Fifty-fourth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 27 - 30, 2016

Minimum Number of Sensors to Ensure Observability of Physiological Systems: a Case Study

Yuankun Xue ◇ Sérgio Pequito ⋆ Joana R. Coelho † Paul Bogdan ◇ George J. Pappas ⋆

***Abstract*— Physiological signals are often spatiotemporal dependent. Some of these signals include electroencephalogram (EEG), electromyogram (EMG) or electrocardiogram (ECG) signals, just to name a few. Coupled-fractional dynamical systems have shown to capture such properties, but methods to determine where to place sensors for these dynamical systems are still lacking. In this paper, we show that the problem of determining the minimum number of sensors such that the global dynamics can be recovered from the collected data is NP-hard. Subsequently, we propose a polynomial algorithm to obtain suboptimal solutions with optimality guarantees. Finally, as a case study, we illustrate the proposed selection mechanism in the context of determining a small subset of sensors to retrieve the overall dynamics in the context of EEG, EMG and ECG signals.**

I. Introduction

Advances in wearable technology allow us to measure all sorts of physiological signals. Fitbit is probably the most familiar to the public in general. It measures data such as the number of steps per day, quality of sleep, steps climbed, and other personal metrics; and Emotiv is a company that develops brain-computer interfaces based on electroencephalography (EEG) technology. Nonetheless, these and other technologies are still in their infancy, and far from allowing to predict, for instance, a cardiac arrest or the beginning of an epileptic seizure, respectively. The benefits generated by such capability are expected to revolutionize healthcare, in the sense that it will allow an impairment with a smartphone via wireless or bluetooth, capable of performing an emergency call, or stimulate locally certain points of the human body to mitigate the effects that follow each occurrence.

Towards that auspicious future, we propose to explore a class of models that will enable the characterization of certain physiological signal dynamics, and allow us to retrieve the overall dynamics associated with certain technologies by only considering a subset of its signals. These models rely on the assumption that acquired signals of the same physiological phenomena have *spatiotemporal properties* that can be captured by the model. This is the case for the electrocardiogram (ECG) signal that records the electrical voltages generated by the heart over a period of time, using electrodes placed across the patient's body: limbs and chest. Commonly, there are between twelve and sixteen ECG electrodes placed on the surface of the patient's body. These sensors capture the potential differences between electrodes that arise from heart-muscle depolarization during each cardiac cycle, measured from different angles in a frame whose heart sets the origin [1] – which emphasizes the *spatial* and *temporal interdependence* of the process. Therefore, one may wonder if wearable technology, such as Fitbit, which matches one of the leads locations, would suffice to predict abnormal cardiac activity, see Figure 1.

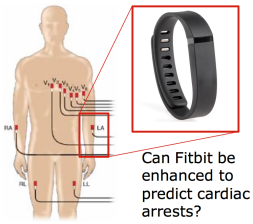


Fig. 1. Can the Fitbit sensor be used to retrieve the overall activity of the remaining sensors, and, in particular, assess if one is about to have a cardiac arrest?

Similarly, the EEG sensors measure electrical potentials (brain waves) resulting from ionic current within the neurons of the brain. Due to the nature of the electrodes used, i.e., different sensors may collect data associated with same spacial activity, and, thus, capturing *interdependent temporal* and *spatial* signals. The frequency of the brain waves recorded from the surface of the scalp can range from once per few seconds to more than 50Hz. The aspect of these waves is dependent on the activity of the correspondent cerebral cortex being acquired, and they can change drastically between states of wakefullness, sleep, coma or in brain diseases such as epilepsy [1]. Finally, an electromyography (EMG) signal detects the electrical potential generated by the skeletal muscle cells when these cells are electrically or neurologically activated, which results in a technique for evaluating and recording the electrical activity produced by skeletal muscles. These signals can then contain dynamical fingerprints that are associated with a configuration of muscle activation, leading to the execution of a specific task. Once again, due to the interdependence of the different biological systems (muscles) in the human body, it is easy to understand that the *spatial* and *temporal dependence* is present. Furthermore, these signals can be analyzed to detect and predict medical abnormalities, activation level, or recruitment order, or to

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. P.B. acknowledges the support by NSF 1453860 and 1331610 grants. In addition, it was supported in part by AFOSR Complex Networks Program and in part by AFOSR MURI CHASE.

◇ Ming Hsieh Department of Electrical Engineering, Viterbi School of Engineering, University of Southern California, USA

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

† Department of Biomedical Engineering, Instituto Superior Técnico, University of Lisbon, Portugal

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

analyze the biomechanics of human or animal movement.

Unlike current mathematical modeling approaches that rely on memoryless assumptions, the statistical analysis of physiological processes (ECG, EEG, blood glucose) demonstrates that they possess significant degree of long-range memory and fractality. In particular, numerous recent studies show that physiological processes can be more accurately modeled via *fractal order dynamical systems* [2]–[6]. Nonetheless, these efforts have either only demonstrated that the statistics of the physiological dynamics is of fractal nature (e.g., the autocorrelation function or the power spectrum exhibit a power law behavior), or only accounted for the temporal dependence of the signals (e.g., demonstrating that there is some form of persistence in time between changes in the magnitude of the physiological process). Subsequently, the spatial dependence existing between various physiological processes cannot be leveraged to obtain information about the signals that are not being directly measured.

To overcome these limitations, we propose to use *coupled discrete-time fractal order dynamical systems* (CDFODS) [7], [8], to capture the spatiotemporal characteristics existing between physiological processes. Furthermore, if we aim to retrieve the state of the system from the measurements alone, commonly referred to as *static observability*, the solvability of a set of p measurements equations to recover an n-dimensional parameter is required. Hence, requiring at least as many measurements as the number of unknowns, $p \geq n$, in general. Subsequently, by leveraging the existence of a model, one can retrieve its states from the collected data obtained from a small subcollection and the model if the system is (dynamically) *observable* [9] – see Section III for details. In this paper, we address the problem of determining the smallest subcollection of variables that need to be measured to ensure observability of the CDFODS.

