2017 American Control Conference
Sheraton Seattle Hotel
May 24–26, 2017, Seattle, USA

Discrete-time Fractional-order Multiple Scenario-based Sensor Selection

Sérgio Pequito ⋆ Andrew Clark ⋄ George J. Pappas ⋆

***Abstract*— In this paper, we address the problem of deploying sensors to estimate the state of a plant described by discrete-time fractional-order system. More specifically, we assume that these systems' parameters and disturbance/measurement noise characteristics describe possible scenarios. Therefore, the goal of this paper is that of selecting a subset of sensors that will optimally perform (in a minimum squared error sense) among multiple (finite) scenarios. In particular, we show this problem to be NP-hard, and we provide a bisection-type algorithm with suboptimality guarantees. Furthermore, we show that no other algorithm ensures better optimality bound for this problem unless P=NP. Finally, we present some simulations that illustrate the applicability of the main results in an electroencephalogram data associated with different tasks.**

I. INTRODUCTION

Applications ranging from process control, automotive industry, power systems, health, aircraft and traffic control are often modeled by parametrizable dynamical systems [1], [2]. The existence of such models is crucial to assess the state of the system from only few measurements obtained from the system's output, i.e., its observability. In fact, from a design point of view, to ensure that the state of the system is recovered it is important to decide which collection of state variables should be measured. The problem of *sensor placement* aims to determine a subset of state variables that need to be measured, while ensuring that a given performance metric is achieved. For instance, in [3] the metric considered relates with the mean square error associated with linear time-invariant system under any disturbance and measurements (with positive definite second moments).

In practical scenarios, the parameters describing the dynamical systems are often not accurately known, or the disturbance and measurement noises might follow distributions with different characterizations. Therefore, one needs to extend the previous research to the case where a designer faces a finite collection of possible *scenarios* that describe a specific dynamics and the characteristics of the disturbance/measurement noises. More specifically, we aim to explore the sensor placement under the case where a finite collection of scenarios is possible, and we want to maximize the performance across the different scenarios. In particular, in what follows we explore the scenarios where the dynamic systems are described by discrete-time fractional-order systems [5], whose class contain the linear time-invariant systems as a particular case. In addition, these models are motivated by their ability to characterize the behavior of neurological phenomena, such as electroencephalogram (EEG) [6]. Thus, one can envision an application where different EEG recordings are associated with a finite number of tasks (to be captured by different scenarios), and one wants to determine a set of sensors with provable performance across different tasks – which we explore in further detail in the illustrative example. Furthermore, this enables a principled approach to develop a new generation of brain-machine interfaces, since it will enable a more accurate state estimation and, hence, more efficient and reliable actuation strategies.

In [7], the sensor placement for multi-scenario is addressed when the plant is disregarded, and the sensors are assumed as random variables. More specifically, the goal is to determine a subset of these (with pre-specified cardinality) that maximizes the mutual information of the sensors with an underlying random process. An alternative to the proposed problem is that of determining the minimum number of sensors required for provable observability in the case where the parameters are assumed unknown and can be arbitrary. This problem was explored in [8], where the authors determine the minimum number of sensors to ensure a discrete-time fractional-order dynamics to satisfy structural observability, which ensures observability for almost all numerical realizations. Finally, the present work extends that in [9], where the the selection of sensors considers a single scenario. In particular, in [9], the authors propose an objective function that captures the performance in terms of the minimum square error, and show this to be submodular. Nonetheless, as we show (see Section III-B), under the assumption that we have different scenarios, the objective is no longer submodular, which motivates the introduction of novel set of strategies to address the problem, as we propose in this paper.

The main contributions of this paper are as follows: (*i*) we propose a robust sensor placement for discrete-time fractional-order systems; (*ii*) we show the problem to be NP-hard; (*iii*) we provide a bisection algorithm with optimality guarantees; (*iv*) we show that no other algorithm can outperform the one proposed unless P=NP; and (*v*) we illustrate the main results in the context of electroencephalogram data associated with different tasks.

The remainder of this paper is organized as follows. In Section II, we provide our setup and problem formulation. In Section III, we present our main results. Finally, in Section IV, we illustrate how our main results can be applied in the context of real EEG data collected during the performance of different tasks.

This work was supported in part by the TerraSwarm Research Center, one of six centers supported by the STARnet phase of the Focus Center Research Program (FCRP) a Semiconductor Research Corporation program sponsored by MARCO and DARPA, and by NSF grant CNS-1544173.

