2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Multi-Agent Global Optimization with Decision Variable Coupling

Lorenzo Sabug, Jr., Fredy Ruiz, *Senior Member, IEEE,* and Lorenzo Fagiano, *Senior Member, IEEE*

***Abstract*— A cooperative, multi-agent global optimization problem is considered, where the global cost function is the sum of the agents' private, non-convex costs. In contrast to all previously considered setups, evaluating the private costs involves a *global* experiment, using a common instance of the decision vector. This is relevant when each agent can only control a part ("subvariable") of the decision vector, but its private cost is also affected by the other subvariables. A novel cooperative optimization method using Set Membership identification and consensus-based techniques is proposed, to make all agents agree on the next global decision vector to be tested. A trade-off between exploitation close to the best point found and exploration around the search set is achieved, even without explicitly sharing the private costs' information. Statistical tests show that the proposed distributed method is competitive with respect to a centralized one.**

I. Introduction

Due to technical demands brought by the rise of sensor network systems, multi-area power control, and multi-agent systems, distributed optimization is subject of strong interest by academia and industry. Different problem setups and techniques have been considered in the literature [1], [2], which can be grouped in two problem setups. The first, with a survey in [2], is where all agents $\mathcal{A}_i$ evaluate their own (not necessarily identical) copies of the decision vector, in a setup which we refer to as parallel evaluation. The second one involves agents controlling their respective subvariables $\boldsymbol{x}_i$ of $\boldsymbol{x}$, i.e., $\boldsymbol{x} = (\boldsymbol{x}_i)_{i=1}^{N_\mathcal{A}}$ with $N_\mathcal{A}$ being the number of agents. Furthermore, each agent has its own cost function, decoupled from the others, i.e. $f_i(\boldsymbol{x}_i)$. The cost functions are thus evaluated separately, either sequentially, or in parallel. For problems where $\boldsymbol{x}_i$ are coupled by linear equalities, augmented Lagrangian methods (ALM) and alternating direction method of multipliers (ADMM) are highly popular. ADMM-based methods are proposed for many problem variations, for which reviews [3] and [4] provide more details.

In some important cases, however, the private cost of an agent depends also on the decisions of other agents, and the cost functions can be evaluated only altogether by means of a (often expensive and/or time-consuming) *global* experiment or process as illustrated in Fig. 1. This can be in cases where each agent has control of only a subset $\boldsymbol{x}_i$ of the global decision variable $\boldsymbol{x}$, but the entire $\boldsymbol{x}$ affects the agent's respective f_i. For instance, in wind farms, the positioning combination of *all* turbines can affect the performance of *each* turbine, and consequently the overall farm power output. Furthermore, in complex factories, the respective performance of all machines are affected by a commonly adopted setting, in an *a priori* unknown and non-trivial fashion. In such cases, all agents have to agree on the global decision variable $\boldsymbol{x}$, before performing the experiment. In addition, the prohibitive evaluation costs limits the calculation/estimation of gradients using additional experiments. This problem is related to continuous distributed constraint optimization (C-DCOP)[1] in the multi-agent systems literature [5], [6], in the sense that agents must agree on a global decision variable (DCOP literature assumes that each $\mathcal{A}_i$ controls a distinct subvariable $\boldsymbol{x}_i$); however, C-DCOP involves the explicit sharing of cost evaluations among agents. Federated learning (FL)-based techniques have also recently attracted interest, as shown in the reviews [7] and [8]. At the heart of FL is a *central server* (CS) which coordinates the optimization of the mean of the individual functions of the agents (or as called in the FL literature, "clients", which tend to be in the order of thousands). The central server sets the test variable $\boldsymbol{x}$ to a subset of its clients, and collects the resulting function values. An obvious drawback to this approach is the CS being a single point of failure, for which a temporary malfunction can lead to failure of the optimization process. Furthermore, FL allows that the clients can set their own $\boldsymbol{x}$ independently, and perform *local* experiments/evaluations.

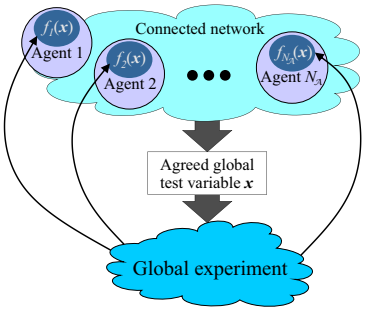

Fig. 1. Problem context considered in this paper.

The setup we consider in this paper is different from the literature in three main aspects. First, only a single vector $\boldsymbol{x}$ is evaluated in a *global* experiment (in contrast to problem setups for parallel evaluation or FL, where agents can evaluate the global objective using differently-valued

All authors are with Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy. e-mail: {lorenzojr.sabug, fredy.ruiz, lorenzo.fagiano}@polimi.it.

This research has been supported by Fondazione Cariplo under grant n. 2022-2005, project "NextWind - Advanced control solutions for large scale Airborne Wind Energy Systems", by the Italian Ministry of Enterprises and Made in Italy in the framework of the project 4DDS (4D Drone Swarms) under grant no. F/310097/01-04/X56., and by the Italian Ministry of University and Research under grants "P2022927H7 - DeepAirborne - Advanced Modeling, Control and Design Optimization Methods for Deep Offshore Airborne Wind Energy" and "P2022EXP2W Learning-based Model Predictive Control by Exploration and Exploitation in Uncertain Environments".

[1]The term "constraint" in DCOP literature refers to the cost functions that 'constrain' the selection of the decision variable values among agents, and should not be confused with constraints as used in the optimization literature.

copies/instances of $\boldsymbol{x}$). Second, the entire $\boldsymbol{x}$ can affect the agent's f_i, and not just the agent's respective subset of variables. This is in contrast to ALM/ADMM-associated setups where each agent can evaluate f_i independently based on its subvariable $\boldsymbol{x}_i$. Third, our work differs from current C-DCOP literature because agents need to agree on the next decision variable $\boldsymbol{x}$ *without* sharing their respective evaluations; i.e., we consider what is called in DCOP literature as *internal constraint privacy* [9]. Furthermore, this consensus needs to be performed before the global experiment execution, while DCOP can allow agents to independently evaluate their costs depending on their couplings.

