Randomized Sensor Selection for Nonlinear Systems With Application to Target Localization

Shaunak D. Bopardikar, Osama Ennasr, and Xiaobo Tan

***Abstract*—Given a nonlinear dynamical system, this letter considers the problem of selecting a subset of the total set of sensors that has provable guarantees on standard metrics related to the nonlinear observability Gramian. The key contribution is a simple randomized algorithm that samples the sensors uniformly without replacement, and yields probabilistic guarantees on the minimum eigenvalue or the inverse of the condition number of the nonlinear observability Gramian relative to that of the complete set of sensors. Numerical studies reveal that the utility of the theoretical results lies in the regime of large total number of sensors wherein the combinatorial nature of the problem presents a significant computational challenge. The results are demonstrated numerically on a problem of moving target localization using an extended Kalman filter in two scenarios: one using range sensors and another with time-difference-of-arrival measurements. A graceful degradation of performance with a decreased number of sensors is observed when compared to the use of all of the sensors for localization. It is also observed that for certain metrics, the proposed approach provides an improvement over a heuristic that selects the sensors in a greedy manner based on the contribution of an additional sensor toward the observability Gramian metric.**

***Index Terms*—Localization, probability and statistical methods, sensor networks, randomized algorithms.**

I. Introduction

SENSOR selection is a classic problem arising in robotics wherein the goal is to choose sensors out of a set in order to efficiently and effectively perform tasks such as target tracking, localization and estimation. In several applications, it is essential to understand fundamental limits on how the quality of the estimate varies with the number of sensors. This letter considers the problem of analyzing the observability of a general smooth nonlinear system as a function of the number of sensors selected. We analyze a simple randomized scheme based on uniform sampling without replacement, derive probabilistic lower bounds on two widely-used metrics of observability, and evaluate the scheme on a target tracking scenario using two types of sensors – range sensing and time difference of arrival (TDOA). The sensor selection problem addressed in this letter is closely related to the deployment of mobile (robotic) sensors wherein all potential locations can be treated as the total set of sensors and the sampled subset comprises the locations chosen for deployment.

A. Related Work

Sensor selection is a well studied problem within the robotics community. Early works include reference [1] that proposes stochastic dynamic programming to predict and evaluate different robot actions, reference [2] that uses mobile sensors to track multiple targets, and references [3], [4] that propose the use of bounded uncertainty models to compute near-optimal sensor placements. More recent contributions include reference [5] which addresses selection of sensors with Boolean outputs, reference [6], which uses the knowledge of target tracks in a multi-objective framework to perform sensor selection, and reference [7], which employs optimization of hinge-loss functions in an application to indoor localization, to name a few. Sensor selection has also been studied extensively in the control community (see survey [8]). Recently there has been a large body of work such as [9]–[11] in the area of sensor and/or actuator selection in order to optimize several metrics for linear systems. Also related is a line of work extending these ideas to the problem of sensor or actuator scheduling as addressed in [12]–[14]. Our recent work in [15] applied randomized sampling *with replacement* to provide lower bounds on several eigenvalue-based metrics of the observability Gramian for linear dynamical systems.

While there has been a lot of work around linear systems, extensions to nonlinear settings have been fewer. In this context, references [16] and [17] propose the use of the inverse of the condition number to study the observability of a nonlinear system and for target tracking. Reference [18] uses the Cramer-Rao lower bound for optimizing the sensor array for source localization using TDOA measurements. In references [19], [20], distributed localization of a moving target using TDOA was proposed. Reference [21] addresses the sensor selection problem wherein the dynamics is modeled using a partial differential equation. For systems in which one can excite the initial state in different directions, there is a rich literature on the use of empirical Gramians that are constructed using the trajectories of outputs [22], [23]. Reference [24] uses these empirical Gramians to perform sensor placement, but the problem of optimal choice of sensors has not been solved with theoretical guarantees. One limitation of existing works is that, to prove theoretical properties on (sub)-optimality of greedy approaches, one requires a submodularity property [9], which is not satisfied by the metrics relevant to sensor selection such as the ones considered in this

Manuscript received February 24, 2019; accepted June 25, 2019. Date of publication July 12, 2019; date of current version July 24, 2019. This letter was recommended for publication by Associate Editor U. Frese and Editor C. Stachniss upon evaluation of the reviewers' comments. This work was supported by National Science Foundation (NSF) under Grants ECCS 1446793 and IIS 1848945. *(Corresponding author: Shaunak D. Bopardikar.)*

The authors are with the Electrical and Computer Engineering Department, Michigan State University, East Lansing, MI 48824 USA (e-mail: shaunak@egr.msu.edu; ennasros@msu.edu; xbtan@egr.msu.edu).