⋆ Department of Electrical and Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania

⋄ Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, MA, 01609, USA

II. PROBLEM STATEMENT

Consider N scenarios described by discrete-time fractional-order systems as follows:

$$\begin{aligned} \Delta_i x_{k+1} &= A_i x_k + w_k^i, \quad i=1,\dots,N, \\ y_k^i &= \mathbb{I}(\mathcal{J}) x_k + v_k^i, \quad k=0,1,\dots, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ ($n \in \mathbb{N}$) is the state vector, x_0 the initial condition, and $y_k \in \mathbb{R}^n$ is the measured output vector, where $\mathbb{I}(\mathcal{J})$ is the $n \times n$ diagonal matrix whose entry j is one if $j \in \mathcal{J}$ and zero otherwise. Notice that, in particular, a sensor j is used (or selected) if $j \in \mathcal{J}$. Furthermore, for each scenario i, we have the following: A_i describes the coupling dynamics capturing the spatial dependency; w_k^i the process noise; v_k^i the measurement noise; $\Delta_i \equiv \text{diag}\,(\Delta_i^{\alpha_1}, \Delta_i^{\alpha_2}, \dots, \Delta_i^{\alpha_n})$, is the diagonal matrix operator with elements the $\{\Delta_i^{\alpha_1}\}_{i=1}^n$, where $\Delta_i^{\alpha_j}$ is the discrete fractional-order difference operator given by

$$\Delta^{\alpha_j} \equiv \sum_{m=0}^{k+1} (-1)^j \binom{\alpha_j}{m} x_{k-m+1},$$

and $\binom{\alpha_j}{m} = \frac{\Gamma(\alpha_j+1)}{\Gamma(m+1)\Gamma(\alpha_j-m+1)}$, where $\alpha_j > 0$ is the fractional-order exponent, and $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the Gamma function. In particular, we notice that α_j captures the temporal dependency of the process associated with x_j.

In fact, for each scenario the system (1) can be re-written in a closed-form as follows [8]:

$$x_k = G_k^i x_0 + \sum_{j=0}^{k-1} G_{k-1-j}^i w_j, \quad k \geq 1, \tag{2}$$

where

$$G_k^i \equiv \begin{cases} I, & k=0, \\ \sum_{j=0}^{k-1} A_j^i G_{k-1-j}^i, & k \geq 1, \end{cases}$$

where $A_0^i = A^i$, and A_j^i is a diagonal matrix whose l-th entry is $(-1)^j \binom{\alpha_l}{j+1}$.

Now, consider a vector of measurements $\bar{y}_k \equiv (y_0^\intercal, y_1^\intercal, \dots, y_k^\intercal)^\intercal$, the vector of process noises $\bar{w}_k^i \equiv ((w_0^i)^\intercal, (w_1^i)^\intercal, \dots, (w_k^i)^\intercal)^\intercal$ and the vector of measurement noises $\bar{v}_k^i \equiv ((v_0^i)^\intercal, (v_1^i)^\intercal, \dots, (v_k^i)^\intercal)^\intercal$. Whereas the vector $\bar{y}_k$ is known, the vectors $\bar{w}_k^i$ and $\bar{v}_k^i$ are not, as well as the initial condition x_0. Nonetheless, we can consider the following (usual) mild assumptions: the initial condition be unknown and modelled by a random variable whose expected value is x_0 and its covariance is $\mathbb{C}(x_0) \succ 0$; the disturbance w_k^i and noise v_k^i to be described by zero-mean random variables, whose covariance is described respectively by $\mathbb{C}(w_k^i) \succ 0$ and $\mathbb{C}(v_k^i) = \sigma_i^2 I$, with $\sigma_i > 0$, for all $k \geq 0$. Further, for all $k, k' \geq 0$ such that $k \neq k'$, let the x_0, w_k and v_k, as well as, the w_k and $w_{k'}$, and v_k and $v_{k'}$ to be uncorrelated. Therefore, under these assumptions, the best linear estimator $\hat{x}_{k'}$ for $x_{k'}$, with k' belonging to a *observation interval* $[0,k] = \{0,1,\dots,k\}$, is given as follows:

$$\hat{x}_{k'} = L_{k'} \hat{z}_{k-1}, \tag{3}$$

where

$$\begin{aligned} \hat{z}_{k-1} = \mathbb{E}(z_{k-1}) + \mathbb{C}(z_{k-1})\mathcal{O}_k^\intercal(\mathcal{J}) \left[\mathcal{O}_k(\mathcal{J})\mathbb{C}(z_{k-1})\mathcal{O}_k^\intercal(\mathcal{J}) \right. \\ \left. +\sigma^2 I\right]^{-1} (\bar{y}_k - \mathcal{O}_k(\mathcal{J})\mathbb{E}(z_{k-1}) - \mathbb{E}(\bar{v}_k)) \end{aligned} \tag{4}$$

is the minimum square estimate of $z_{k-1} = (x_0^\intercal, (\bar{w}_{k-1}^i)^\intercal)^\intercal$, $\mathbb{E}(x)$ is its expected value of x, and $\mathbb{C}(x) \equiv \mathbb{E}([x-\mathbb{E}(x)][x-\mathbb{E}(x)]^\intercal)$ its covariance; and $\mathcal{O}_k(\mathcal{J}) = [L_0^\intercal \mathbb{I}(\mathcal{J}), L_1^\intercal \mathbb{I}(\mathcal{J}), \dots, L_k^\intercal \mathbb{I}(\mathcal{J})]^\intercal$, $L_j = [G_j, G_{j-1}, \dots, G_0, \mathbf{0}]$ for $j \geq 0$, and $\mathbf{0}$ is a zero matrix with appropriate dimensions. Furthermore, the minimum mean square error of our estimate $\hat{x}_{k'}$ under the scenario i is given by