Responding to our currently-considered problem, we resort to using a zeroth-order technique; specifically, we propose a new multi-agent-based method based on our previous work on Set Membership Global Optimization (SMGO) [10], [11]. In this paper, we formulate a multi-agent version, dubbed MA-SMGO, to address the challenge of deciding the global decision vector, while not having to divulge the private functions, nor the evaluation values of the agents. MA-SMGO is built on the structure of SMGO, which trades off between exploitation and exploration, with the addition of consensus-based steps to ensure that such trade-off is related to the global objective, and not just any agent's private objective. We derive the theoretical convergence properties of the distributed approach, and subject it to statistical tests, showing that it is in fact very competitive when compared to the centralized version.

This paper is organized in seven sections. We discuss nomenclature, the problem setup and formulation, and the Set Membership approach in Section II. Our proposed multi-agent method is discussed with its subroutines in Section III, and its theoretical properties are discussed in Section IV. We show an illustrative test in Section V, and present benchmark tests and comparative analysis in Section VI. Finally, we provide concluding remarks in Section VII.

II. Preliminaries

A. Problem Formulation

We consider a set of agents $\boldsymbol{\mathcal{A}} = \{\mathcal{A}_i, i = 1, ..., N_{\boldsymbol{\mathcal{A}}}\}$ aiming to minimize a global black-box function f which is the sum of their respective private functions f_i:

$$f(\boldsymbol{x}) = \sum_{i=1}^{N_{\boldsymbol{\mathcal{A}}}} f_i(\boldsymbol{x}),\ \forall \boldsymbol{x} \in \mathcal{X}. \tag{1}$$

where $\boldsymbol{x} \in \mathcal{X}$ is the decision vector, and $\mathcal{X} \subset \mathbb{R}^D$ is a compact, convex search set. We suppose that the search set $\mathcal{X}$ is known to all agents in $\boldsymbol{\mathcal{A}}$. In most cases, $\mathcal{X}$ is a hyperrectangle composed of the respective ranges of the individual tuning variables, i.e. $x_d \in [\underline{x}_d, \overline{x}_d], d = 1, \ldots, D$.

We assume that each $\mathcal{A}_i$ has no *a priori* knowledge regarding the functional form of f_i, or even its convexity nor its unimodality (presence of a unique global minimum). Instead, we take an assumption regarding the regularity of each f_i:

Assumption 1: All functions f_i are locally Lipschitz continuous in the search set $\mathcal{X}$, with their respective finite (but unknown) constants γ_i, i.e. $f_i \in \mathcal{F}(\gamma_i), i = 1, \ldots, N_{\boldsymbol{\mathcal{A}}}$ where $\mathcal{F}(\eta) \doteq \left\{h : |h(\boldsymbol{x}) - h(\boldsymbol{y})| \leq \eta \|\boldsymbol{x} - \boldsymbol{y}\|,\ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}\right\}$.

In addition to the assumptions regarding the black-box functions, we take two assumptions on the communication topology of the agents:

Assumption 2: The network of $\boldsymbol{\mathcal{A}}$ is fully connected.

Assumption 3: The evaluation time of all private black-box functions $f_i(\boldsymbol{x})$, and consequently, the global $f(\boldsymbol{x})$, is long enough so that the time devoted to inter-agent communication is negligible.

Assumptions 2 and 3 are highly reasonable for real-world problems where the evaluation-related experiments and/or simulations run on a time-scale much larger than the time required by the agents to successfully complete their communications, even with packet delays/losses and their appropriate recovery mechanisms. Given the currently (and cheaply) available communication technologies, this situation is often found in practice.

We consider an iterative optimization procedure to approximately find a global minimizer of $f(\boldsymbol{x})$, and denote with n the iteration number. A crucial assumption of this paper, different from the literature, follows:

Assumption 4: At any iteration n all private functions f_i are evaluated using the same decision variable $\boldsymbol{x}^{(n)}$, which is known to all agents, that is,

$$z_i^{(n)} = f_i(\boldsymbol{x}^{(n)}) + \epsilon_i^{\langle n \rangle}, i = 1, \cdots, N_{\boldsymbol{\mathcal{A}}},$$

with $\epsilon_i^{\langle n \rangle}$ Unknown But Bounded (UBB) additive noise with finite but unknown bounds $\overline{\epsilon}_i$ such that $|\epsilon_i^{\langle n \rangle}| \leq \overline{\epsilon}_i$.

This assumption means that while all functions values are revealed independently (and privately) to each agent, the evaluated variable $\boldsymbol{x}^{(n)}$ is global, and hence is the main coupling mechanism among all agents.

Each $\mathcal{A}_i$ maintains a data set of samples up to iteration n,

$$\boldsymbol{X}_i^{\langle n \rangle} \doteq \left\{\mathring{\boldsymbol{x}}_i^{(1)}, \ldots, \mathring{\boldsymbol{x}}_i^{(n)}\right\}, \tag{2}$$

with each sample $\mathring{\boldsymbol{x}}_i^{(k)}$ being a tuple of the evaluated point and its corresponding f_i measurement

$$\mathring{\boldsymbol{x}}_i^{(k)} \doteq \left(\boldsymbol{x}^{(k)}, z_i^{(k)}\right). \tag{3}$$

We now define the best sampled point

$$\boldsymbol{x}^{*\langle n \rangle} \doteq \arg \min_{k=\{1,\ldots,n\}} \sum_{i=1}^{N_{\boldsymbol{\mathcal{A}}}} z_i^{(k)}, \tag{4}$$

which is the sampled point that produced the best *global* objective f (which is a sum of components, see (1)). The best sampled value w.r.t. each $\mathcal{A}_i$ is chosen as

$$z_i^{*\langle n \rangle} = z_i^{(k)},\ \text{s.t.}\ \left(\boldsymbol{x}^{*\langle n \rangle}, z_i^{(k)}\right) \in \boldsymbol{X}_i^{\langle n \rangle}, \tag{5}$$

which we note is not necessarily the best sampled value for the individual cost of $\mathcal{A}_i$, but is rather associated with $\boldsymbol{x}^{*\langle n \rangle}$ in the same tuple (see (4)). Accordingly, we define the absolute optimum value as $z^* = \min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$. We are now in a position to state the problem considered in this paper.

