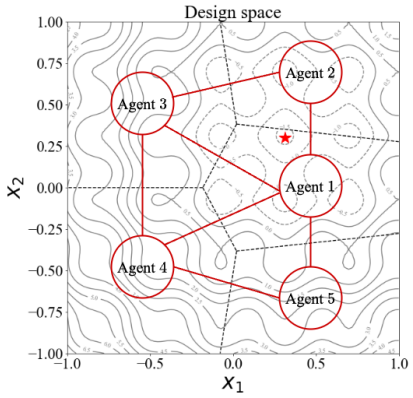


(b) Mean of convergence speed by Model 1

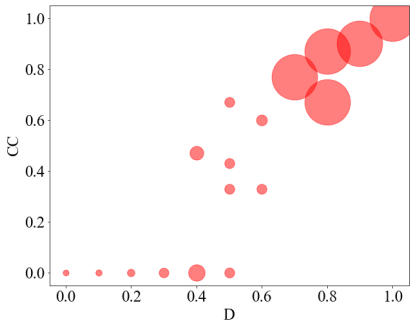


(c) Mean of convergence speed by Model 2

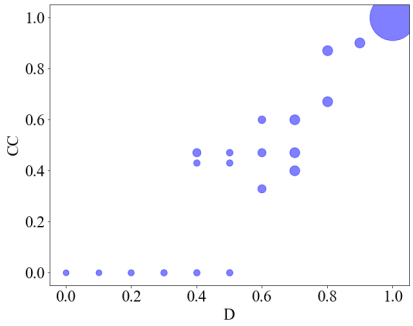
FIGURE 9: SCENARIO 1



(a) Segmentation 2



(b) Mean of convergence speed by Model 1



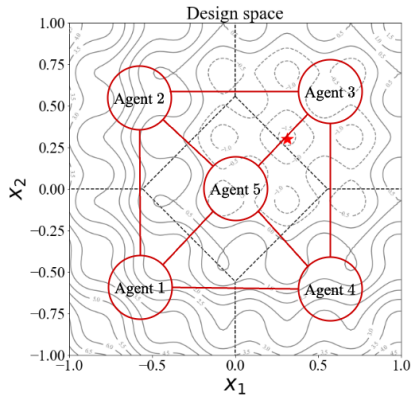
(c) Mean of convergence speed by Model 2

FIGURE 10: SCENARIO 2

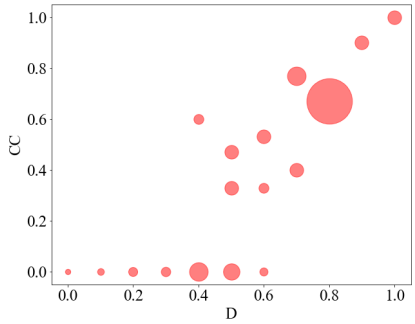
6. CONCLUSION

In order to analyze the impact of communication networks and costs on the performance of distributed MAS, our study developed a DMABO framework that enhances collective decision-making by segmenting the design space, while achieving the trade-off between connectivity and communication cost. Utilizing Lloyd's algorithm, we partitioned the design space, allocating specific regions to agents within the distributed MAS. To analyze the impact of communication networks among agents on performance, we proposed two distinct models: a range-limited model that forms networks based on the adjacency matrix of partitioned regions and a range-free model that established random communication links without neighborhood constraints. Furthermore, we adopted network density as a measure to quantify communication costs, enabling us to explore the impact of communication costs on MAS performance in DSE. Our findings indicated that the range-limited communication model significantly improved performance while effectively managing communication costs. These results offered valuable insights into achieving the trade-off between the communication structure and its associated costs for the best performance in MAS for DSE.

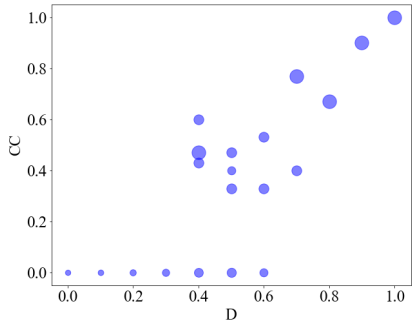
Future work of this study could take several directions to address its limitations and further explore the potential of DMABO. Here are some avenues for future research. First, future studies could investigate how different types of objective functions, especially those with varying levels of complexity, affect the performance of DMABO. Second, our current research was limited to a team size of five agents. Second, we plan to examine the effects of increasing the number of agents on the efficiency and communication costs of the DMABO framework. Third, while this study used Lloyd's algorithm to divide the design space, alternative methods could offer different advantages. Future research could compare various space division methods and study their influence on the system's performance. For example, one could use preexisting design data to inform the initial space division or dynamically reallocate agents based on the data that agents collect in real time to improve overall performance.