This letter has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors. This video illustrates the performance of the randomized sampling approach (Algorithm 1) proposed in the letter on a target localization problem using time-difference-of-arrival sensors.

Digital Object Identifier 10.1109/LRA.2019.2928208

letter. Reference [25] has applied submodular optimization for selecting input nodes for consensus of networks with negative edges to maximize the minimum eigenvalue of the graph Laplacian, which is non submodular.

The randomization viewpoint has been used extensively to solve complex control design problems over the past two decades; see [26], [27] and [28], to name a few. These techniques adopt a sampling-based approach and provide probabilistic bounds on the resulting controller performance. In contrast, our approach leverages insights and analysis techniques from the field of random matrix theory. We refer the interested reader to the tutorial letter [29] for a detailed survey in the field.

B. Contributions

Given a nonlinear dynamical system, we consider the problem of selecting a subset of the sensors out of the total set in order to guarantee observability of the resulting system. We first formalize the performance of a simple randomized scheme based on *uniform sampling without replacement* and derive high-probability lower bounds on the minimum eigenvalue and the inverse of the condition number of the observability Gramian. The lower bounds are dependent on the state of the system and are relative to the *global value* obtained by using *all* of the available sensors. This is a key distinction with respect to the literature on sensor selection where the performance has largely been characterized relative to the optimal (but difficult to compute) subset of a given cardinality. The difference with respect to our recent work [15] is that the latter work relies on sampling with replacement, which cannot guarantee that a sensor does not get chosen more than once, and therefore, that work is more appropriate for the sensor placement problem.

We then illustrate the approach and the results in an example of tracking a moving target using range sensors and TDOA sensors, respectively. In particular, the location of the target, i.e., the state of the target dynamics, is estimated with an Extended Kalman Filter (EKF) based on the sensor measurements. We present studies that provide insight into the range of problem parameters (e.g., the size of the domain, the total number of sensors used, etc.) for which the lower bound on each of the considered Gramian metrics holds with high probability. In particular, *the utility of our theoretical results lies in the regime of large total number of sensors wherein the combinatorial nature of the problem presents a significant computational challenge*. We compare the steady-state estimation error covariance of the EKF applied to the subset of sensors to that of the EKF applied to all of the sensors, and observe a graceful degradation of performance with respect to the number of sensors selected. Finally, we provide a comparison of the empirical performance of our approach with that of a heuristic, which selects the sensors in a greedy manner based on the contribution of an additional sensor toward the observability Gramian metrics, and an improvement with respect to certain metrics is observed.

C. Organization of This Paper

This letter is organized as follows. Section II provides a formal statement of the problem and reviews an existing result from randomized Gramian computation. Section III formalizes our algorithm and includes the main theoretical results and their mathematical proofs. Section IV describes the application of the proposed method for localization and presents numerical results of the proposed approach. Finally, Section V summarizes our findings and identifies directions for future research.

II. Problem Formulation and Background

In this section, we present the mathematical description of the sensor selection problem considered in this letter and provide some background information on observability of nonlinear dynamical systems.

A. Model

Consider a network of sensors placed in an environment. The state of the underlying system to be estimated at time t is denoted by $x_t \in \mathcal{X} \subseteq \mathbb{R}^n$, whose evolution is modeled as

$$\begin{aligned} x_{t+1} &= f(x_t) + w_t, \\ y_{i,t} &= h_i(x_t) + v_{i,t}, \end{aligned} \tag{1}$$

where $y_{i,t} \in \mathbb{R}^m$, denotes the vector of measurements from the i-th sensor at time t, $w_t \sim \mathcal{N}(0, Q), v_{i,t} \sim \mathcal{N}(0, R)$ represent the process noise and the measurement noise, respectively. We assume that there are M sensors in total.