Problem 1: Given Assumptions 1-4, build a distributed method to generate a sequence of the global decision variable $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(3)}, \ldots$, such that for any desired finite precision

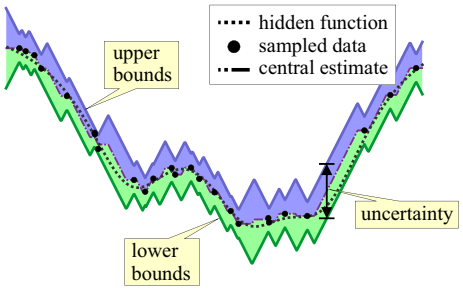

Fig. 2. Set Membership-based model and quantities from a data set.

$\varepsilon > 0$, there exists a finite number of (synchronous and global) evaluations $n_\varepsilon < \infty$ such that

$$z^{*\langle n_\varepsilon \rangle} - f(\boldsymbol{x}^*) < \varepsilon.$$

The problem is particularly challenging because no agent has complete information regarding f, due to the privacy assumption. This means that even after sampling, $\mathcal{A}_i$ can only infer information on its own f_i, but cannot do so on the other component functions f_j, which is an added complication to the one in centralized ("conventional") black-box optimization. Hence, some coordination is required among agents to ensure that the exploitation and exploration of the global objective f is done and not just of any private f_i.

B. Set Membership (SM) approach overview

For self-consistency, we briefly recall the Set Membership (SM) function estimation approach [12], recently generalized in the kinky inference framework [13], to build approximate models of the hidden functions f_i given $\boldsymbol{X}_i^{\langle n \rangle}$ (see (2)-(3)) and Assumption 1. From the data set $\boldsymbol{X}_i^{\langle n \rangle}$, a noise bound estimate $\tilde{\epsilon}_i^{\langle n \rangle}$, and a Lipschitz constant estimate $\tilde{\gamma}_i^{\langle n \rangle}$, we can build the following upper and lower bounds to f_i:

$$\overline{f}_i^{\langle n \rangle}(\boldsymbol{x}) = \min_{\mathring{\boldsymbol{x}}^{(k)} \in \boldsymbol{X}_i^{\langle n \rangle}} \left(z_i^{(k)} + \tilde{\epsilon}_i^{\langle n \rangle} + \tilde{\gamma}_i^{\langle n \rangle} \left\| \boldsymbol{x} - \boldsymbol{x}^{(k)} \right\| \right), \quad (6)$$

$$\underline{f}_i^{\langle n \rangle}(\boldsymbol{x}) = \max_{\mathring{\boldsymbol{x}}^{(k)} \in \boldsymbol{X}_i^{\langle n \rangle}} \left(z_i^{(k)} - \tilde{\epsilon}_i^{\langle n \rangle} - \tilde{\gamma}_i^{\langle n \rangle} \left\| \boldsymbol{x} - \boldsymbol{x}^{(k)} \right\| \right). \quad (7)$$

If $\tilde{\epsilon}_i^{\langle n \rangle}$ and/or $\tilde{\gamma}_i^{\langle n \rangle}$ are not *a priori* known, these can be estimated from $\boldsymbol{X}_i^{\langle n \rangle}$ itself, with the method proposed in [14]. Aside from the bounds (6)-(7), we can also quantify an estimate for f_i at iteration n as

$$\tilde{f}_i^{\langle n \rangle}(\boldsymbol{x}) = \frac{1}{2} \left(\overline{f}_i^{\langle n \rangle}(\boldsymbol{x}) + \underline{f}_i^{\langle n \rangle}(\boldsymbol{x}) \right). \quad (8)$$

Lastly, the uncertainty at any $\boldsymbol{x}$ is given as

$$\lambda_i^{\langle n \rangle}(\boldsymbol{x}) = \overline{f}_i^{\langle n \rangle}(\boldsymbol{x}) - \underline{f}_i^{\langle n \rangle}(\boldsymbol{x}). \quad (9)$$

A simple illustration of the SM-based quantities can be found in Fig. 2.

III. Proposed Consensus-Based Multi-Agent Optimization Method

To address Problem 1, we propose additional routines to the Set Membership Global Optimization (SMGO) algorithm [11], to accommodate the special setup described in Section II. Agents in $\boldsymbol{\mathcal{A}}$ perform consensus operations to agree on the main operation modes of SMGO, and consequently, also on the next sampled point $\boldsymbol{x}^{(n+1)}$. The following discussion is in the viewpoint of an ego agent $\mathcal{A}_i$.

A. Data set update and identification of best point

Consider the arrival of a new sample $\mathring{\boldsymbol{x}}_i^{(n)}$, given by the relevant entries of the (globally) evaluated point $\boldsymbol{x}^{(n)}$, and the (private) function value $z_i^{(n)}$ (see (3)). The updated $\boldsymbol{X}_i^{\langle n \rangle}$ is iteratively built as $\boldsymbol{X}_i^{\langle n \rangle} = \boldsymbol{X}_i^{\langle n-1 \rangle} \cup \mathring{\boldsymbol{x}}_i^{(n)}$. We now proceed to identify the best sampled point $\boldsymbol{x}^{*\langle n \rangle}$. As $\boldsymbol{x}^{*\langle n \rangle}$ relates with the *global* objective f, a consensus is done with other agents $\mathcal{A}_j$ to decide whether to update $\boldsymbol{x}^{*\langle n \rangle}$, while hiding information on f_i. Agent $\mathcal{A}_i$ determines the improvement of the new local value w.r.t. the existing best sampled one $z_i^{*\langle n \rangle}$, $\delta_i^{\langle n \rangle} = z_i^{*\langle n-1 \rangle} - z_i^{(n)}$. The calculated improvement $\delta_i^{\langle n \rangle}$ is then broadcasted to all $\mathcal{A}_j$ via *anonymous*[2] packets, and in turn, $\mathcal{A}_i$ collects all $\delta_j^{\langle n \rangle}$. The net (global) improvement is then computed as

$$\delta^{\langle n \rangle} = \overbrace{\delta_i^{\langle n \rangle}}^{\substack{\text{improvement} \\ \text{w.r.t. } f_i}} + \overbrace{\sum_j \delta_j^{\langle n \rangle}}^{\substack{\text{improvement values} \\ \text{collected from all } \mathcal{A}_j}}.$$

If the net improvement is positive, $\mathcal{A}_i$ independently updates its stored best point $\boldsymbol{x}^{*\langle n \rangle} = \boldsymbol{x}^{(n)}$, along with the respective $z_i^{*\langle n \rangle} = z_i^{(n)}$.