$$\text{mmse}(x_{k'}) \equiv \text{tr}\left(L_{k'} \Sigma_{z_{k-1}}^i(\mathcal{J}) L_{k'}^\intercal\right), \tag{5}$$

where $\Sigma_{z_{k-1}}^i(\mathcal{J})$ is the error covariance of $\hat{z}_{k-1}$ when a subcollection $\mathcal{J} \subset [n] = \{1,\dots,n\}$ sensors is considered, and it is described by

$$\begin{aligned} \Sigma_{z_{k-1}}^i(\mathcal{J}) &\equiv \mathbb{E}\left((z_{k-1}-\hat{z}_{k-1})(z_{k-1}-\hat{z}_{k-1})^\intercal\right) \\ &= \mathbb{C}(z_{k-1}) - \mathbb{C}(z_{k-1})\mathcal{O}_k^\intercal(\mathcal{J}) \\ &\quad \left(\mathcal{O}_k(\mathcal{J})\mathbb{C}(z_{k-1})\mathcal{O}_k^\intercal(\mathcal{J}) + \sigma_i^2 I\right)^{-1} \mathcal{O}_k(\mathcal{J})\mathbb{C}(z_{k-1}), \end{aligned} \tag{6}$$

where the covariances implicitly depend on the error characteristics associated with the ith scenario. Subsequently, we can consider as an estimation error metric the η-confidence ellipsoid of $z_{k-1} - \hat{z}_{k-1}$ [10]. More specifically, the minimum volume ellipsoid that contains $z_{k-1} - \hat{z}_{k-1}$ with probability η when a subset $\mathcal{J}$ of sensors is selected, that can be described by $\mathcal{E}_\epsilon^i \equiv \{z : z^\intercal \Sigma_{z_{k-1}}^i(\mathcal{J}) z \leq \epsilon\}$, where $\epsilon \equiv F_{\chi^2_{n(k+1)}}^{-1}(\eta)$ and $F_{\chi^2_{n(k+1)}}$ is the cumulative distribution function of a χ-squared random variable with $n(k+1)$ degrees of freedom [11]. Thus, the volume of $\mathcal{E}_\epsilon^i$,

$$\text{vol}(\mathcal{E}_\epsilon^i) \equiv \frac{(\epsilon\pi)^{n(k+1)/2}}{\Gamma\left(n(k+1)/2+1\right)} \det\left((\Sigma_{z_{k-1}}^i)^{1/2}(\mathcal{J})\right), \tag{7}$$

where the Gamma function quantifies the estimation's error of $\hat{z}_{k-1}$, and as a result, for any $k' \in [0,k]$, of $\hat{x}_{k'}$ as well, since per (3) the optimal estimator for z_{k-1} defines the optimal estimator for $x_{k'}$ [11]. Henceforth, if we consider the logarithm of (7), then we obtain

$$\log \text{vol}(\mathcal{E}_\epsilon^i) = \beta + 1/2 \log\det\left(\Sigma_{z_{k-1}}^i(\mathcal{J})\right); \tag{8}$$

where β is a constant that depends only on $n(k+1)$ and ϵ, in accordance to (7), and, as a result, we refer to the $\log\det\left(\Sigma_{z_{k-1}}^i(\mathcal{J})\right)$ as the $\log\det$ *estimation error of the Kalman-like filter* under scenario i [9]. This function is *supermodular* and *non-increasing* (formally defined in Section III) that enable the use of algorithms to attain an approximate solution with some optimality guarantees [9].

Therefore, in order to extend the previous formations to the multi-scenario setup, we address the following problem.

A. Scenario-based Sensor Placement

Given N scenarios described in (1), our aim is to determine the r sensors that solve the following problem

$$\begin{array}{lll} \text{minimize} & \underset{i\in\{1,\dots,N\}}{\text{maximize}} & \log\det\left(\Sigma^i_{z_{k-1}}(\mathcal{J})\right) \\ & \text{subject to} & |\mathcal{J}|\leq r, \quad \mathcal{J}\subseteq\{1,\dots,n\}. \end{array} \qquad (\mathcal{P}_1)$$

Nonetheless, the minimum-maximum operations across different submodular objectives is not necessarily submodular [7]. In fact, in the next section, we show that for the present setup this is also the case. Therefore, new mechanisms are required to address the problem are required, which are the basis of the main results of our paper.

III. MAIN RESULTS

We present the main results of the present paper. First, in Section III-A, we show that the proposed problem is NP-hard (see Theorem 1). Then, in Section III-B, we provide an example where common strategies to deal with similar problems will fail in the general setup proposed in this paper. Next, in Section III-C, we provide our approach, whose procedure is summarized in Algorithm 1. Furthermore, in Section **??**, we show that such approach possesses optimality guarantees (Theorem 3), and these cannot be improved by any other polynomial algorithm (Theorem 4).

A. Computational Complexity

We start by showing that the problem at hand is computationally challenge. In fact, it can be shown that there is *no* polynomial-time algorithm providing any approximation guarantee unless P=NP.