B. Nonlinear Observability

Given model (1), one can define the observability matrix, $O := \begin{bmatrix} O_1^T & O_2^T & \dots & O_M^T \end{bmatrix}^T$, where

$$O_i := \begin{bmatrix} \nabla L_f^0(h_i) \\ \nabla L_f^1(h_i) \\ \vdots \\ \nabla L_f^{d_i}(h_i) \end{bmatrix}, \tag{2}$$

for each $i \in \{1, \dots, M\}$, and d_i is the smallest positive integer for which $\text{rank}(O_i) = n$. The operator

$$L_f^k(h_i) = (\nabla L_f^{k-1}(h_i))^T f, \qquad \text{with } L_f^0(h_i) = h_i.$$

Here, ∇ denotes the partial derivative with respect to x. If (1) is linear time-invariant, then O_i reduces to the standard time-invariant observability matrix.

For ease of exposition, we assume a homogeneous sensor network (i.e., sensors with identical attributes), $d_1 = \dots = d_M = d$. It can be checked that for $M \geq 4$, with the choice of $d_i = 1$, the matrix O has full rank in $n = 2$ dimensions. Note that O_i is a function of the state x, since h_i is a nonlinear function of x, which implies that O is also a function of x.

Given a subset, $S = \{i_1, i_2, \dots, i_N\} \subseteq \{1, \dots, M\}$, where $N \leq M$, the observability matrix for the set S is

$$O_S(x_t) := \begin{bmatrix} O_{i_1}(x_t) \\ O_{i_2}(x_t) \\ \vdots \\ O_{i_N}(x_t) \end{bmatrix}.$$

C. Observability Metrics

This letter analyzes two metrics for observability of system (1), namely the minimum eigenvalue $\lambda_{\min}(\cdot)$ of the observability Gramian and the inverse of the condition number of the observability Gramian, i.e.,

$$\mu(O_S(x_t)) := \frac{\sigma_{\min}(O_S(x_t))}{\sigma_{\max}(O_S(x_t))} = \sqrt{\frac{\lambda_{\min}(W_S)}{\lambda_{\max}(W_S)}},$$

where $\sigma_{\min}(\cdot)/\sigma_{\max}(\cdot)$ are the minimum / maximum singular values, $\lambda_{\min}(\cdot)/\lambda_{\max}(\cdot)$ are the minimum / maximum eigenvalues, and $W_S := O_S^T O_S$. Since the condition number is a function of x_t, it is unreasonable to expect that one can find a sensor subset S such that for a given $\alpha \in [0,1)$,

$$\mu(O_S(x)) \geq \alpha, \quad \forall x \in \mathcal{X}.$$

Instead, what we can expect is, given *any* $x \in \mathcal{X}$ in the domain of interest, we would like to select $S \subseteq \{1,\dots,M\}$ so that

$$\mu(O_S(x)) \geq \alpha \mu(O(x)). \tag{3}$$

In other words, can we find a subset S of the sensors that will function at least within a certain fraction of the observability using *all* of the sensors? Even this formulation is a combinatorial optimization problem and, without any additional assumptions on the sensing models, it is computationally complex to solve (3). In what follows, we present an approach based on randomized selection that can provide probabilistic guarantees on a solution to (3).

III. Randomized Algorithm and its Analysis

This section formalizes a randomized sampling approach and presents the theoretical properties along with formal proofs of the claims.

A. Algorithm and Results

We propose a very simple randomized selection algorithm without replacement. The key idea is to select the first sensor uniformly randomly out of the M sensors. Then select the second sensor uniformly randomly out of the remaining $M-1$, and so on. The result is a randomly chosen subset $S = \{i_1,\dots,i_N\}$, for which

$$\sigma_{\min}(O_S) = \sqrt{\lambda_{\min}(W_S)},\ \mu(O_S) = \sqrt{\frac{\lambda_{\min}(W_S)}{\lambda_{\max}(W_S)}}.$$

This procedure is formalized in Algorithm 1.

While Algorithm 1 is very simple and does not seem to make use of the distinction between sensors, its major advantage is that we obtain the following probabilistic guarantee on its performance.