B. Generation of candidate points and SM-based estimates

We iteratively generate discrete candidate points within $\mathcal{X}$ which, together with their SM-based estimates, will be used by the algorithm's exploitation and exploration routines to choose the next sampling point $\boldsymbol{x}^{(n+1)}$. For each new sample $\boldsymbol{x}^{(n)} \in \boldsymbol{X}_i^{\langle n \rangle}$, we generate candidate points along the coordinate directions up to the bounds of $\mathcal{X}$, and additional ones along the segments that connect $\boldsymbol{x}^{(n)}$ to any other sample $\boldsymbol{x}^{(m)} \in \boldsymbol{X}^{\langle n \rangle} \setminus \boldsymbol{x}^{(n)}$. Such new candidate points are then appended to the existing ones to build the candidate points set $\boldsymbol{E}^{\langle n \rangle}$. More details on candidate points generation are given in [11].

Remark 1: Because the sampled points $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}$ are known to all agents (see Assumption 4), a deterministic candidate points generation as proposed in [11], instead of a randomized one, results in synchronization of $\boldsymbol{E}^{\langle n \rangle}$ across $\boldsymbol{\mathcal{A}}$, without need for communications.

C. Coordination-based sampling point selection

Similar to our previous work in [11], the selection of $\boldsymbol{x}^{(n+1)}$ is based on a sequence of exploitation- and (if necessary,) exploration-related routines, described in the next subsections. Additionally, to address the multi-agent problem setup treated in this paper, we propose a new inter-agent coordination approach, which is used for both routines and discussed here.

[2]Without the assumption of anonymous packets, any agent can in fact infer the shape of another local objective f_j, differing only by a simple offset.

1) Selection of proposed candidate points: Agent $\mathcal{A}_i$ selects $N_c \geq 1$ candidate points from $\boldsymbol{E}^{\langle n \rangle}$, making a proposal set $\boldsymbol{\Theta}_i$ which is broadcasted to the other agents $\mathcal{A}_j$. The proposed candidate points are chosen according to a merit function $\delta_i(\cdot)$, defined differently for the exploitation and exploration modes, which will be detailed in Sections III-D and III-E. The first candidate point $\boldsymbol{\theta}_{i[1]}$ is selected as

$$\boldsymbol{\theta}_{i[1]} = \arg \max_{\boldsymbol{\theta} \in \boldsymbol{T}_{i[1]}^{\langle n \rangle}} \delta_i(\boldsymbol{\theta}). \tag{10}$$

where $\boldsymbol{T}_{i[1]} = \boldsymbol{E}^{\langle n \rangle} \cap \mathcal{S}$, with $\mathcal{S} \subseteq \mathcal{X}$ being a finite-volume, contiguous set (defined differently for each sampling mode). To choose the next proposal points $\boldsymbol{\theta}_{i[2]}, \ldots, \boldsymbol{\theta}_{i[N_c]}$, we define *exclusion balls* $\mathcal{B}(\boldsymbol{\theta}, r_\theta) \doteq \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{\theta}\| \leq r_\theta\}$, around the previously-chosen points ($r_\theta = \mu d(\mathcal{S})$ where $\mu > 0$ and $d(\cdot)$ is the diameter of a set), to avoid selecting point clusters for $\boldsymbol{\Theta}_i^{\langle n \rangle}$, which might have very similar merits. The following sets $\boldsymbol{T}_{i[k]}$ are then iteratively built as

$$\boldsymbol{T}_{i[k]} = \boldsymbol{T}_{i[k-1]} \setminus \bigcup_{k'=1}^{k-1} \mathcal{B}(\boldsymbol{\theta}_{i[k']}, r_\theta), \tag{11}$$

from which the next proposed points are selected as $\boldsymbol{\theta}_{i[k]} = \arg \max_{\boldsymbol{\theta} \in \boldsymbol{T}_{i[k]}} \delta_i(\boldsymbol{\theta})$, $k = 2, \ldots, N_c$. Finally, we build the proposal set

$$\boldsymbol{\Theta}_i = \left\{ \boldsymbol{\theta}_{i[k]}, k = 1, \ldots, N_c \right\}. \tag{12}$$

2) Coordination step: In this subroutine, $\mathcal{A}_i$ sends $\boldsymbol{\Theta}_i$ to all $\mathcal{A}_j$, and receives all $\boldsymbol{\Theta}_j$, resulting in all agents having a copy of the aggregate set $\boldsymbol{\Theta} = \cup_{i=1}^{N_{\mathcal{A}}} \boldsymbol{\Theta}_i$, with $|\boldsymbol{\Theta}| \leq N_c N_{\mathcal{A}}$. Afterwards, $\mathcal{A}_i$ tests each $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ for its *local* merit, and the results are collected in the augmented set

$$\mathring{\boldsymbol{\Theta}}_i \doteq \left\{ \left(\boldsymbol{\theta}, \delta_i(\boldsymbol{\theta})\right) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \right\}. \tag{13}$$

All agents exchange their respective $\mathring{\boldsymbol{\Theta}}_i$ and store their received copies in the set $\mathring{\boldsymbol{\Theta}} = \bigcup_{i=1}^{N_{\mathcal{A}}} \mathring{\boldsymbol{\Theta}}_i$. The aggregate (global) merit for each $\boldsymbol{\theta} \in \mathring{\boldsymbol{\Theta}}$ is then calculated,

$$\delta(\boldsymbol{\theta}) = \sum_{j=1}^{N_{\mathcal{A}}} \delta_j(\boldsymbol{\theta}). \tag{14}$$

Lastly, the candidate point for the (exploitation or exploration) sampling mode is selected as

$$\boldsymbol{x}_\theta = \arg \max_{\boldsymbol{\theta} \in \mathring{\boldsymbol{\Theta}}} \delta(\boldsymbol{\theta}). \tag{15}$$

Note that due to the aforementioned exchanges, the set $\mathring{\boldsymbol{\Theta}}$ and point $\boldsymbol{x}_\theta$ are synchronized among all agents in $\boldsymbol{\mathcal{A}}$.