Theorem 1: Define $f_i(\mathcal{J}) \triangleq \log\det(\Sigma^i_{z_{k-1}}(\mathcal{J})$ and $f(\mathcal{J}) \triangleq \max\{f_i(\mathcal{J}) : i=1,\dots,N\}$. For any function $\gamma : \mathbb{Z}_+ \to \mathbb{Z}_+$, if there is a polynomial-time algorithm that returns a set S satisfying $f(\mathcal{J}) \geq \gamma(n)f(\mathcal{J}^*)$, where $\mathcal{J}^*$ is the optimal solution to $\mathcal{P}_1$, then P=NP.

Proof: As a preliminary, define the *hitting set problem* as follows. Let U denote a finite set, and let $T_1,\dots,T_m$ denote a collection of subsets of U. The hitting set problem is to select a minimum-size set S such that $S\cap T_i \neq \emptyset$ for all $i=1,\dots,m$. The hitting set problem is known to be NP-hard [12]. The approach is to show that, if a polynomial time algorithm that returns a set S satisfying $f(\mathcal{J}) \geq \gamma(n)f(\mathcal{J}^*)$ exists, then there is a polynomial-time algorithm for the hitting set problem. Let U and $T_1,\dots,T_m$ define an instance of the hitting set problem. An instance of the worst-case sensor placement problem can be constructed from U and $T_1,\dots,T_m$ as follows. Choose $\alpha=1$. Let $n=|U|$ and let $V=\{1,\dots,n\}$ denote the set of sensor locations. Define $\mathbb{C}(x_0)=I$, $\mathbb{C}(w_k)=I$, and define $A^{(l)}$ for $l=1,\dots,m$ by

$$(A^{(l)})_{ij} = \begin{cases} 1, & i\in T_l, j\neq i \\ (n+1), & i\in T_l, j=i \\ 0, \text{else} \end{cases} \qquad (9)$$

Let $\mathcal{J}$ be the output of the polynomial-time algorithm. Now, if $\mathcal{J}$ is not a solution to the hitting set problem, then $O_{k,S}=0$ for some k, and hence $f_l(\mathcal{J})=0$. Conversely, if $\mathcal{J}$ is a hitting set, then $O^T_{k,S}O_{k,S}$ is nonzero and positive semidefinite, and hence $f_l(S)>0$ for all $l=1,\dots,m$. Since by assumption $f(\mathcal{J}) \geq \gamma(n)f(\mathcal{J}^*)$ for some $\gamma(n)>0$, we must have $f(\mathcal{J})>0$, implying that $\mathcal{J}$ is a solution to the hitting set problem. Hence if there exists a hitting set of cardinality r, then that set will be returned by the algorithm.

We can then obtain a minimum-size hitting set in polynomial time by iterating over all $r=1,\dots,n$ and finding the minimum r where the solution to $\mathcal{P}_1$ returned by the algorithm satisfies $f(\mathcal{J})>0$. This set $\mathcal{J}$ is a minimum-size hitting set. This implies existence of a polynomial-time algorithm for the NP-hard hitting set problem, and thus P=NP. ■

As a consequence, it is recurrent to resort to greedy algorithms that will provide an approximate solution to $\mathcal{P}_1$. Among the possible strategies is that of showing that the objective is submodular, which leads to well known greedy algorithms that ensure some optimality guarantees. Nonetheless, in the next subsection, we show that our problem does not possess such properties, and, in fact, the greedy algorithms to obtain an approximate solution can perform arbitrarily poorly upon the initial setup of the problem, i.e., the optimality gap can be made arbitrarily large.

B. Sub-Optimality Gap

In this subsection, we show that the proposed problem cannot be directly address by the submodular optimization schemes previously proposed in the literature. More specifically, we provide a simple counter-example (with small state dimension) that presents a decay in performance, i.e., achieved value by the objective set function in $\mathcal{P}_1$, for a variation of the parameters in the different scenarios.

Recall, that a particular instance of discrete-time fractional-order systems is that of discrete linear time-invariant systems. Therefore, let us consider two scenarios described by linear-time invariant systems (where the fractional-order derivative is the usual derivative) as follows

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 10 \\ 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \end{pmatrix},$$

and

$$A_2 = \begin{pmatrix} 0.1136 & 1.4790 & -0.2339 & 8.5306 \\ 9.0953 & -0.8608 & -1.0570 & 0.1922 \\ -0.4677 & 10.7847 & -0.2841 & -0.8223 \\ -0.1249 & 0.3086 & 9.9133 & -0.0942 \end{pmatrix},$$

where A_2 results from A_1 with each entry perturbed by standard normal noise. In addition, we assume that the disturbance/measurement and initial errors to be the same in both secnarios; more specifically, $\mathbb{C}(x^i_0)=i10^{-1}I_4$ where I_4 is the 4×4 identity matrix and $\mathbb{C}(w^i_k)=i10^{-1}I_4$ and $\sigma_i=0.1$ for $i=1,2$. Simply speaking, we are assuming that there is only uncertainty in the system dynamic's model.