Furthermore, in this paper, we propose to extend the use of submodularity tools to find an approximation of the minimum number of sensors in the CDFODS context (which we prove to be NP-hard). Also, to the best of our knowledge, this is the first time that this problem is studied in the setup mentioned. Submodular tools are used across multiple fields of science, for instance, mathematics, economics, circuit theory, operation research and machine learning [10]–[13]. In particular, examples of machine learning applications include static sensor selection [14], [15] where a dynamic model is not explicitly considered; natural language processing [16]; robotics applications [17], [18]; and spatiotemporal processes modeled as linear-time-invariant models under uncertainty [19], just to name a few. The problem of determining the minimum number of sensors to ensure observability in linear-time invariant systems has been explored in [20], and its optimal solutions in [21]. In addition, we notice that the previous solutions [20], [21] considered the continuous-time integer (non-fractional) order systems. It is also important to notice that although [20], [21] propose approximations that resemble the greedy algorithms known to approximate those whose objective are given by a submodular function, they do not explicitly use these notions as we propose in this paper. The closest work to the current paper is that in [8], where similar models where considered to model EEG, but assumed that the parameters are not known and can be considered as free unknown variables. Furthermore, the minimum number of sensors can be determined optimally, and polynomially, by resorting to structural systems theory. Nonetheless, the observability of the system can only be ensured for almost all parametric settings, whereas in several applications those conditions may not hold. In particular, it only requires a couple of sensors for the EEG setting, whereas for a given setting it may require more sensors to ensure observability (as we present in detail in the case studies).

The contributions of this paper are threefold: (*i*) we show that the problem of determining the minimum number of sensors required to ensure observability of CDFODS is NP-hard; (*ii*) we propose to use submodularity theory to approximate the results, which yields optimality guarantees; and (*iii*) we present, as a case study, its application in the context of three physiological signals (i.e., EEG, ECG and EMG). Hence, validating the proposed model and explicitly exploring the applicability of the main theoretical findings.

The remaining of the paper is organized as follows: Section II discusses the CDFODS and the mathematical formulation of the problem of determining the minimum number of sensors to obtain an observable system. In Section III, we revisit some of the properties of the CDFODS and submodularity theory. The theoretical results of this paper are presented in Section IV. Finally, in Section V, we exploit the proposed framework and results in the context of several physiological applications.

II. Problem Statement

Consider a model dynamics described by a linear discrete-time fractional-order system as follows

$$\begin{bmatrix} D^{\alpha_1} & & \\ & \ddots & \\ & & D^{\alpha_n} \end{bmatrix} x[k+1] = Ax[k], \qquad (1)$$

where $k = 0, 1, \ldots, K$, $\alpha_i \in \mathbb{R}^+$ for $i = 1, \ldots, n$, and D^{α_i} is the discretized fractional-order operator [9], and $x[0] = x_0 \in \mathbb{R}^n$ the initial condition. Further, for brevity, we refer to the dynamics in (1) as $\mathcal{F}(A; \boldsymbol{\alpha}, K)$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$.

In addition, assume that a collection of sensors is deployed to collect data about the state of the system. If a sensor is able to capture a linear combination of state variables, then the collection of sensors measurements can be represented as follows:

$$y[k] = Cx[k], \qquad (2)$$

where C is a matrix with appropriate dimensions encoding the linear combination of measured states. Nonetheless, as it often occurs in several setups (in particular, in physiological systems), we can assume that the sensors are dedicated, i.e., they measure a single state, and the output captured can be described as follows:

$$y[k] = \mathbb{I}_n^{\mathcal{J}} x[k], \qquad (3)$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

where $\mathcal{J}$ denotes the indices of the state variables being measured, and $\mathbb{I}_n^{\mathcal{J}}$ denotes the matrix that consists of the rows of the $n \times n$ identity matrix with indices in $\mathcal{J}$.

A system $\mathcal{F}(A; \boldsymbol{\alpha}, K)$ and a collection of measurements (3) is said to be (dynamically) *observable* at time $k = 0$ if and only if there exists some K such that the state x_0 can be uniquely determined from the knowledge of y_k and $\mathcal{F}(A; \boldsymbol{\alpha}, K)$. Subsequently, the goal of this paper is to determine the minimum $\mathcal{J}$ in (3) that, together with $\mathcal{F}(A; \boldsymbol{\alpha}, K)$, ensures observability of the system at time $k = 0$. This problem can be formally stated as follows.

Minimum Sensor Placement Problem

Given $\mathcal{F}(A; \boldsymbol{\alpha}, K)$, determine the minimum number of dedicated sensors $\mathcal{J}^*$ such that

$$\begin{aligned} \mathcal{J}^* = \arg \min_{\mathcal{J} \subset \{1,\dots,n\}} \quad & |\mathcal{J}| \\ \text{s.t.} \quad & (\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n^{\mathcal{J}}) \text{ is observable.} \end{aligned} \tag{4}$$

○

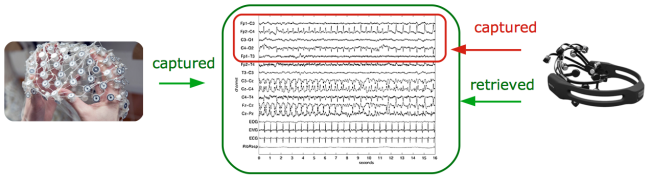


Fig. 2. Is it possible to replace the cap with 64 sensors by the EMOTIV-EPOC (depicted on the right-hand side) with 14 sensors, while being able to retrieve the overall data using the dynamic's model?

Figure 2 illustrates one possible application of this problem, i.e., determine if a subcollection of sensors allow to retrieve the overall evolution of EEG signals. In particular, determining the smallest collection of such sensors will lead to the development of energy efficient EEG wearables. In Section V, we will discuss in further detail this particular application.

III. CDFODS Observability and Submodularity

In this section, we first review some properties of the CDFODS and the characterization of the feasibility space of (4), followed by a brief rehash of submodularity properties.

We start by observing that the closed-form to (1) can be described as follows.