D. Exploitation

For this mode, each $\mathcal{A}_i$ selects candidate points in a small region around the current best point $\boldsymbol{x}^{*\langle n \rangle}$, with the best estimated improvement from $z_i^{*\langle n \rangle}$. Accordingly, we use the coordination method in Section III-C with the following parameters:

- selection area $\mathcal{S}$ is set to the trust region $\mathcal{T}^{\langle n \rangle}$, which expands or contracts according to the *a posteriori* improvement of the current sample w.r.t. the existing best sample, see [11] for more details on this mechanism.
- the merit function is defined as

$$\delta_i(\boldsymbol{\theta}) \doteq z_i^{*\langle n \rangle} - \tilde{f}_i^{\langle n \rangle}(\boldsymbol{\theta}) + \beta \left[\lambda_i^{\langle n \rangle}(\boldsymbol{\theta}) - 2\tilde{\epsilon}_i^{\langle n \rangle} \right]_+, \tag{16}$$

  which is the estimated local improvement of $\boldsymbol{\theta}$ w.r.t. the *local* best value (see (5)), compensated from the estimated noise. This is based on the estimated f_i value at $\boldsymbol{\theta}$, a linear sum of the lower bound and central estimate (with the weighting factor $\beta = 0.1$ as in [11]).

After the coordination step (15) with cost as in (16), all agents have agreed on a common testing point $\boldsymbol{x}_\theta$. An expected improvement test eventually decides if the agents will proceed with evaluating $\boldsymbol{x}_\theta$, using the condition

$$\delta(\boldsymbol{x}_\theta) \geq \eta. \tag{17}$$

with $\delta(\cdot)$ being understood as the global estimated improvement (14) (given the local improvement merit (16)), and η being a (previously agreed or specified) constant threshold. If (17) is satisfied, the next experiment will be done with the agreed point $\boldsymbol{x}_\theta$, i.e., $\boldsymbol{x}^{(n+1)} = \boldsymbol{x}_\theta$. Otherwise, we select $\boldsymbol{x}^{(n+1)}$ via exploration instead, as discussed next.

Remark 2: The expected improvement test in MA-SMGO uses the global merit δ, in contrast to SMGO which uses the lower bound $\underline{f}^{\langle n \rangle}$, see e.g. [10]. This is to avoid another coordination step just to exchange $\underline{f}_i^{\langle n \rangle}$ among agents.

E. Exploration

The exploration routine proceeds in a similar consensus-based fashion using the coordination in Section III-C, but now using $\mathcal{S} = \mathcal{X}$, and the merit function

$$\delta_i(\boldsymbol{\theta}) \doteq \lambda_i^{\langle n \rangle}(\boldsymbol{\theta}) + \kappa(\tau^{\langle n \rangle}(\boldsymbol{\theta})). \tag{18}$$

The first term $\lambda_i^{\langle n \rangle}(\boldsymbol{\theta})$ is the SM-based uncertainty measure at $\boldsymbol{\theta}$ (see (9)). The second term, on the other hand, depends on the candidate point age $\tau^{\langle n \rangle}(\boldsymbol{\theta})$, calculated as the number of iterations elapsed from the creation of $\boldsymbol{\theta}$ (see Section III-B), up to the present iteration n. The operation $\kappa(\cdot)$ can be any strictly increasing function, and for most cases, it can simply be a small positive factor multiplied by $\tau^{\langle n \rangle}(\boldsymbol{\theta})$. This second term is added to ensure theoretical convergence, based on the arguments of [11]. In contrast to exploitation, where an expected improvement test is performed, the best candidate point selected from the exploration routine is directly assigned as the next sampling point, $\boldsymbol{x}^{(n+1)} = \boldsymbol{x}_\theta$.

IV. Theoretical Convergence

We prove here the theoretical convergence of the proposed MA-SMGO to the absolute minimum value z^* up to any finite precision ε, without the need to expose any individual agent evaluations. In this section, we use similar arguments as in [11]. The first lemma proves that MA-SMGO will eventually terminate any sequence of successive exploitation moves.

Lemma 1: Successive MA-SMGO exploitation samplings will terminate in finite iterations.

Proof: See the Appendix. ■

Lemma 2: MA-SMGO will generate a dense samples distribution, as $n \to \infty$.

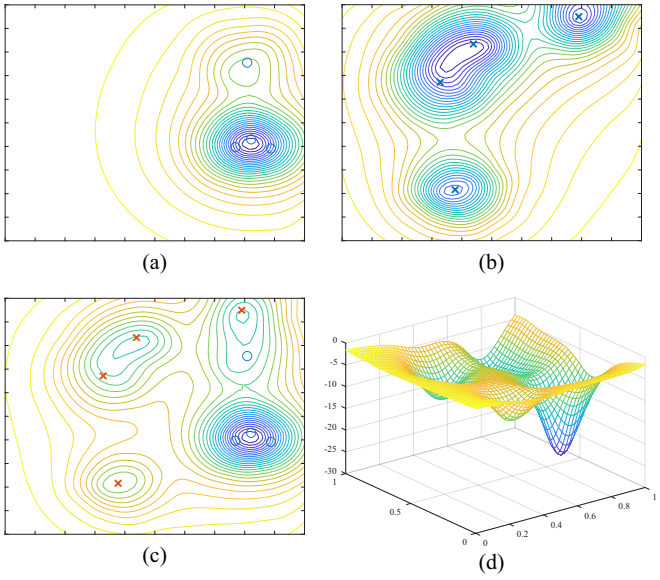

Fig. 3. Illustrative test: contour plots of the (private) agent functions (a) f_1 and (b) f_2, (c)-(d) the contour and surface plots of the overall function f.