(a) Segmentation 3



(b) Mean of convergence speed by Model 1



(c) Mean of convergence speed by Model 2

FIGURE 11: SCENARIO 3

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from the National Science Foundation through the grants CMMI-2321463 and CMMI-2419423.

REFERENCES

[1] Nardi, Luigi, Koeplinger, David and Olukotun, Kunle. "Practical design space exploration." *2019 IEEE 27th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*: pp. 347–358. 2019. IEEE.

[2] Snoek, Jasper, Larochelle, Hugo and Adams, Ryan P. "Practical bayesian optimization of machine learning algorithms." *Advances in neural information processing systems* Vol. 25 (2012).

[3] Icard, Thomas F. "Bayes, bounds, and rational analysis." *Philosophy of Science* Vol. 85 No. 1 (2018): pp. 79–101.

[4] Chen, S., Bayrak, A. E. and Sha, Z. "Multi-Agent Bayesian Optimization for Unknown Design Space Exploration." *ASME 2023 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*.

[5] Yue, Xubo, Kontar, Raed Al, Berahas, Albert S, Liu, Yang, Zai, Zhenghao, Edgar, Kevin and Johnson, Blake N. "Collaborative and Distributed Bayesian Optimization via Consensus: Showcasing the Power of Collaboration for Optimal Design." *arXiv preprint arXiv:2306.14348* (2023).

[6] Yang, Zewen, Sosnowski, Stefan, Liu, Qingchen, Jiao, Junjie, Lederer, Armin and Hirche, Sandra. "Distributed learning consensus control for unknown nonlinear multi-agent systems based on gaussian processes." *2021 60th IEEE Conference on Decision and Control (CDC)*: pp. 4406–4411. 2021. IEEE.

[7] Dai, Xiaobing, Yang, Zewen, Xu, Mengtian and Hirche, Sandra. "Distributed Event-Triggered Online Learning for

Multi-Agent System Control using Gaussian Process Regression." *arXiv preprint arXiv:2304.05138* (2023).

[8] Klaesson, Filip. "Distributed Bayesian Optimization in Multi-Agent Systems." (2020).

[9] Russo, Daniel J, Van Roy, Benjamin, Kazerouni, Abbas, Osband, Ian, Wen, Zheng et al. "A tutorial on thompson sampling." *Foundations and Trends® in Machine Learning* Vol. 11 No. 1 (2018): pp. 1–96.

[10] Hernández-Lobato, José Miguel, Requeima, James, Pyzer-Knapp, Edward O and Aspuru-Guzik, Alán. "Parallel and distributed Thompson sampling for large-scale accelerated exploration of chemical space." *International conference on machine learning*: pp. 1470–1479. 2017. PMLR.

[11] Garcia-Barcos, Javier and Martinez-Cantin, Ruben. "Fully distributed Bayesian optimization with stochastic policies." *arXiv preprint arXiv:1902.09992* (2019).

[12] Young, M Todd, Hinkle, Jacob, Ramanathan, Arvind and Kannan, Ramakrishnan. "Hyperspace: Distributed bayesian hyperparameter optimization." *2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*: pp. 339–347. 2018. IEEE.

[13] Peralta, Federico, Reina, Daniel Gutierrez and Toral, Sergio. "Water quality online modeling using multi-objective and multi-agent Bayesian Optimization with region partitioning." *Mechatronics* Vol. 91 (2023): p. 102953.

[14] Han, Yang, Zhang, Ke, Li, Hong, Coelho, Ernane Antônio Alves and Guerrero, Josep M. "MAS-based distributed coordinated control and optimization in microgrid and microgrid clusters: A comprehensive overview." *IEEE Transactions on Power Electronics* Vol. 33 No. 8 (2017): pp. 6488–6508.

[15] Wang, Xilu, Jin, Yaochu, Schmitt, Sebastian and Olhofer, Markus. "Recent advances in Bayesian optimization." *ACM Computing Surveys* Vol. 55 No. 13s (2023): pp. 1–36.

[16] Frazier, Peter I. "A tutorial on Bayesian optimization." *arXiv preprint arXiv:1807.02811* (2018).

[17] Aurenhammer, Franz. "Voronoi diagrams—a survey of a fundamental geometric data structure." *ACM Computing Surveys (CSUR)* Vol. 23 No. 3 (1991): pp. 345–405.

[18] Lloyd, Stuart. "Least squares quantization in PCM." *IEEE transactions on information theory* Vol. 28 No. 2 (1982): pp. 129–137.

[19] De Laat, Maarten, Lally, Vic, Lipponen, Lasse and Simons, Robert-Jan. "Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis." *International Journal of Computer-Supported Collaborative Learning* Vol. 2 (2007): pp. 87–103.