Lemma 1 ([9], [22]): The solution to (1) is given by

$$x[k+1] = G_{k+1} x[0], \tag{5}$$

where

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

with $A_0 = A$ and $A_j = \text{diag}(-(-1)^{j+1}\binom{\alpha_1}{j+1}, \dots, -(-1)^{j+1}\binom{\alpha_n}{j+1})$. Please note that α_i is a non-integer number and the term $\binom{\alpha_i}{j} = \frac{\Gamma(\alpha_i+1)}{\Gamma(j+1)\Gamma(\alpha_i-j+1)}$ is expressed via the Gamma function: $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ [23]. ◇

Subsequently, a sequence of K measurements can be described as follows (by invoking Lemma 1):

$$\begin{aligned} y[1] = Cx[1] = CG_0x[0],\ y[2] = Cx[2] = CG_1x[0], \\ \dots,\ y[K] = Cx[K] = CG_{K-1}x[0]. \end{aligned}$$

Subsequently, we can rewrite it as

$$y_{0:K} = \underbrace{[(CG_0)^\intercal\ (CG_1)^\intercal\ \dots\ (CG_{K-1})^\intercal]^\intercal}_{\mathcal{O}(\mathcal{F}(A;\boldsymbol{\alpha},K),C)} x[0], \tag{6}$$

where $y_{0:K} = [y^\intercal[0]\ y^\intercal[1]\ \cdots\ y^\intercal[K]]^\intercal$, and the matrix $\mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), C)$ is commonly referred to as *observability matrix*. In order to retrieve $x[0]$, we can first premultiply both hand sides of (6) by $\mathcal{O}^\intercal \equiv \mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), C)$, which leads to

$$\mathcal{O}^\intercal y_{0:K} = \mathcal{O}^\intercal \mathcal{O} x[0].$$

Thus, if $\mathcal{O}^\intercal\mathcal{O}$ is invertible, then we obtain a closed-form solution to $x[0]$, i.e.,

$$x[0] = (\mathcal{O}^\intercal\mathcal{O})^{-1}\mathcal{O}^\intercal y_{0:K}. \tag{7}$$

As a consequence, we obtain the following results:

Theorem 1 ([9], [22]): The system described by (1) and (2) is observable if and only if there exists a finite time K such that rank $(\mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), C)) = n$. ◇

Theorem 2 ([9], [22]): If the system described by (1) and (2) is observable, then the initial state x_0 can be retrieved as in (7). ◇

Remark 1: Due to the nature of Theorem 1, it may occur that a specific set of K measurements do not yield observability. Notwithstanding, it is possible that there exits $K' > K$ measurements by the same collection of sensors such that Theorem 1 holds. In other words, the observability criterion implicitly considers the tradeoffs between how much data is collected, i.e., the time allowed to collect data, and the number of sensors acquiring it. Furthermore, this analysis differs from that performed with linear time-invariant systems, where we know that a solution exists for K' equal to at most n [24]. ◇

In Section IV, we show how the aforementioned characterization of the feasibility space of (4) can be used to obtain the solution to the minimum sensor placement problem. Briefly, towards this goal, we will resort to submodularity theory [10]. Given a finite set $\mathcal{V}$ with n objects, which we denote by $|\mathcal{V}| = n$, we can define a function $f : 2^{\mathcal{V}} \to \mathbb{R}^+$ that associates a positive scalar with each subset of $\mathcal{V}$. This function is said to be *submodular* if it satisfies the so-called *diminishing returns* property, i.e., for all $\mathcal{X} \subset \mathcal{Y} \subset \mathcal{V}$ and $v \notin \mathcal{Y}$, we must have:

$$f(\mathcal{X} \cup \{v\}) - f(\mathcal{X}) \geq f(\mathcal{Y} \cup \{v\}) - f(\mathcal{Y}).$$

In other words, the increment by considering a new element in a smaller sized set is at least as high as adding it to a superset of the latter. Furthermore, when f is submodular, one can use greedy algorithms that yield approximate solutions that are often very close to the optimal solution [25], [26].

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

IV. Minimum Sensor Placement

In this section, we present the main results of this paper. More precisely, in Theorem 3 we show our problem to be NP-hard. Notwithstanding, we present a greedy strategy (see Algorithm 1) that has polynomial computational complexity, and approximates the original problem with optimality guarantees, as described in Theorem 4.

Theorem 3: The minimum sensor placement problem for CDFODS, as described in (4), is NP-hard. ◇

Proof: The proof follows by noticing that there exists a set of coefficients $\boldsymbol{\alpha}$ that leads to $A_j = \mathbb{I}_n$ for $j = 1, \ldots, K$. Therefore, (4) reduces to a linear time-invariant system, and the problem reduces to the minimum sensor placement problem for linear-time invariant systems presented in [20], that is NP-hard. Thus, because (4) contains (as subclass of problems) one that is NP-hard, it follows that (4) is also NP-hard. ■

Next, we show that the feasibility space of (4) is given by a constraint, (see Theorem 1), that is submodular.

Lemma 2: Given a CDFODS $\mathcal{F}(A; \boldsymbol{\alpha}, K)$, the function

$$f(\mathcal{J}) = \text{rank}(\mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n^{\mathcal{J}}))$$

is submodular in $\mathcal{J} \subset \{1, \ldots, n\}$. ◇

Proof: We define $\mathcal{V} = \{1, 2, \ldots, nK\}$ as the set of row vector indices for observability matrix $\mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n)$. Let $\mathcal{V}_{\mathcal{J}_A} \subseteq \mathcal{V}$ be the set of row vector indices for $\mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n^{\mathcal{J}_A})$, where $\mathcal{J}_A \subset \{1, \ldots, n\}$. Let $r(\mathcal{V}_{\mathcal{J}_A})$ be the rank of the matrix that contains as rows the row vectors $\{g_i\}_{i \in \mathcal{V}_{\mathcal{J}_A}}$ indexed by $\mathcal{V}_{\mathcal{J}_A}$. Hence, $r : 2^{\mathcal{V}} \to \mathbb{Z}^+ \subset \mathbb{R}^+$ is the rank function of observability matrix and $r(\mathcal{V}_{\mathcal{J}_A})$ is the cardinality of maximum independent subset of vectors contained within the set of vectors indexed by $\mathcal{V}_{\mathcal{J}_A}$. Therefore, $r(\mathcal{V}_{\mathcal{J}_A}) = \text{rank}(\mathcal{O}(\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n^{\mathcal{J}_A}))$.