Theorem III.1: For *any* given $x \in \mathcal{X}$, let $B(x) := \max_{i\in\{1,\dots,M\}} \lambda_{\max}(W_i(x))$ and $W(x) := \sum_{i=1}^{M} W_i(x)$. Then, with Algorithm 1, the matrix O_S satisfies the following guarantees for any given $\epsilon \in (0,1)$:

1) With the minimum eigenvalue metric,

$$\mathbb{P}\left(\lambda_{\min}(O_S(x)) \geq (1-\epsilon)\lambda_{\min}(O(x))\right) \geq 1 - n\left(\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}}\right)^{\frac{N\lambda_{\min}(W(x))}{MB(x)}}.$$

Algorithm 1: Randomized Sensor Selection.

1: **Input:** Matrices $O_i, \forall i \in \{1,\dots,M\}$ defined in (2), number N of samples of sensors.
2: Initialize $S = \emptyset$.
3: **while** $|S| \leq N$ **do**
4: Uniformly randomly choose $i \in \{1,\dots,M\} \setminus S$.
5: $S = S \cup \{i\}$.
6: **end while**
7: **Output:** The matrix $O_S \in \mathbb{R}^{Nd\times n}$.

2) With the inverse condition number metric,

$$\mathbb{P}\left(\mu(O_S(x)) \geq \sqrt{\frac{1-\epsilon}{1+\epsilon}}\mu(O(x))\right) \geq 1 - n\left(\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}}\right)^{\frac{N\lambda_{\min}(W(x))}{MB(x)}} - n\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{N\lambda_{\max}(W(x))}{B(x)M}}. \tag{4}$$

We highlight a few remarks about this result below.

Remark III.2 (Role of ϵ): The choice of the accuracy parameter $\epsilon \in (0,1)$ is arbitrary and therefore, can be treated as *user defined*. If we seek to make ϵ closer to 1, then the probability of achieving (3) increases, but the lower bound becomes conservative in practice, and vice versa when ϵ is chosen closer to 0. Further discussion on this aspect is provided in the numerical studies presented in Section IV.

Remark III.3 (Dependence on x): Theorem III.1 holds for *any fixed* value of $x \in \mathcal{X}$. This fact is reflected in the value of the parameter B which is dependent on x and in particular, on the ratio $\lambda_{\min}(W(x))/B(x)$, and therefore, the probability that Problem 3 is solved will be different for each x. To obtain a uniform probability, one needs to take a supremum over all allowable $x \in \mathcal{X}$, which may lead to conservativeness in the number of sensors N required to achieve a desired accuracy.

Remark III.4 (Sensor fraction): The probabilistic guarantee of the approach depends on the fraction N/M of sensors selected and is quantified in the level of probabilistic guarantee. The fraction N/M does not affect the lower bound on $\mu(O_S(x))$, which depends only on the choice of ϵ. This observation will be evident in the numerical studies.

Theorem III.1 immediately yields the following sample complexity bounds for N.

Corollary III.5 (Sample complexity bounds): For any $\delta \in (0,1)$, Algorithm 1 solves Problem (3) with probability $1-\delta$ if, for the minimum eigenvalue metric,

$$N \geq \frac{M}{\epsilon + (1-\epsilon)\ln(1-\epsilon)} \ln\frac{n}{\delta} \sup_{x\in\mathcal{X}} \frac{B(x)}{\lambda_{\min}(W(x))},$$

and for the inverse condition number as metric,

$$N \geq \frac{M}{-\epsilon + (1+\epsilon)\ln(1+\epsilon)} \ln\frac{2n}{\delta} \sup_{x\in\mathcal{X}} \frac{B(x)}{\lambda_{\min}(W(x))}.$$

If N satisfies the bound in Corollary III.5, then the probabilistic guarantee in Theorem III.1 is satisfied for any operating point x. The result does not say anything about the probability that for *all* $x \in \mathcal{X}$, the matrices $O_S(x)$ simultaneously are as desired. It now remains to be seen how useful this bound will turn out

to be for specific sensor and motion models in applications of interest and for what problem parameter values these bounds lead to non-trivial values. This will be the focus of Section IV.

As will be evident in the proof of Theorem III.1, to the best of our knowledge, the analysis tools provide a compact bound on the probability of satisfaction only for the case of uniform sampling. Extension of these tools to the case of non-uniform selection probabilities is a topic of current research and requires a non-trivial extension of a key result in randomized matrix approximation.