In Table III-B, we present the values of the objective in $\mathcal{P}_1$ for the possible subcollection of sets with at most two elements due to restricting $r=2$, and the instance of time consider is $k=3$. Notice that by performing the usual greedy strategies used for submodular functions on $\mathcal{P}_1$ with $r = 2$, one would first select sensor that minimizes the maximum achieved value of performance (measured as in the objective in $\mathcal{P}_1$) across the two different scenarios. Therefore, one first chooses sensor 4, and in the second step would choose sensor 2, leading to a final selection of $\{2, 4\}$ that incurs in -82.0580, whereas the optimal value is -82.1284 achieved when the pair $\{1, 3\}$ is considered. Consequently, there is an optimality gap of 0.0704, which can be made larger by increasing $\mathbb{C}(x_0^i)$ and $\mathbb{C}(w_k^i)$ by a multiplicative factor, while decreasing by another the term σ_i, which follows from the analysis of (6). ◦

$\mathcal{J}$	$F_1(\mathcal{J})$	$F_2(\mathcal{J})$	$\max\{F_1, F_2\}$
$\emptyset$	-36.8414	-36.8414	-36.8414
$\{1\}$	-65.2056	-64.3435	-64.3435
$\{2\}$	-65.2056	-64.0357	-64.0357
$\{3\}$	-65.2056	-64.9990	-64.9990
$\{4\}$	-65.2056	-65.2516	-65.2056
$\{1, 2\}$	-78.4066	-77.9726	-77.9726
$\{1, 3\}$	-82.9736	-82.1284	-82.1284
$\{1, 4\}$	-78.4066	-78.7959	-78.4066
$\{2, 3\}$	-78.4066	-78.0992	-78.0992
$\{2, 4\}$	-82.9736	-82.0580	-82.0580
$\{3, 4\}$	-78.4066	-78.1341	-78.1341

TABLE I

EVALUATION OF OBJECTIVE SET FUNCTION IN $\mathcal{P}_1$.

C. Proposed Approach

In this subsection, we propose a bisection algorithm with optimality guarantees when the constraints are relaxed. Towards this goal, notice that $\mathcal{P}_1$ can be re-written as follows:

$$\begin{aligned} \underset{c \in \mathbb{R}}{\text{minimize}} \quad & c \\ \text{subject to} \quad & \log\det\left(\Sigma^i_{z_{k-1}}(\mathcal{J})\right) \le c, \ i \in \{1, \ldots, N\}, \\ & |\mathcal{J}| \le r, \quad \mathcal{J} \subseteq \{1, \ldots, n\}. \end{aligned} \tag{$\mathcal{P}_2$}$$

Now, given a value of c, it follows that $\mathcal{P}_2$ is closely related with the following optimization problem

$$\begin{aligned} \underset{\mathcal{J} \subseteq \{1,\ldots,n\}}{\text{minimize}} \quad & |\mathcal{J}| \\ \text{subject to} \quad & \log\det\left(\Sigma^i_{z_{k-1}}(\mathcal{J})\right) \le c, \ i \in \{1, \ldots, N\}. \end{aligned} \tag{$\mathcal{P}_3$}$$

More specifically, if there exists $\mathcal{J}$ such that $|\mathcal{J}| \le r$, then $\mathcal{J}$ achieves feasibility of the constraints in $\mathcal{P}_2$. Nonetheless, $\mathcal{P}_3$ problem is also computationally challenging as we state in the following result.

Theorem 2: The problem $\mathcal{P}_3$ is NP-hard. ⋄

Proof: The proof follows by noticing that under a single scenario, obtain as particular case the *minimal sensor placement problem under budget constraint* for LTI systems, see [13] for details. ■

Hereafter, we show that it is possible to reformulate $\mathcal{P}_3$ such that a suboptimal solution with optimality guarantees is available. Subsequently, one just requires to execute such strategy for different values of c, until an optimal value of c is achieved.

In order to obtain a suboptimal solution to $\mathcal{P}_3$, we notice that the main issue is the existence of several submodular constraints that have to be simultaneously satisfied. Let $F_i(\mathcal{J}) = \log\det\left(\Sigma^i_{z_{k-1}}(\mathcal{J})\right)$, then these constraints can be replaced by an upper-truncated function $F_i^c(\mathcal{J}) = \max\{F_i(\mathcal{J}), c\}$ that is also submodular [14], and averaged out by considering $\bar{F}^c(\mathcal{J}) = \frac{1}{N}\sum_{i=1}^N F_i^c(\mathcal{J})$. Subsequently, for c, it follows that $\mathcal{F}_i^c(\mathcal{J}) \le c$ if and only if $\bar{F}^c(\mathcal{J}) = c$. Thus, we obtain the following problem:

$$\begin{aligned} \underset{\mathcal{J} \subseteq \{1,\ldots,n\}}{\text{minimize}} \quad & |\mathcal{J}| \\ \text{subject to} \quad & \bar{F}^c(\mathcal{J}) = \bar{F}^c(\{1, \ldots, n\}), \end{aligned} \tag{$\mathcal{P}_4$}$$

whose solution is also a solution to $\mathcal{P}_3$. Furthermore, $\mathcal{P}_4$ is a particular instance of the *submodular covering problem* that is also known to be NP-hard [15]. Notwithstanding, one can use Algorithm 1, which provides the following guarantees on the solution to $\mathcal{P}_4$.