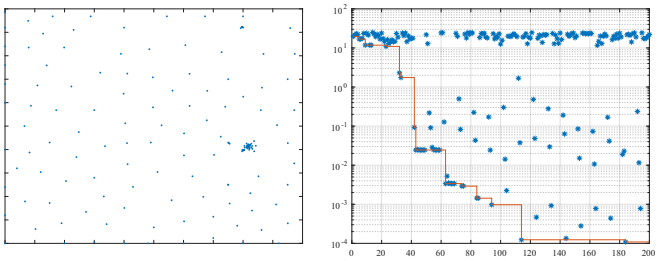

Fig. 4. Illustrative test: (a) spread of sampled points, (b) history of sampled values with iteration history of best (overall) sampled value

TABLE I
ILLUSTRATIVE TEST: OPTIMIZATION PARAMETERS

Description	Value
Dimensionality D	2
Number of agents $N_{\mathcal{A}}$	2
Number of iterations N	200
Search set $\mathcal{X}$	$[0,1]^2$
Exploitation improvement threshold η	0.05
Number of candidate point proposals N_c for each agent	4

Proof: Given the presence of a monotonically increasing age-based merit κ, and the fact that upon exchange of exploration-related merits, MA-SMGO agents will synchronize the exploration point, the proof of this lemma follows the same argument as that given in Lemma 2 of [11]. ■

Lemma 3: Given any finite desired precision $\varepsilon > 0$, there exists a finite number of iterations $n_\varepsilon < \infty$ such that

$$z^{*\langle n_\varepsilon \rangle} < z^* + \varepsilon.$$

Proof: The proof is given in Theorem 1 of [11], which can be consulted for further details. ■

V. ILLUSTRATIVE TEST

We demonstrate the effectiveness of the proposed approach, by testing it in an illustrative two-dimensional problem, and by putting in a comparative test on several benchmark problems of different characteristics.

In this example, we use MA-SMGO to solve a simple problem of minimizing a black-box f with $D = 2$ and $N_{\mathcal{A}} = 2$, such that each (private) f_i is a four-peak function

$$f_i(\boldsymbol{x}) = -\sum_{j=1}^{4} \psi(\boldsymbol{x}, \boldsymbol{c}_{i[j]}) \tag{19}$$

where an individual peak is defined as

$$\psi(\boldsymbol{x}, \boldsymbol{c}_{i[j]}) \doteq \frac{10}{1 + 100(\boldsymbol{x} - \boldsymbol{c}_{i[j]})^2}. \tag{20}$$

Each peak center $\boldsymbol{c}_{i[j]}$ has been randomly generated within the search set. Fig. 3 shows the contour plots of the respective agent functions, and of the global (sum) one. Furthermore, it shows a 3D surface of the example f, showing several local minima, and a deeper global one. More detailed specifications for this illustrative test are summarized in Table I.

Fig. 4 shows the distribution of sampled points throughout the search set. The spread of the samples shows an increased concentration around a local and the global minima, implying exploitation around these regions. Fig. 4 also shows the history of the optimality gap (defined as the difference between the current best sampled value $z^{*\langle n \rangle}$ and the optimum value $f(\boldsymbol{x}^*)$) in log scale. We observe that the samples started to exploit in the region around $\boldsymbol{x}^*$ from $n = 40$ to $n = 60$, reflected by the optimality gap going less than 0.1 in these iterations. After around $n = 70$, MA-SMGO agents tended to explore around $\mathcal{X}$, shown by the higher optimality gaps of the sampled values. However, the best global value has continued to improve, pushing down the optimality gap close to 1×10^{-4} at the end of the example run.

VI. BENCHMARK COMPARISON TESTS

We evaluated MA-SMGO in different benchmark tests, comparing its iteration-based optimization performance with that of the centralized ("classical") SMGO.

A. Benchmark problems

We consider 5 benchmark optimization problems summarized in Table V based on the fundamental peaks function (20), where each peak center $\boldsymbol{c}_{i[j]}$ has been randomly generated within $\mathcal{X}$. We test the proposed MA-SMGO in different problem setups:

1) $f(\boldsymbol{x})$ is composed of decoupled functions, i.e., $f(\boldsymbol{x}) = \sum_{i=1}^{N_{\mathcal{A}}} f_i(\boldsymbol{x}_i)$. This means that each function f_i is only dependent on $\boldsymbol{x}_i$, and agent $\mathcal{A}_i$ can, in theory, independently optimize it. For example, for P1, we have the general form

$$f(\boldsymbol{x}) = f_1(\boldsymbol{x}^{[1:2]}) + f_2(\boldsymbol{x}^{[3:4]}) + f_3(\boldsymbol{x}^{[5:6]}),$$

where $\boldsymbol{x}^{[p:q]}$ is the subvariable of $\boldsymbol{x}$ composed of the p-th up to its q-th elements. Problem P3 takes the same form, but with $D = 8$.

TABLE II

OVERVIEW OF THE BENCHMARK PROBLEMS CONSIDERED IN THE STATISTICAL TESTS

Problem	D	$N_{\mathcal{A}}$	f_1	f_2	f_3	f_4	f_5
P1	6	3	$\sum_{i=1}^{6}\psi(\boldsymbol{x}^{[1:2]},\boldsymbol{c}_{1i})$	$\sum_{i=1}^{6}\psi(\boldsymbol{x}^{[3:4]},\boldsymbol{c}_{2i})$	$\sum_{i=1}^{6}\psi(\boldsymbol{x}^{[5:6]},\boldsymbol{c}_{3i})$	–	–
P2			$\sum_{i=1}^{12}\psi(\boldsymbol{x},\boldsymbol{c}_{1i})$	$\sum_{i=1}^{12}\psi(\boldsymbol{x},\boldsymbol{c}_{2i})$	$\sum_{i=1}^{12}\psi(\boldsymbol{x},\boldsymbol{c}_{3i})$	–	–
P3	8	4	$\sum_{i=1}^{8}\psi(\boldsymbol{x}^{[1:2]},\boldsymbol{c}_{1i})$	$\sum_{i=1}^{8}\psi(\boldsymbol{x}^{[3:4]},\boldsymbol{c}_{2i})$	$\sum_{i=1}^{8}\psi(\boldsymbol{x}^{[5:6]},\boldsymbol{c}_{3i})$	$\sum_{i=1}^{8}\psi(\boldsymbol{x}^{[7:8]},\boldsymbol{c}_{4i})$	–
P4			$\sum_{i=1}^{16}\psi(\boldsymbol{x},\boldsymbol{c}_{1i})$	$\sum_{i=1}^{16}\psi(\boldsymbol{x},\boldsymbol{c}_{2i})$	$\sum_{i=1}^{16}\psi(\boldsymbol{x},\boldsymbol{c}_{3i})$	$\sum_{i=1}^{16}\psi(\boldsymbol{x},\boldsymbol{c}_{4i})$	–
P5		5	$\sum_{i=1}^{8}\psi(\boldsymbol{x}^{[1:2]},\boldsymbol{c}_{1i})$	$\sum_{i=1}^{8}\psi(\boldsymbol{x}^{[3:4]},\boldsymbol{c}_{2i})$	$\sum_{i=1}^{16}\psi(\boldsymbol{x},\boldsymbol{c}_{3i})$	$\sum_{i=1}^{12}\psi(\boldsymbol{x}^{[1:4]},\boldsymbol{c}_{4i})$	$\sum_{i=1}^{12}\psi(\boldsymbol{x}^{[5:8]},\boldsymbol{c}_{5i})$