B. Proofs of the Theoretical Claims

We begin with a useful result from randomized matrix approximation to conclude about properties of the extreme eigenvalues of sums of random matrices. Most matrix concentration inequalities, e.g., Ahlswede-Winter or Bernstein [29], are valid for sampling *with replacement*. The inequality below is a technique originally by Gross and Nesme, based on Hoeffding-like inequalities, for transferring results from sampling with replacement to sampling without replacement.

Theorem III.6: (Matrix Chernoff [30]). Let $\mathcal{X}$ denote a finite set of symmetric positive semi-definite matrices with dimension n and suppose that

$$\max_{X\in\mathcal{X}} \lambda_{\max}(X) \leq B.$$

Sample c matrices $X_1, X_2, \ldots, X_c$, uniformly at random from $\mathcal{X}$ and without replacement. Compute,

$$\mu_{\min} := c\lambda_{\min}(\mathbb{E}[X_1]), \text{ and } \mu_{\max} := c\lambda_{\max}(\mathbb{E}[X_1]).$$

Then, for any $\epsilon \in [0, 1)$,

$$\mathbb{P}\left(\lambda_{\min}\left(\sum_{j=1}^{c} X_j\right) \leq (1-\epsilon)\mu_{\min}\right) \leq n\left(\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}}\right)^{\frac{\mu_{\min}}{B}},$$

and for any $\epsilon \geq 0$,

$$\mathbb{P}\left(\lambda_{\max}\left(\sum_{j=1}^{c} X_j\right) \geq (1+\epsilon)\mu_{\max}\right) \leq n\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{\mu_{\max}}{B}}.$$

We now present the proof of Theorem III.1.

Proof of Theorem III.1: For the sake of brevity, we drop the dependence of x in W_S. We begin with

$$W_S = O_S^T O_S = \sum_{j=1}^{N} O_{i_j}^T O_{i_j}$$

where $W_j = O_{i_j}^T O_{i_j} \in \mathbb{R}^{n\times n}$ is a randomly chosen matrix out of the set of matrices $\mathcal{X} := \{W_1, \ldots, W_M\} \setminus \{W_{i_1}, \ldots, W_{i_{j-1}}\}$.

Using Theorem III.6 for the matrices in $O_S^T O_S$, we have

$$\mu_{\min} = N\lambda_{\min}\left(\frac{1}{M}W\right) = \frac{N}{M}\lambda_{\min}(W),$$

$$\mu_{\max} = N\lambda_{\max}\left(\frac{1}{M}W\right) = \frac{N}{M}\lambda_{\max}(W).$$

The first claim now follows directly by applying the inequality for $\lambda_{\min}$. For the second claim, we use Boole's inequality (union bound) and conclude that

$$\mathbb{P}\left(\lambda_{\min}\left(\sum_{j=1}^{N} O_{i_j}^T O_{i_j}\right) > (1-\epsilon)\frac{N}{M}\lambda_{\min}(W)\right.$$

$$\left.\bigcap \lambda_{\max}\left(\sum_{j=1}^{N} O_{i_j}^T O_{i_j}\right) < (1+\epsilon)\frac{N}{M}\lambda_{\max}(W)\right) \geq 1$$

$$- n\left(\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}}\right)^{\frac{N\lambda_{\min}(W)}{MB}} - n\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{N\lambda_{\max}(W)}{BM}},$$

for any $\epsilon \in (0, 1)$. The claim now follows from algebra. ■

We now present the proof of Corollary III.5.

Proof of Corollary III.5: The first claim follows by imposing the requirement that

$$n\left(\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}}\right)^{\frac{N\lambda_{\min}(W(x))}{MB(x)}} = \delta$$

$$\Leftrightarrow \ln\frac{n}{\delta} = \frac{N\lambda_{\min}(W(x))}{MB(x)}\ln\left(\frac{(1-\epsilon)^{1-\epsilon}}{\mathbf{e}^{-\epsilon}}\right)$$

$$\Leftrightarrow \ln\frac{n}{\delta} = \frac{N\lambda_{\min}(W(x))}{MB(x)}(\epsilon + (1-\epsilon)\ln(1-\epsilon)),$$

which, upon taking the supremum over x, yields the desired sample complexity bound.