Algorithm 1 Approximation Algorithm for $\mathcal{P}_4$

Input: The data required to $\Sigma^i_{z_{k-1}}(\mathcal{J})$ as in equation (6), i.e., $\{(\{A_i, \mathbb{C}(x_0^i), \{\mathbb{C}(w_j^i)\}_{j=1,\ldots,k}\})\}_{i=1,\ldots,m}$, a maximum number r of sensors considered, and an approximation ε of the optimal value attained in $\mathcal{P}_2$

Output: an approximate solution $\mathcal{J}$ to $\mathcal{P}_4$

$c_{\max} = \max_{i=1,\ldots,m} \left(\log\det(\Sigma^i_{z_{k-1}}(\emptyset))\right)$;

$c_{\min} = \min_{i=1,\ldots,m} \left(\log\det(\Sigma^i_{z_{k-1}}(\{1, \ldots, n\}))\right)$;

while $|c_{\max} - c_{\min}| \ge \varepsilon$ **do**

  $c = \frac{c_{\max} + c_{\min}}{2}$;

  $\mathcal{J} = \emptyset$;

  while $\bar{F}^c(\mathcal{J}) \ge c$ **do**

    for each $i \in \{1, \ldots, n\}$

      $\Delta_i = \bar{F}^c(\mathcal{J} \cup \{i\}) - \bar{F}^c(\mathcal{J})$;

    end for

    $\mathcal{J} = \mathcal{J} \cup \{\arg \max_{s \in \{1,\ldots,n\}} \Delta_s\}$;

  end while

  if $|\mathcal{J}| > r$ **then**

    $c_{\min} = c$;

  else

    $c_{\max} = c$;

  end if

end while

Theorem 3: Let $\mathcal{S}'$ be a solution obtained using Algorithm 1 for a given $c \in \mathbb{R}^+$ within a small ε-neighborhood of the optimal value (i.e., $\varepsilon << 1$), and $\mathcal{S}^*$ an optimal solution to $\mathcal{P}_4$, then

$$|\mathcal{S}'| \le \left(1 + \ln\left\{\frac{\bar{F}^c(\{1, \ldots, n\}) - \bar{F}^c(\emptyset)}{\bar{F}^c(\{1, \ldots, n\}) - \bar{F}^c(\mathcal{S}^{T-1})}\right\}\right)|S^*|,$$

where $\mathcal{S}^{T-1} \subset \{1, \ldots, n\}$ at the iteration $T-1$, before $\mathcal{S}' = \mathcal{S}^T$ at iteration T is obtained. Furthermore, the computational complexity of Algorithm 1 is given by $\mathcal{O}(Nk^3n^5|\ln(\varepsilon)|)$, where k is the finite-time horizon of the estimate sought. ■

Proof: The bound follows by invoking property (iii) of Theorem 1 in [15]. The complexity follows from noticing that we need to perform N times the evaluation of the objective, i.e., the number of scenarios, with the set $\mathcal{S}'$ together with one element of $\{1,\ldots,n\}\setminus\mathcal{S}'$ at each iteration; thus, requiring a total of $\mathcal{O}(Nn^2)$. In addition, each time a different subset is considered a inverse of a matrix has to be computed to obtain $\Sigma^i_{z_{k-1}}(\mathcal{J})$ (see (6)), as well as its inverse, which requires $\mathcal{O}((kn)^3)$ operations, for a specific floating-point precision. Finally, we notice that the outer for-loop corresponds to the bissection method, where the number of iterations required to ensure a given approximation error ε from the optimal value is greater or equal to $\frac{\log(c_{\max}-c_{\min})-\log(\varepsilon)}{2}$, where $c_{\max}$ and $c_{\min}$ are the endpoints of the interval used as starting point to the bissection methods, as described in Algorithm 1. Hence, the overall computational complexity follows. ■

Furthermore, unless P=NP, there is no algorithm that improves on the optimality guarantees of Theorem 3, as shown in the following.

Theorem 4: Let $\delta:\mathbb{Z}_+\rightarrow\mathbb{Z}_+$, and suppose that there is a polynomial-time algorithm for $\mathcal{P}_2$ that returns a set S' satisfying

$$\frac{|S'|}{|S^*|}\leq 1+\log\left\{\frac{\overline{F}^c(\{1,\ldots,n\})-\overline{F}^c(\emptyset)}{\overline{F}^c(\{1,\ldots,n\})-\overline{F}^c(\mathcal{S}^{T-1})}\right\}-\delta(n) \tag{10}$$

for any instance of the sensor placement problem, then P=NP. ◇

Proof: Let U and $T_1,\ldots,T_m$ define an instance of the hitting set problem, and construct an instance of the robust sensor placement problem with $\alpha=1$, $\mathbb{C}(x_0)=I$, $\mathbb{C}(w_k)=I$, and $A^{(l)}$ defined for $l=1,\ldots,N$ by (9). Select a parameter c by

$$c=\min_{l,\mathcal{J}}\left\{\log\det\left(\Sigma^l_{z_{k-1}}(\mathcal{J})\right):\mathcal{J}\cap T_l\neq\emptyset\right\}.$$

We have that $c>0$ since any hitting set satisfies $\log\det\left(\Sigma^l_{z_{k-1}}(\mathcal{J})\right)>0$. Under this choice of c, the value of $F^c_i(\mathcal{J})$ is equal to c if $\mathcal{J}\cap T_i\neq\emptyset$ and 0 otherwise. Hence, $F^c(\mathcal{S}^{T-1})\leq\frac{1}{N}((N-1)c)$, $F^c(\emptyset)=0$, $F^c(V)=1$.