Now, consider two subset $\mathcal{V}_{\mathcal{J}_A}, \mathcal{V}_{\mathcal{J}_B} \subseteq \mathcal{V}$. Let $\mathcal{I}$, $\mathcal{I}_{\mathcal{AB}}$, $\mathcal{I}_{\mathcal{A}}$ be the maximum independent subset of $\mathcal{V}_{\mathcal{J}_A} \cap \mathcal{V}_{\mathcal{J}_B}$, $\mathcal{V}_{\mathcal{J}_A} \cup \mathcal{V}_{\mathcal{J}_B}$ and $\mathcal{V}_{\mathcal{J}_A}$, respectively. By definition, we have $\mathcal{I} \subseteq \mathcal{I}_{\mathcal{A}} \subseteq \mathcal{I}_{\mathcal{AB}}$. To prove the submodularity of the rank function r, we need to show that

$$r(\mathcal{V}_{\mathcal{J}_A}) + r(\mathcal{V}_{\mathcal{J}_B}) \geq r(\mathcal{V}_{\mathcal{J}_A} \cup \mathcal{V}_{\mathcal{J}_B}) + r(\mathcal{V}_{\mathcal{J}_A} \cap \mathcal{V}_{\mathcal{J}_B}) \quad (8)$$

or, equivalently,

$$r(\mathcal{V}_{\mathcal{J}_B}) \geq |\mathcal{I}_{\mathcal{AB}}| - |\mathcal{I}_{\mathcal{A}}| + |\mathcal{I}|, \quad (9)$$

due to the fact that $|\mathcal{I}| = r(\mathcal{V}_{\mathcal{J}_A} \cap \mathcal{V}_{\mathcal{J}_B})$, $|\mathcal{I}_{\mathcal{AB}}| = r(\mathcal{V}_{\mathcal{J}_A} \cup \mathcal{V}_{\mathcal{J}_B})$ and $|\mathcal{I}_{\mathcal{A}}| = r(\mathcal{V}_{\mathcal{J}_A})$. In addition, observe that $\mathcal{I}_{\mathcal{AB}} \cap \mathcal{V}_{\mathcal{J}_B}$ is also a set of independent vectors and a subset of $\mathcal{V}_{\mathcal{J}_B}$. Therefore, we have

$$r(\mathcal{V}_{\mathcal{J}_B}) \geq |\mathcal{I}_{\mathcal{AB}} \cap V_{\mathcal{J}_B}|, \quad (10)$$

and since $\mathcal{I}_{\mathcal{AB}} \cap \mathcal{V}_{\mathcal{J}_B} = \mathcal{I}_{\mathcal{AB}} \setminus (\mathcal{V}_{\mathcal{J}_A} \setminus \mathcal{I})$, it follows that

$$r(V_{\mathcal{J}_B}) \geq |\mathcal{I}_{\mathcal{AB}} \setminus (\mathcal{V}_{\mathcal{J}_A} \setminus \mathcal{I})| = |\mathcal{I}_{\mathcal{AB}}| - |\mathcal{I}_{\mathcal{A}}| + |\mathcal{I}|. \quad (11)$$

Hence, the function $f(\mathcal{J}) = r(\mathcal{V}_{\mathcal{J}})$ is submodular. ■

As a consequence of Lemma 2, the problem 4 can be approximately solved via greedy strategy that repeatedly adds a sensor that maximizes the increase of the rank of the observability matrix until Theorem 1 holds. Algorithm 1 implements such approach. Briefly, this algorithm is initiated with an empty set $\tilde{\mathcal{J}}$ to which (at each iteration) will be added an index. Towards this goal, every index is considered to assess how much the the rank of the observability matrix increases, and one of the indices that maximizes such difference is, subsequently, added to $\tilde{\mathcal{J}}$. The loop repeats until the rank stops increasing (which always does after at most n iterations), so one obtains an approximation $\tilde{\mathcal{J}}$ to (4), with computational complexity and optimality guarantees described in the next result.

Algorithm 1 Greedy Algorithm to (4)

Input: CDFODS $\mathcal{F}(A; \boldsymbol{\alpha}, K)$;
Output: $\mathcal{J}^*$ that is an approximate solution to (4);

```
Initialize J̃ = ∅;
repeat
  Initialize Δr* = 0;
  r_0 = rank(O(F(A; α, K), I_n^J̃));
  for i = 1 to n do
    J' = J̃ ∪ {i};
    Δr(i) = rank(O(F(A; α, K), I_n^J')) − r_0;
    if Δr(i) > Δr* then
      Δr* = Δr(i);
      i* = i;
    end if
  end for
  if Δr* > 0 then
    J̃ = J̃ ∪ {i*};
  end if
until Δr* = 0.
```

Theorem 4: Algorithm 1 approximates the sensor placement problem for CDFODS described in (4) in polynomial-time, i.e., with computational complexity $\mathcal{O}(n^5)$. Furthermore, it achieves a solution that is at most 33% worse than the optimal solution despite the state dimension. ◇

Proof: Algorithm 1 is a greedy method that progressively picks up the index of a column vector from the identity matrix and adds it to $\mathcal{J}$ such that the maximal increase of rank for observability matrix $O_k(\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n^{\mathcal{J}})$ in each iteration is obtained. The algorithm will terminate when no further increase is possible. Therefore, observe that the outer loop will at most be executed n times as $r(\mathbb{I}_n^{\mathcal{J}}) = |J|$, offers a lower bound for the rank of the observability matrix. Furthermore, no more than n indices will be added to $\mathcal{J}$ until the row rank of $O_k(\mathcal{F}(A; \boldsymbol{\alpha}, K), \mathbb{I}_n^{\mathcal{J}})$ is n, i.e., the system is observable. In each iteration, the algorithm will loop over all n indices and check their possible contribution to the rank increase when the corresponding column vector is added to $\mathbb{I}_n^{\mathcal{J}}$. Therefore, any operation in the algorithm will be performed no more than n^2 times before rank of observability matrix is full-rank. Since the rank verification of the observability matrix is the most expensive operation and the remaining operations (i.e., comparison, set union