TABLE III

STATISTICAL BENCHMARK TEST: OPTIMIZATION PARAMETERS

Description	**Value**
Number of iterations N	$100D$
Search set $\mathcal{X}$	$[0,1]^D$
Exclusion radius constant μ	0.1
Exploitation improvement threshold η	0.05
Number of candidate point proposals N_c for each agent	$2D$

2) The global function f is composed of globally-coupled component functions. This setup is considered in Problems P2 and P4.
3) f has decoupled component functions, and the rest of the component functions are globally coupled. Problem P5 has this setup, with $\mathcal{A}_1$ and $\mathcal{A}_2$ having decoupled functions, and $\mathcal{A}_4$ and $\mathcal{A}_5$ having a larger set of decision variables. Lastly, $\mathcal{A}_3$ has to optimize its own f_3 with the entire $\boldsymbol{x}$.

B. Comparative test parameters

We compare three different optimization approaches:

1) **Random:** a baseline naïve approach, in which all agents will simply decide on a random global test variable $\boldsymbol{x}^{(n)}$ at each iteration.
2) **MA-SMGO:** our proposed technique that involves consensus-based data exchange among agents, while keeping their own local functions private. The optimization parameters are given in Table III.
3) **SMGO:** our previous work from [11] will serve as a centralized approach. This is with the assumption that all the agent functions are instead *public*, and the global f can be sampled/modelled by a centralized entity. For SMGO, we set the expected improvement threshold to a constant $\eta = 0.05$, instead of adapting with the global Lipschitz constant estimate (as in [11]). This way, we keep the difference between SMGO and MA-SMGO only in the aspect of being centralized/decentralized.

We do not have a competitor decentralized method for comparison, because existing ones do not fit the problem description and assumptions described in Section II. In running the above three methods, we resort to a statistical test comprising of 25 independent trials, each of which is run with N iterations depending on the problem dimensionality (see Tables V and III). Each run starts with randomly generated sampling points, and all tested methods (Random, MA-SMGO, SMGO) have the same set of initial points for all of their respective trials, to maintain fairness.

C. Statistical results

Table IV summarizes the best found global objective value at the end of the 25 independent runs, with the respective means and standard deviations for each problem. It was unsurprising that random sampling resulted in the worst mean values across all the problems considered. Furthermore, the centralized SMGO has resulted in the best mean result in all problems except P2 (with comparable standard deviations), which is also expected because we removed the privacy assumption of all agents. Access to all agent functions enables SMGO to model the global function f, instead of each agent making a model of their respective f_i. Furthermore, SMGO has access to all the candidate points for calculating the merits, instead of only the proposed ones exchanged among agents.

Even with the limitations brought by agent function privacy, MA-SMGO resulted in a highly competitive result with SMGO, with even a best result with P2, a six-dimensional function with non-separable objectives. The difference of MA-SMGO and SMGO mean results for P1 and P2 (both with $D = 6$) is within less than 1, while for P3-P5 ($D = 8$) it is around 3. In fact, for all the tested problems except P4, the MA-SMGO and SMGO average results were within each other's standard deviations, and far from that of Random. These results demonstrate the effectiveness of the proposed MA-SMGO consensus-based subroutines to address the need to optimize a global objective, without having to expose the individual agents' respective evaluation results.

TABLE IV

BENCHMARK COMPARISON: MEANS AND STANDARD DEVIATIONS OF RESULTING BEST VALUES ACROSS 25 INDEPENDENT TRIALS.

Problem	**Random**	**MA-SMGO**	**SMGO**
P1	-36.58 ($\pm$2.53)	-42.67 ($\pm$4.39)	-43.54 ($\pm$2.03)
P2	-9.68 ($\pm$1.07)	-13.12 ($\pm$2.83)	-12.98 ($\pm$3.67)
P3	-58.04 ($\pm$4.28)	-65.27 ($\pm$4.75)	-68.22 ($\pm$7.74)
P4	-10.14 ($\pm$0.25)	-11.32 ($\pm$0.63)	-14.80 ($\pm$1.05)
P5	-39.09 ($\pm$1.90)	-44.56 ($\pm$2.63)	-47.41 ($\pm$3.22)

VII. CONCLUSION

This paper proposes a new optimization method to address problem setups where a global objective is expressed as a sum of agent private functions, but for which the global decision variable affects all the private functions. The discussed method introduces consensus-based subroutines to the Set

Membership Global Optimization (SMGO) method, enabling agents to coordinate with each other the next global sampling point, without having to share their respective private data sets and models. The resulting method, referred to as MA-SMGO, is demonstrated in an illustrative test and statistically compared with the centralized SMGO, showing its competitive iteration-based optimization results, even with the limitations brought by multi-agent setup and agent privacy.

Current ongoing work deals with the case where agent functions are not fully coupled with the entire decision variable but only its subset. Furthermore, investigations are underway on two aspects: learning the couplings (or decouplings) of the agent private functions with respect to the decision subvariable/s, to further minimize the required inter-agent communication; and techniques to make the technique scalable w.r.t. dimensionality and number of participating agents.