By assumption, there is a polynomial-time algorithm satisfying

$$\begin{aligned}\frac{|S'|}{|S^*|}&\leq 1+\log\left\{\frac{\overline{F}^c(\{1,\ldots,n\})-\overline{F}^c(\emptyset)}{\overline{F}^c(\{1,\ldots,n\})-\overline{F}^c(\mathcal{S}^{T-1})}\right\}-\delta(n)\\&\leq 1+\log N-\delta(n)\end{aligned}$$

implying that there is a polynomial-time algorithm for hitting set problem with an optimality bound strictly better than $\log N$, and hence P=NP by [16]. ■

IV. Examples and Discussion

First, we notice that by considering Algorithm 1 when the scenario described in Section III-B is consider leads to the optimal solution depicted in blue in Table I. Therefore, we can argue that the proposed methods works in a setup where previously available tools fail. Next, we consider a 64-channel electroencephalogram (EEG) data set associated with different motor and imagery tasks [17], acquired with the BCI2000 system with a sampling rate of 160Hz [18]. More specifically, in these tasks, the subjects sit in front of a screen where targets might appear at the right/left/top/bottom side of the screen. Upon identifying the target, each subject is asked to open and close the corresponding fists or feet as a function of where the target appears. Different individuals performed 14 experimental runs consisting of one minute with eyes open, one minute with eyes closed, and three two-minute runs of 4 interacting tasks with the target: (*Task 1*) open and close left or right fist as the target appears on either left or right side of the screen; (*Task 2*) imagine opening and closing left or right fist as the target appears on either left or right side of the screen; (*Task 3*) open and close both fists or both feet as the target appears on either the top or the bottom of the screen and (*Task 4*) imagine opening and closing both fists or both feet s the target appears on either the top or the bottom of the screen.

We consider one individual, and estimated the coupling matrices and different fractional-order parameters associated with the different tasks (i.e., scenarios) [5]. In addition to the four different scenarios, we assumed that the disturbance covariance across time is given by the identity matrix, whereas the initial state covariance captures the spatial impact of the disturbance, i.e., is given by $\mathbb{C}(x_0)=A\mathbb{C}(w_0)A^\intercal$. Furthermore, we assumed the measurements error's covariance to be described by $0.1\mathbb{I}_{64}$, and we aimed to consider an estimate after $k\in\{1,4\}$ time steps, while prescribing $r\in\{1,2,\ldots,14\}$ sensors, and within an approximation of 0.1 of the optimal value.

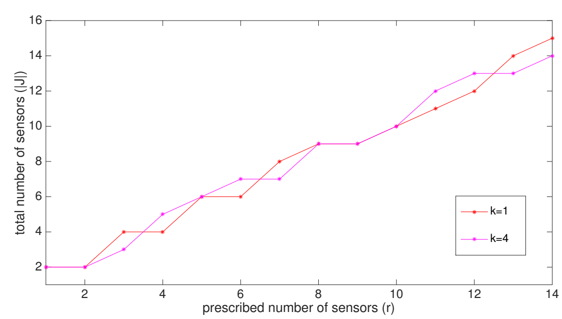

Fig. 1. This figure exhibits the total number of sensors obtained with the proposed algorithm when the prescribed number of sensors is r, and k time steps are considered for estimation.

The results are plotted in Fig. 1 and Fig. 2, where we depict the number of sensors obtained and the objective value achieved, respectively. It is worth noticing that more time estimate, i.e., to collect data and compute an estimate, the better is estimation of the state, i.e., lower uncertainty of the state estimate, which is supported by the fact that lower values of c are achieved with a larger number of time steps used for estimation. Specifically, by employing Algorithm 1 when we seek to estimate the state in one time step with 14 sensors (the same number of sensors utilized in some of the commercially available technology, e.g. (Emotiv-Epoc) [19]), we obtain $\mathcal{J}' = \{25, 63, 41, 11, 42, 9, 28, 49, 43, 19, 46, 44, 15, 60, 1\}$, whose order is consistent with the selection of sensors. In addition, the final value of c equals -191.7725. In Figure 3, we depict the location of the sensors determined by our approach and those currently used in Emotiv-Epoc. Therefore, one may argue that the location of the sensors has to be revisited aiming the use of system theoretical properties that could be leveraged to develop better and more reliable brain-machine interfaces.