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

and assignment), the cost of the rank verification operation determines the overall time complexity of the algorithm. Checking the rank of a matrix can be done using rank-revealing QR factorization with a complexity of $\mathcal{O}(n^3)$, up to a specified numerical precision. Therefore, the overall complexity of the algorithm is $\mathcal{O}(n^5)$. In addition, this algorithm is the greedy algorithm used for submodularity function when the number of sensors is also constraint to be less or equal to a specified value. Notwithstanding, due to the fact that we have always a solution in at most n steps, our problem is virtually indistinguishable from the case where at each iteration k, we were considering the constrained version with at most k sensors. Hence, the algorithm ensures similar optimality guarantees to the constraint version of the problem for arbitrary submodular functions [10]; in particular, it achieves the mentioned bound. ■

In the next section, as a case study, we use of the main results presented in this section in the context of three different physiological systems.

V. Simulation Results

Now, we conduct an exhaustive application of the previous results to three different physiological signals: EMG, EEG and ECG. Towards this goal, the first step consists in the identification of the parameters for their dynamics as CFFODS are obtained as described in [7]. In each subsection, we described in detail the data used, and we argue on the quality of the estimated models, to which we report the findings using the main results of this paper.

A case-study investigation of EMG signals

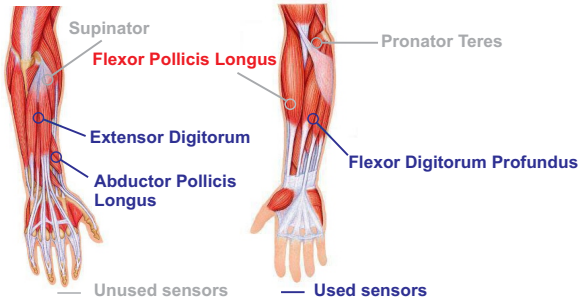


Fig. 4. Sensor distribution in the 6-channel iEMG signal clinical measurement experiment. The sensors in blue represent the minimal deployment of sensors that ensure the global dynamics of iEMG in all 6 muscles can be retrieved. The sensors in grey represent the unused sensors while the muscular activity at the location of the sensor in red is simulated based on the identified fractional order system, and compared to the data acquired.

We analyze the intramuscular EMG (iEMG) signals from clinical experiment to compare muscle contractions of transradial amputees to those of non-amputated subjects. All subjects are asked to perform forearm movements. The iEMG signals are recorded at different sites of the forearm muscles as shown in Figure 4: (*i*) extensor digitorum (ED); (*ii*) flexor digitorum profundus (FDP); (*iii*) abductor pollicis longus (APL); (*iv*) flexor pollicis longus (FPL); (*v*) pronator teres (PT); and (*vi*) supinator (SUP). Each subject is inserted with fine wire electrodes for measurement purpose. Subjects are asked to relax 6 seconds, then do the finger flexion at a consistent strength for 10 seconds. The entire process is repeated twice. The ADInstruments data acquisition system sampled the iEMG at 4 KHz after applying a 2 KHz low pass filter, and a 10 Hz high pass filter to minimize any motion artifacts from electrodes or leads. As an example, we consider iEMG signals measured when the subject 3 is performing the finger flexion at a constant strength.

After the identification of the parameters that account for the spacial-temporal characterization of the signals. It is interesting to notice that the coupling matrix captures the spatial dependencies among different motor muscles involved in the forearm movement and the fractional order exponents of different iEMG channels lie in $[-0.14, 0.19]$ that support the nature of fractional dynamics [7].

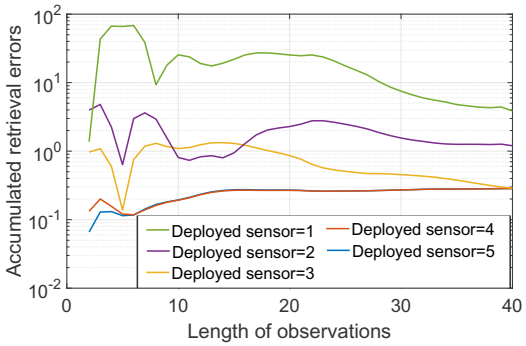


Fig. 5. The figure shows the accumulated errors in retrieved system's initial states change as a function of the number of used sensors and the observation length. The increase in number of sensor deployed provides information gain that leads to the decrease in overall reduced deviation of states retrieved compared to the actual measurements.

Based on the identified system $\mathcal{F}(A; \boldsymbol{\alpha}, K)$, we perform the greedy method proposed in Algorithm 1 to retrieve a small collection of sensors required such that the fractional order system is observable. In Figure 4, we depict in blue the identified sensors that ensure retrieval of initial states for unobserved muscular activities. In Figure 3-(a-1), we report the retrieved initial states over all 6 sensing channels compared against the actual measurements. These states were recovered using (7), when the sensors ED, FDP and APL are considered. In particular, we notice that we were able to successfully recover the initial states of sensors SUP, PT and FPL. To show the capability of retrieving the initial states, we simulate the muscular dynamics over first 10-second session of finger flexion at FPL based on the retrieved system states and the inferred innovation series from our estimation. To give an intuition, we show the comparison between the signal recorded and the simulated activities in Figure 3-(a-2). As we can see from the figure, the simulated dynamics at the non-measured FPL channel fits well to the actually recorded muscular activities; thus, validating the capability of the proposed setup (i.e., the model and the sensing capabilities determined by Algorithm 1) to recover global dynamics in the iEMG experiment.