[20] Holland, Paul W and Leinhardt, Samuel. "Transitivity in structural models of small groups." *Comparative group studies* Vol. 2 No. 2 (1971): pp. 107–124.

APPENDIX A. THE STANDARD DEVIATIONS OF THE CONVERGENCE SPEED IN THE SCENARIO 1-3.

TABLE 1: MEAN AND STD. OF CONVERGENCE SPEED BY MODEL 1, SCENARIO 1.

D ,	0	0.1	0.2	0.3		0.4		0.5					0.6			0.7	0.8	0.9	1
CC	0	0	0	0	0.6	0	0.47	0	0.33	0.43	0.47	0.67	0.33	0.6	0.87	0.77	0.67	0.9	1
Sample size	10	10	10	7	3	6	4	2	4	1	2	1	1	3	3	10	10	10	10
Mean	19	15	14.33	11.8	13	12.8	13	6	17	12	9	9	17	12.33	8	4	8	8	8
Std.	0	2.31	2.92	2.13	0	3.05	0	0	0	0	3.3	0	0	3.3	0	0	0	0	0

TABLE 2: MEAN AND STD. OF CONVERGENCE SPEED BY MODEL 2, SCENARIO 1.

D	0	0.1	0.2	0.3		0.4		0.5			0.6			0.7		0.8		0.9	1
CC	0	0	0	0	0.6	0	0.6	0	0.47	0.67	0	0.33	0.53	0.33	0.4	0.77	0.67	0.87	0.9	1
Sample size	10	10	10	7	3	5	5	3	5	2	2	3	5	3	4	3	6	4	10	10
Mean	19	17	16	14	12	14	14	14.5	14.76	13	20	15	10	12	12	14	10.5	13	11.6	8
Std.	0	1.67	3.10	1.10	0	3.67	0	4.5	0.47	0	0	0	0	0	1.63	0	2.06	0	0.8	0

TABLE 3: MEAN AND STD. OF CONVERGENCE SPEED BY MODEL 1, SCENARIO 2.

D	0	0.1	0.2	0.3	0.4		0.5				0.6		0.7	0.8		0.9	1	0.9	1
CC	0	0	0	0	0	0.47	0	0.33	0.43	0.47	0.33	0.6	0.77	0.67	0.87	0.9	1	0.9	1
Sample size	10	10	10	10	6	4	3	2	2	3	2	4	10	5	5	10	10	10	10
Mean	12	11.33	11	8.71	6.67	9	8	8	8	8	8	7.67	6	6	6	6	6	11.6	8
Std.	0	1.10	1.63	1.67	0.94	1.41	0	0	0	0	0	0.47	0	0	0	0	0	0.8	0

TABLE 4: MEAN AND STD. OF CONVERGENCE SPEED BY MODEL 2, SCENARIO 2.

D	0	0.1	0.2	0.3	0.4			0.5			0.6		0.7			0.8		0.9	1
CC	0	0	0	0	0	0.43	0.47	0	0.47	0.67	0.33	0.47	0.4	0.47	0.6	0.67	0.87	0.9	1
Sample size	10	10	10	7	6	2	2	5	3	2	3	5	3	4	3	7	3	10	10
Mean	12	12	11.4	11	11	11	9	11	11	10.67	9	9	8	8	8	8	8	8	6
Std.	0	0	0.8	0.82	0.82	0	2.83	0	0	1.25	0	0	0	0	0	0	0	0	0

TABLE 5: MEAN AND STD. OF CONVERGENCE SPEED BY MODEL 1, SCENARIO 3.

D	0	0.1	0.2	0.3	0.4		0.5			0.6		0.7		0.8	0.9		1
CC	0	0	0	0	0	0.6	0	0.33	0.43	0.33	0.4	0.4	0.77	0.67	0.9	1	1
Sample size	10	10	10	10	6	4	4	4	2	3	3	3	7	10	10	10	10
Mean	14	12.8	12.2	9.2	7.5	9	7.67	8	8	10	9	8	7.5	7	8	8	8
Std.	0	1.47	1.47	1.72	0.5	0	0.47	0	0	0	0	0	0.5	0	0	0	0

TABLE 6: MEAN AND STD. OF CONVERGENCE SPEED BY MODEL 2, SCENARIO 3.

D	0	0.1	0.2	0.3	0.4				0.5		0.6		0.7		0.8		1
CC	0	0	0	0	0	0.43	0.47	0.6	0	0.47	0	0.53	0.4	0.77	0.67	0.9	1
Sample size	10	10	10	10	5	3	1	1	5	1	4	3	5	5	10	10	10
Mean	14	12.8	12.2	10.8	9.5	9	8	9	9.33	9	10	9	9	8	8	8	8
Std.	0	1.47	1.47	0.75	0.5	0	0	0	0.47	0	0	0	0	0	0	0	0