For the second claim, observe that

$$\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}} \leq \frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}$$

$$\Rightarrow -\left(\frac{\mathbf{e}^{-\epsilon}}{(1-\epsilon)^{1-\epsilon}}\right)^{\frac{N\lambda_{\min}(W)}{MB}} \geq -\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{N\lambda_{\min}(W)}{MB}}$$

from which it follows that the right hand side of (4) is lower bounded by

$$1 - n\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{N\lambda_{\min}(W)}{MB}} - n\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{N\lambda_{\max}(W)}{BM}}$$

$$\geq 1 - 2n\left(\frac{\mathbf{e}^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\frac{N\lambda_{\min}(W)}{MB}},$$

where we used the fact that $\lambda_{\min}(\cdot) \leq \lambda_{\max}(\cdot)$. Imposing the requirement that the second term equals δ and upon following similar algebra as in the proof of the first claim, we obtain the following sufficient condition on N,

$$N \geq \frac{M}{-\epsilon + (1+\epsilon)\ln(1+\epsilon)}\ln\frac{2n}{\delta}\frac{B(x)}{\lambda_{\min}(W(x))}.$$

The claim now follows by taking a supremum over x. ■

IV. Application to Target Localization

In this section, we numerically study the effectiveness of Algorithm 1 with nonlinear observation functions in two scenarios. The first scenario includes the use of range sensors while the second scenario uses TDOA sensors. In both scenarios, we compare the estimation performance of the resulting EKF with

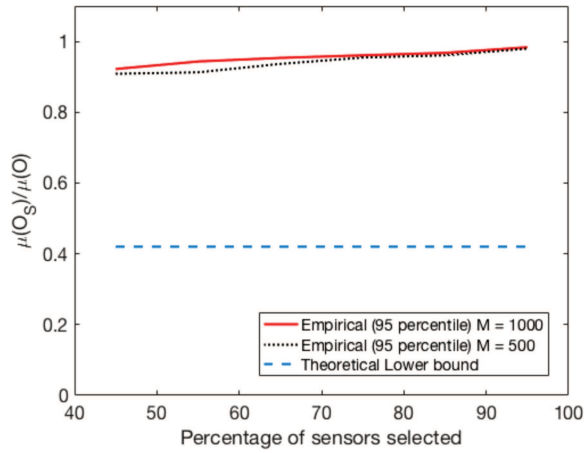

Fig. 1. The ratio, $\frac{\mu(O_S)}{\mu(O)}$ for different sizes of S for the range sensors scenario. The empirical curve corresponds to the 95 percentile, i.e., only 5% of the realizations lie below this curve.

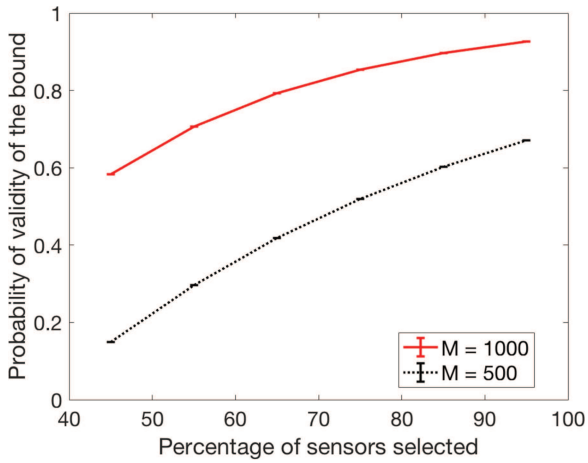

Fig. 2. Probabilistic lower bound from Theorem III.1 for the range sensors scenario. The error bars indicate ± 1 standard deviation.

that of using all of the sensors. Without any loss of generality, we consider a linear motion model in the examples, which is inspired out of similar motion models for mobile targets considered in [19], [20].