To explore the influence of the length of the time-series, i.e., the number of measurements samples collected, on the

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

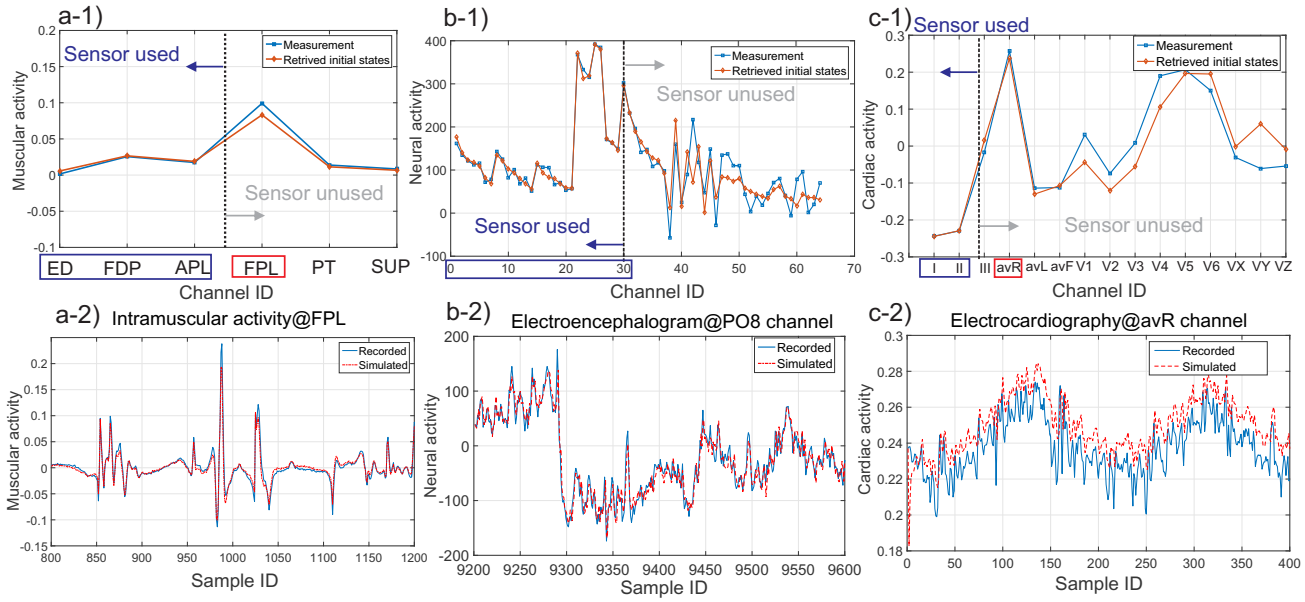


Fig. 3. The retrieved initial states of the system using (7) across different signals are presented in the top row. We show the simulated states evolution of sensors that were not considered in the experiments and compare them with actual measurements at these sensors. In the first column, Figure 3-(a-1) shows The initial states of unused sensors are recovered by the proposed algorithm based on the measurements from minimal subset of sensors (ED, APL and FDP). In Figure 3-(a-2), the simulated data based on the retrieved initial states at unused FPL-sensor is compared against the actual measurements during first 0.4 second of finger flexion. Similarly, Figure 3-(b) and Figure 3-(c) show the initial states retrieved for unused sensors and the simulated dynamics at one of these sensors during the experiments of EEG and ECG, respectively.

accuracy of retrieved global dynamics, we measured the accumulated errors of retrieved system initial states, i.e., sum of the deviation from the actual recordings over all sensing channels, which we depict in Figure 5. An important observation is that the increase in number of sensors deployed will help reduce the estimation errors, that are namely due to approximation errors and small weights in the coupling between some of the sensors' signals. Similar conclusions can be achieved when we observe the physiological process for a longer period of time, given a fixed number of sensors used. However, there exists a phase-change phenomenon as the number of sensors used decreases: one can always retrieve the system's initial states with similar accuracy when at least 3 sensors are used, and enough observations are made. However, there is a jump in the accumulated errors that persists over observation horizon if less than 3 sensors are deployed, which suggests the existence of a lower bound for the number of sensors to retrieve global dynamics (i.e., minimal observability), as predicted by the theoretical setting explored in this paper.

A case-study investigation of EEG signals

To explore the minimal observability of a neuro-physiological system, we apply the proposed Algorithm 1 to a 64-channel electroencephalogram data set which records the brain activity of 109 subjects when they are performing motor and imagery tasks. In the experiment, each subject sits in front of a screen where targets might appear at the right/left/top/bottom side of the screen. Upon noticing the target, each subject is asked to open and close the corresponding fists or feet as a function of where the target appears. Each individual performed 14 experimental runs consisting of one minute with eyes open, one minute with eyes closed, and three two-minute runs of interacting with the target. The data set is collected by BCI2000 system with a sampling rate of 160Hz [27], [28].

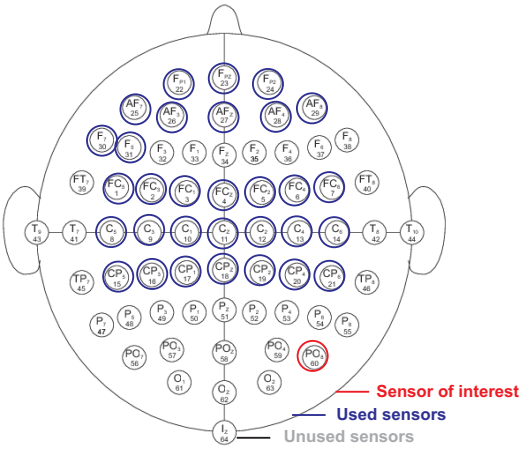


Fig. 6. The 64-channel geodesic sensor distribution for measurement of EEG. The sensors in blue represent the minimum number of sensors and their deployment to guarantee that the CDFODS in (1), whose states correspond to the channel measurements, is observable. The sensor in red is used as a sanity check on the evolution of the identified CDFODS. The results are compared against the recorded activity in Figure 3-(b-2).