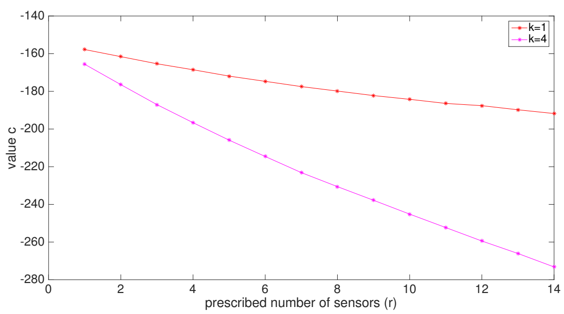

Fig. 2. This figure describes the objective value c achieved for each problem where the time horizon is k.

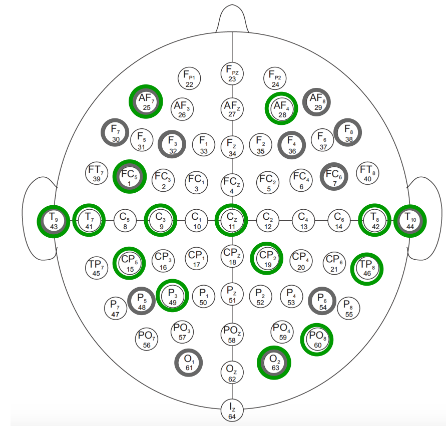

Fig. 3. This figure provides a comparison between the location of the sensors in the EMOTIV-EPOC EEG cap (depicted in gray) versus the locations associated with the solution of the proposed problem (depicted in green) – see details in Section IV.

V. Conclusions and Future Research

We considered the problem of sensors placement to estimate the state of a plant described by discrete-time fractional-order system. More specifically, we assumed that these systems' parameters and disturbance/measurement noise characteristics describe possible scenarios, and the goal was to select a subset of sensors that will optimally perform (in a minimum squared error sense) among multiple (finite) scenarios. We showed this problem to be NP-hard, and provided a bisection algorithm with suboptimality guarantees, and showed that no other algorithm ensures better performance under the same set of assumptions. Finally, as proof-of-concept we presented some simulations that illustrate the applicability of the main results in a electroencephalogram data associated with four different tasks. In addition, we conclude that the location of sensors in some of the current technology can be re-design to obtain optimality with respect to state estimation-metrics, as explored in this paper.

Future research focus in exploring the existence of metrics that capture estimation of the state, as well as the identification of the models. In addition, we seek to derive faster computationally schemes for possible online implementation, e.g. recursive schemes. Also, we are interested in understanding performance issues in the context of closed-loop systems, with potential application on the development of a principled analysis and design of brain-machine-brain interfaces.

References

[1] T. Kailath, *Linear Systems*, ser. Prentice-Hall information and system sciences series. Prentice Hall International, 1998.

[2] C.-T. Chen, *Linear System Theory and Design*, 3rd ed. New York, NY, USA: Oxford University Press, Inc., 1998.

[3] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal kalman filtering: Fundamental limits, submodularity, and algorithms," in *Proceedings of the American Control Conference*, 2016.

[4] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor Placement for Optimal Kalman Filtering: Fundamental Limits, Submodularity, and Algorithms," *ArXiv e-prints*, Sep. 2015.

[5] Y. Xue, S. Rodriguez, and P. Bogdan, "A spatio-temporal multi-fractal modeling approach for cyber-physical systems with applications to brain-machine-body interfaces," *In the Proceedings of the Design, Automation and Test in Europe (DATE) Conference*, 2016.

[6] G. Werner, "Fractals in the nervous system: Conceptual implications for theoretical neuroscience," *Frontiers in Physiology*, vol. 1, p. 15, 2010.

[7] A. Krause, B. McMahan, C. Guestrin, and A. Gupta, "Selecting observations against adversarial objectives," in *Proceedings of the Neural Information Processing Systems*, Vancouver, Canada, December 2007.

[8] S. Pequito, P. Bogdan, and G. J. Pappas, "Minimum number of probes for brain dynamics observability," in *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 306–311.

[9] V. Tzoumas, Y. Xue, S. Pequito, P. Bogdan, and G. J. Pappas, "Selecting sensors in biological fractional-order systems," in *Under Review*, 2016.

[10] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[11] S. Venkatesh, *The Theory of Probability: Explorations and Applications*. Cambridge University Press, 2012.

[12] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson, *Introduction to Algorithms*, 2nd ed. McGraw-Hill Higher Education, 2001.

[13] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Near-optimal sensor scheduling for batch state estimation: Complexity, algorithms, and limits," *arXiv preprint arXiv:1608.07533*, 2016.

[14] T. Fujito, "Approximation algorithms for submodular set cover with applications," *IEICE Transactions on Information Systems*, vol. 83, 2000.

[15] L. A. Wolsey, "An analysis of the greedy algorithm for the submodular set covering problem," *Combinatorica*, vol. 2, no. 4, pp. 385–393, 1982.

[16] U. Feige, "A threshold of ln(n) for approximating set cover," *Journal of the ACM (JACM)*, vol. 45, no. 4, pp. 634–652, 1998.

[17] "EEG Motor Movement/Imagery Dataset," *PhysioNet*. [Online]. Available: http://www.physionet.org/pn4/eegmmidb/

[18] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, June 2004.

[19] "Emotiv Epoc website, url=https://emotiv.com/epoc.php."