APPENDIX

Proof of Lemma 1 Consider that all $\mathcal{A}_i \in \boldsymbol{\mathcal{A}}$ have the same set of candidate points $\boldsymbol{E}^{\langle n \rangle}$ (see Remark 1). Furthermore, taking the expected improvement condition (17) and expanding $\delta(\boldsymbol{x}) = \sum_i \delta_i(\boldsymbol{x})$ in the context of exploitation (see Sections III-C and III-D), we use (16) and have

$$\sum_i z_i^{*\langle n \rangle} - \sum_i \tilde{f}_i^{\langle n \rangle}(\boldsymbol{x}) + \beta \sum_i \left[\lambda_i^{\langle n \rangle}(\boldsymbol{x}) - 2\tilde{\epsilon}_i^{\langle n \rangle} \right]_+ \geq \eta. \quad (21)$$

Exploitation samplings are still possible as long as there are points which can pass (17), that is

$$\mathcal{P}^{\langle n \rangle} \doteq \left\{ \boldsymbol{x} \in \mathcal{T}^{\langle n \rangle} : (21) \text{ applies} \right\} \neq \varnothing.$$

Now, for the sake of contradiction, assume that exploitation samplings are performed in infinite succession.

Because $\sum_i z_i^{*\langle n \rangle} = z^{*\langle n \rangle}$ and the (global) sum of the central estimates will not be lower than the lowest sampled value (see for instance Fig. 2), the sum of the first two terms in (21) is non-positive. Hence, (17) is always violated if $\beta \sum_i \left[\lambda_i^{\langle n \rangle}(\boldsymbol{x}) - 2\tilde{\epsilon}_i^{\langle n \rangle} \right]_+ < \eta$. Since $\lambda_i^{\langle n \rangle}(\boldsymbol{x}) \leq 2\tilde{\epsilon}_i^{\langle n \rangle} + 2\tilde{\gamma}_i^{\langle n \rangle} \|\boldsymbol{x} - \boldsymbol{m}\|$ with $\boldsymbol{m} \in \boldsymbol{X}^{\langle n \rangle}$ being the closest sample to $\boldsymbol{x}$, we have

$$\overbrace{\beta \sum_i \left[\lambda_i^{\langle n \rangle}(\boldsymbol{x}) - 2\tilde{\epsilon}_i^{\langle n \rangle} \right]_+}^{a} \leq \overbrace{2\beta \sum_i \tilde{\gamma}_i^{\langle n \rangle} \|\boldsymbol{x} - \boldsymbol{m}\|}^{b} < \eta.$$

In the above inequality, if $b < \eta$, then the improvement test violation $a < \eta$ holds. Now we proceed with $b < \eta$, and rearranging terms, we arrive at

$$2\beta \|\boldsymbol{x} - \boldsymbol{m}\| \sum_i \tilde{\gamma}_i^{\langle n \rangle} < \eta. \quad (22)$$

The above inequality means that for every sample $\boldsymbol{x}^{(i)}$ (not just from exploitation but also exploration), there is a finite-radius ball $\mathcal{B}\left(\boldsymbol{x}^{(i)}, \frac{\eta}{2\beta \sum_i \tilde{\gamma}_i^{\langle n \rangle}} \right)$ that is removed from $\mathcal{P}^{\langle n \rangle}$, given that $\forall n,\ \tilde{\gamma}_i^{\langle n \rangle} \in [\underline{\gamma}_i, \gamma_i]$. Therefore, $\exists n' < \infty : \mathcal{P}^{\langle n' \rangle} = \varnothing$, forcing MA-SMGO to perform exploration sampling. Hence, the contradictory assumption is invalidated, and the lemma is proven. ■

REFERENCES

[1] A. Nedić and J. Liu, "Distributed Optimization for Control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 77–103, 2018.

[2] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. H. Johansson, "A survey of distributed optimization," *Annual Reviews in Control*, no. xxxx, 2019.

[3] Y. Yang, X. Guan, Q.-S. Jia, L. Yu, B. Xu, and C. J. Spanos, "A Survey of ADMM Variants for Distributed Optimization: Problems, Algorithms and Features," pp. 1–37, 2022.

[4] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A Survey of Distributed Optimization and Control Algorithms for Electric Power Systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

[5] J. Fransman, J. Sijs, H. Dol, E. Theunissen, and B. De Schutter, "Distributed bayesian: A continuous distributed constraint optimization problem solver," *J. Artif. Int. Res.*, vol. 76, may 2023. [Online]. Available: https://doi.org/10.1613/jair.1.14151

[6] M. Choudhury, A. Sarker, S. Yaser, M. M. A. A. Khan, W. Yeoh, and M. M. Khan, "A particle swarm inspired approach for continuous distributed constraint optimization problems," *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106280, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0952197623004645

[7] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowledge-Based Systems*, vol. 216, p. 106775, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705121000381

[8] P. Singh, M. K. Singh, R. Singh, and N. Singh, *Federated Learning: Challenges, Methods, and Future Directions*. Cham: Springer International Publishing, 2022, pp. 199–214. [Online]. Available: https://doi.org/10.1007/978-3-030-85559-8_13

[9] T. Grinshpoun, "When you say (dcop) privacy, what do you mean?: Categorization of dcop privacy and insights on internal constraint privacy," in *ICAART 2012 - Proceedings of the 4th International Conference on Agents and Artificial Intelligence*, ser. ICAART 2012 - Proceedings of the 4th International Conference on Agents and Artificial Intelligence, 2012, pp. 380–386, 4th International Conference on Agents and Artificial Intelligence, ICAART 2012 ; Conference date: 06-02-2012 Through 08-02-2012.

[10] L. Sabug, F. Ruiz, and L. Fagiano, "SMGO: A set membership approach to data-driven global optimization," *Automatica*, vol. 133, p. 109890, 2021.

[11] L. Sabug, F. Ruiz, and L. Fagiano, "SMGO-Δ: Balancing Caution and Reward in Global Optimization with Black-Box Constraints," *Information Sciences*, vol. 605, pp. 15–42, 2022.

[12] M. Milanese and C. Novara, "Set Membership identification of nonlinear systems," *Automatica*, vol. 40, no. 6, pp. 957–975, 2004.

[13] J. P. Calliess, S. J. Roberts, C. E. Rasmussen, and J. Maciejowski, "Lazily Adapted Constant Kinky Inference for nonparametric regression and model-reference adaptive control," *Automatica*, vol. 122, p. 109216, 2020.

[14] L. Fagiano and C. Novara, "Learning a Nonlinear Controller From Data: Theory, Computation, and Experimental Results," *IEEE Transactions on Automatic Control*, vol. 61, no. 7, pp. 1854–1868, jul 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7271025/