Following a similar experimental setting to that of the iEMG, spatiotemporal parameters were estimated, and, in particular, the fractal order exponents range from 0.34 to 1.04. Then, using the estimated model, we used Algorithm 1

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

to obtain the sensors placement to achieve system's observability. The minimal set of sensors and their deployment are reported in Figure 6 and depicted in blue. More specifically, 30 out of 64 sensors are required to ensure observability. In addition, we retrieved the system's initial states of all 64 EEG channels by applying equation (7), whose results are reported in Figure 3-(b-1). Furthermore, as illustrated in Figure 3-(b-2), the recovered EEG system's states follow closely the actual measurements. Specifically, we simulate the neural activity at the channel PO_8, whose retrieved initial state differs by less than 4% from recorded value, and compared it to the actual EEG recordings.

In what follows, we performed an experiment to check how the length of observations affects the accuracy of the retrieved states under different neural sensor settings. The results are reported in Figure 10, and, as expected, it is shown that the errors of the retrieved states decrease as we make observations on the system dynamics for a longer period of time given a fixed set of sensors selected, and when we choose more sensors given a fixed length of observations. The similar phase-change phenomenon described for the same experiment using EMG, is also noticed when we use less than 16 sensors in which the accuracy does not improve at the expense of observation length. Further, there exists a sudden increase of errors over all sensing channel when the number of sensors we use drops from 16 to 8.

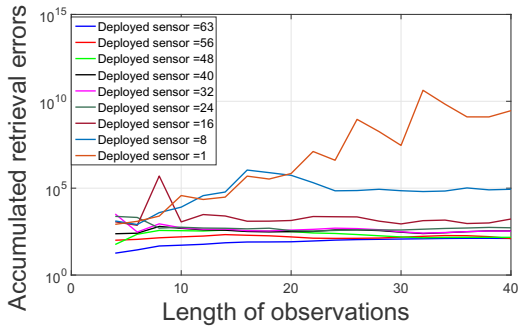


Fig. 7. The accumulated errors of the retrieved neural system's initial states over 64 channels change over observation horizon. A phase-change phenomenon is shown when number of sensors deployed drops from 16 to 8, where the errors do not decrease as more observations are made.

A case-study investigation of ECG signals

To verify the efficacy of our theoretical analysis results, we also consider the application of the algorithm to analyze the electrocardiogram (ECG). The raw clinical ECG data was extracted from the PTB diagnostic ECG database [28]. Data on 52 healthy subjects (13 women, age 48 ± 19 and 39 men, age 42 ± 14) was obtained by the National Metrology Institute of Germany. Each subject's record includes 15 different signals simultaneously acquired: the conventional 12-lead (I, II, III, aV_R, aV_L, aV_F, V_1, V_2, V_3, V_4, V_5, and V_6) and 3 Frank orthogonal leads (V_X, V_Y, and V_Z), as illustrated in Figure 7. Each signal is digitalized at 1000Hz, with a signal bandwith of 0Hz to 1KHz and with 1 μV LSB resolution [29]. As an example, we choose the ECG signal from the subject S0010 from the data set during a 39-second measurement session.

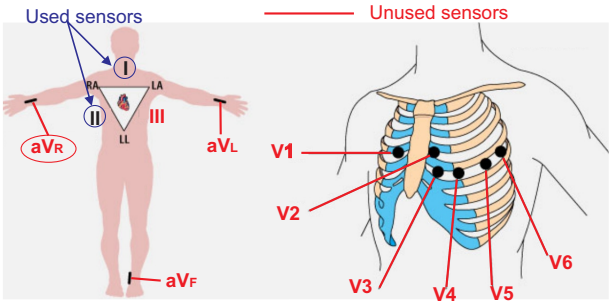


Fig. 8. The deployment of 12-lead ECG system used in the experiment is shown. The sensors in blue (I and II) is the identified minimal subset of sensors to ensure the observability of the fractional order system. We simulate the cardiac activity at aV_R given the identified fractional order system.

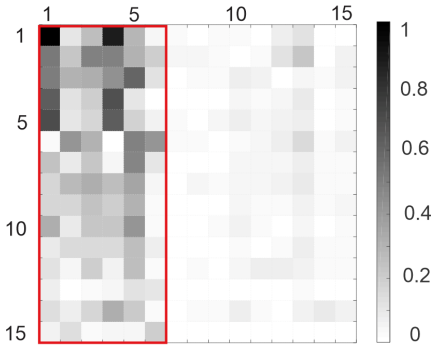


Fig. 9. The coefficients colormap of spatio-coupling matrix A. The distribution of normalized coefficients shows a subset of channels plays dominant roles in the cardiac dynamics.

After, we obtain the estimate of fractal exponents that range from 0.7051 to 1.3773, which supports the fractal nature of the underlying dynamical process. The greedy algorithm (Algorithm 1) shows we do not need more than two sensors (i.e., I and II) to retrieve the dynamics of all 15-channel cardiac activities. We show the recovered system states using the observation from only two sensors in blue in Figure 8-(c-1). An important observation is that the obtained deployment of sensors aligns with the pattern distribution of coefficients in the coupling matrix A, as described in Figure 9. A subset of channels plays the dominant role in the cardiac dynamics and these channels themselves are tightly coupled, such that we can retrieve global dynamics by the observations made upon a subset of them. This is verified by the output of algorithm and validated during our simulation to generate the system states evolution at aV_R as shown in Figure 3-(c-2). The simulated model response is strongly coherent with actual measurement. We set up the similar experiment to show the relation between the observation length and number of sensors deployed in context of a cardiac system. The results are shown in Figure 10 and it is aligned with our previous observations made in the experiments of iEMG and EEG. The lower bound for the number of sensors used to reliably retrieve the global initial states can be identified by checking where the phase-change appears. In the case of ECG, we notice there exists a sudden rise in the errors when only one sensor is considered. In contrast, we can achieve the similar accuracy if 2 or more

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.

sensors are placed in the measurement system as described in Figure 8 given enough observations. As a consequence, one can argue that the answer to the question posed in Figure 1 is negative; hence, the insights obtained with the case study can be leveraged to incorporate new features in wearable devices that will enhance their capabilities and, ultimately, improves people's quality of life.

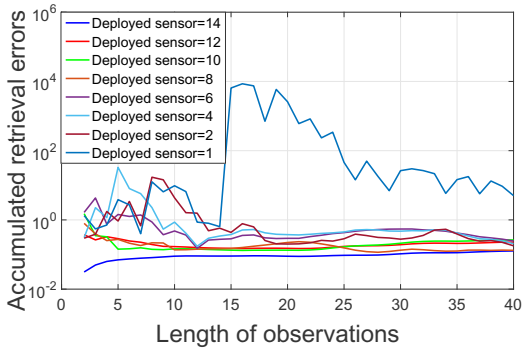


Fig. 10. The accumulated errors are plotted against the length of observations and different settings of sensors. The longer observations improve upon the overall accuracy of system states estimates while it also suggests a lower bound for number of sensors used for observability. This observation aligns with our theoretical analysis on minimal observability.

VI. Conclusions

Coupled-fractional dynamical systems capture the evolution of different physiological signals, such as electroencephalogram, electromyogram or electrocardiogram signals. In this paper, we proposed to determine the smallest collection of signals required to retrieve the overall evolution of the process modeled by the coupled discrete-time fractional-order dynamical system. In particular, we show that the problem is NP-hard, but a submodular approach can be leveraged to obtain approximate solutions with optimality guarantees. Further, we presented a case study the proposed mechanism in the context of the aforementioned physiological signals. Future research will focus in extending the present setup to the case where some uncertainty exists in the parametric models, dynamics, and measurements. In addition, it would be of interest to extend and validate the present findings when multi-scenarios are considered, for instance, several tasks, and the sensor placement provides guarantees across different models (and assumptions).

References

[1] J. E. Hall, *Guyton and Hall textbook of medical physiology*. Elsevier Health Sciences, 2015.

[2] F. Moon, *Chaotic and Fractal Dynamics: An Introduction for Applied Scientists and Engineers*. Wiley, 1992.

[3] B. N. Lundstrom, M. H. Higgs, W. J. Spain, and A. L. Fairhall, "Fractional differentiation by neocortical pyramidal neurons." *Nature Neuroscience*, vol. 11, no. 11, pp. 1335–1342, Nov. 2008.

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

[5] R. G. Turcott and M. C. Teich, "Fractal character of the electrocardiogram: Distinguishing heart-failure and normal patients," *Annals of Biomedical Engineering*, vol. 24, no. 2, pp. 269–293, 1996.

[6] S. Thurner, C. Windischberger, E. Moser, P. Walla, and M. Barth, "Scaling laws and persistence in human brain activity," *Physica A: Statistical Mechanics and its Applications*, vol. 326, no. 3-4, pp. 511 – 521, 2003.

[7] 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," *Proceedings of the Design, Automation and Test in Europe (DATE) Conference*, pp. 642–647, 2016.

[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] M. Bettayeb, S. Djennoune, S. Guermah, and M. Ghanes, "Structural properties of linear discrete-time fractional-order systems," *Proceedings of the 17th World Congress of the International Federation of Automatic Control, Seoul, Korea*, pp. 15 262–15 266, 2008.

[10] F. Bach, "Learning with Submodular Functions: A Convex Optimization Perspective," *ArXiv e-prints*, Nov. 2011.

[11] J. Edmonds, "Matroids and the greedy algorithm," *Mathematical Programming*, vol. 1, no. 1, pp. 127–136, 1971.

[12] L. Lovász, *Mathematical Programming The State of the Art: Bonn 1982*, A. Bachem, B. Korte, and M. Grötschel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1983.

[13] A. Schrijver, "A combinatorial algorithm minimizing submodular functions in strongly polynomial time," *Journal of Combinatorial Theory – Series B*, vol. 80, no. 2, pp. 346–355, Nov. 2000.

[14] A. Krause and C. Guestrin, "Submodularity and its applications in optimized information gathering," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 4, pp. 32:1–32:20, Jul. 2011.

[15] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, pp. 235–284, Jun. 2008.

[16] K. Kirchhoff and J. Bilmes, "Submodularity for data selection in machine translation," *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 131–141, October 2014.

[17] N. A. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Decentralized active information acquisition: Theory and application to multi-robot slam," *Proceedings of the 2015 International Conference on Robotics and Automation*, pp. 4775–4782, May 2015.

[18] A. Meliou, A. Krause, C. Guestrin, and J. M. Hellerstein, "Nonmyopic informative path planning in spatio-temporal models," *Association for the Advancement of Artificial Intelligence*, pp. 602–607, 2007.

[19] F. Schnitzler, J. Y. Yu, and S. Mannor, "Sensor selection for crowdsensing dynamical systems," *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2015, San Diego, California, USA, May 9-12*, 2015.

[20] A. Olshevsky, "Minimal controllability problems," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 3, pp. 249–258, Sept 2014.

[21] S. Pequito, G. Ramos, S. Kar, A. P. Aguiar, and J. Ramos, "The Robust Minimal Controllability Problem," *ArXiv e-prints*, Jan. 2014.

[22] S. Guermah, S. Djennoune, and M. Bettayeb, "Controllability and observability of linear discrete-time fractional-order systems." *Applied Mathematics and Computer Science*, vol. 18, no. 2, pp. 213–222, 2008.

[23] D. Baleanu, J. Machado, and A. Luo, *Fractional Dynamics and Control*. Springer, 2011.

[24] J. Hespanha, *Linear Systems Theory*. Princeton University Press, 2009.

[25] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions – part I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.

[26] H. Lin and J. Bilmes, "Multi-document summarization via budgeted maximization of submodular functions," *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 912–920, 2010.

[27] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "Bci2000: A general-purpose brain-computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[28] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C.-K. Peng, and H. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[29] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet," *Biomedizinische Technik*, vol. 40, no. S1, pp. 317–318, 1995.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:23:23 UTC from IEEE Xplore. Restrictions apply.