A. Range Sensors

We first study the probabilistic guarantees provided by Algorithm 1. In this numerical study, we assume a simple linear model for the state evolution, i.e.,

$$x_{t+1} = f(x_t) = Ax_t + w_t,$$

with A equal to the identity within a square environment, i.e., $\mathcal{X} \in [0,1] \times [0,1]$. The observation model is assumed to be the square of the distance between the state x_t and the location of the range sensor $x^i \in \mathcal{X}, \forall i \in \{1, \ldots, M\}$, i.e.,

$$y_i(x) = h_i(x) + v_{i,t} = (x^i - x)^T (x^i - x) + v_{i,t}.$$

The observability matrix for the i-th sensor is given by

$$O_i = \begin{bmatrix} \nabla h_i \\ \nabla(\nabla h_i^T f) \end{bmatrix} = 2 \begin{bmatrix} x - x^i \\ (2x - x^i)^T A \end{bmatrix}.$$

We assume that the ranging sensors are located uniformly randomly in $\mathcal{X}$. Applying the randomized sampling scheme, we obtain a sensor set S whose inverse condition number has been compared empirically with that of the entire set in Figures 1 and 2. To generate one value of $\frac{\mu(O_S)}{\mu(O)}$, we evaluated 100 Monte Carlo realizations of the sampling algorithm each yielding a different set S. As highlighted in Remark III.4, the curve for the theoretical lower bound $\sqrt{(1-\epsilon)/(1+\epsilon)}$ in Figure 1 is flat because it depends only on the choice of ϵ, which was fixed to 0.7 in these simulations. We observe that for a higher value of M, the accuracy plot remains approximately similar, but the probabilistic lower bound from Theorem III.1 improves significantly, suggesting that this simple sampling scheme provides useful high probability guarantees when the density of sensors is high in a region, such as in applications that involve networks of inexpensive sensors for indoor localization. As discussed in Remark III.2, since the choice of ϵ is arbitrary, if we were to decrease ϵ further, then the lower bound in Figure 1 will increase thereby making the approximation more accurate, but at the expense of a weaker bound on the probabilistic guarantee.

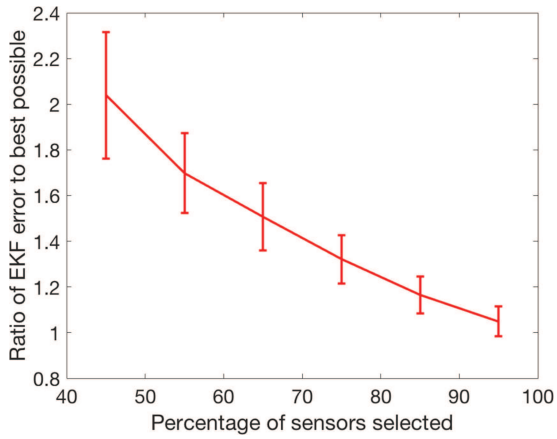

Fig. 3. The ratio, $\frac{\lambda_{\max}(P_{S,T})}{\lambda_{\max}(P_T)}$ for different sizes of S for the range sensors scenario. The error bars indicate ± 1 standard deviation. In this simulation, total number of sensors $M = 100$, the number of steps $T = 10$ and 20 Monte Carlo runs of Algorithm 1 were carried out for every value of N.

We now present the results of application of Algorithm 1 to state estimation using an EKF. If $P_{S,t}$ denotes the estimation error covariance of the EKF using the sensor set S at time t, then the performance of an EKF is measured using $\lambda_{\max}(P_{S,T})$, where T is the estimation horizon. We compared the performance of the EKF using the set S computed using Algorithm 1 to that of the EKF using all of the M sensors in terms of the ratio $\lambda_{\max}(P_{S,T})/\lambda_{\max}(P_T)$.

The trend is summarized in Figure 3, wherein a lower value on the y-axis indicates better performance. We observe that the ratio improves with increasing N as is expected. In particular, with 45% of the total number of sensors, we obtain only a factor of 2 degradation in performance.

We finally compared the performance of the EKF using the set S computed out of Algorithm 1 to that of the EKF using a *greedy algorithm* described as follows:

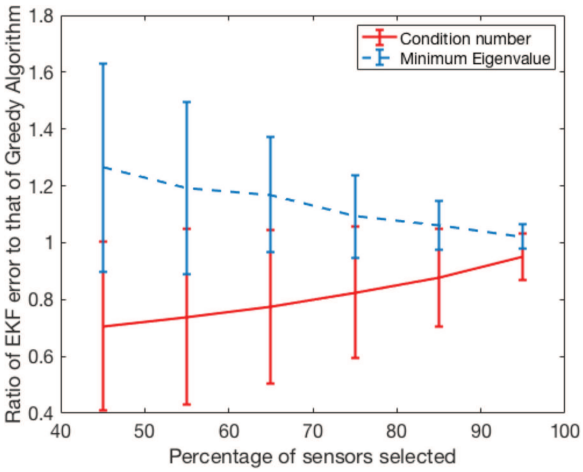

Fig. 4. The ratio, $\frac{\lambda_{\max}(P_{S,T})}{\lambda_{\max}(P_{G,T})}$ for different sizes of S for the range sensors scenario. Here, $P_{G,T}$ denotes the estimation error covariance of the EKF using the output of the greedy algorithm. A ratio less than unity indicates improvement performance of Algorithm 1 over the greedy heuristic. The error bars indicate ± 1 standard deviation. In this simulation, total number of sensors $M = 100$, the number of steps $T = 10$ and 20 Monte Carlo runs of Algorithm 1 were carried out for every value of x and 20 Monte Carlo runs for every value of N.

1) Start with a set of 3 arbitrarily chosen sensors and a fixed value of x such as the initial estimate of the target location.
2) Select the next sensor that maximizes the metric if that sensor is to be added to the existing subset.
3) Continue until the number of sensors equals N.

The greedy algorithm terminates in finite number of iterations, but does not possess any known theoretical guarantees. We used the same number of sensors in our proposed randomized approach. The results are summarized in Figure 4 using both metrics considered in this letter. The choice of the initial target position x is selected uniformly randomly from the environment. We observe that the EKF error is comparable between the two approaches and in particular, using the condition number metric, our approach provides an improvement over the greedy algorithm. We conjecture that this may be attributed to the fact that the greedy algorithm relies greatly on the choice of the point x in the computation of the observability Gramian, whereas our sampling algorithm is agnostic to the choice of x. Therefore, over a trajectory, there may be target locations that lead to a higher estimation error on average for the greedy approach.

B. TDOA Sensors

We now present the results in a scenario that involves the use of TDOA measurements. The model is equivalent to range difference between two sensors when the signal speed is known. Given any two neighboring sensors i and j, the observation model is assumed to be the difference of the distance between the state x_t and the location of the sensor x^i and the distance between the x_t and the location of its neighbor $x^j, \forall j \in \{1, \ldots, M\}$, i.e.,

$$y_{i,j}(x) = h_{i,j}(x) + v_{ij,t} = \|x^i - x\| - \|x^j - x\| + v_{ij,t}.$$

In this approach, a sensor is used as a *reference node*, and the range differences are computed with respect to that reference node. If sensor i is used as the reference, the nonlinear observability matrix for its j-th neighbor is given by

$$O_{i,j} = \begin{bmatrix} \nabla h_{i,j} \\ \nabla(\nabla h_{i,j}^T f) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{x-x^i}{\|x-x^i\|} - \frac{x-x^j}{\|x-x^j\|} \\ \frac{A(2x-x^i)}{\|x-x^i\|} - \frac{(x-x^i)Ax(x-x^i)}{\|x-x^i\|^3} - \frac{A(2x-x^j)}{\|x-x^j\|} + \frac{(x-x^j)Ax(x-x^j)}{\|x-x^j\|^3} \end{bmatrix}.$$

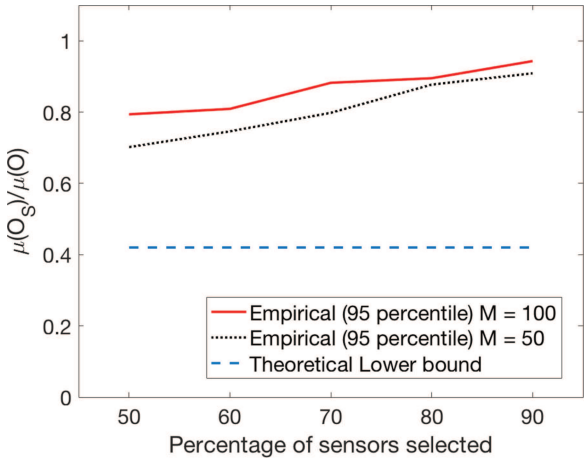

Fig. 5. The ratio, $\frac{\mu(O_S)}{\mu(O)}$ for different sizes of S using the TDOA sensors. The empirical curve corresponds to the 95 percentile, i.e., only 5% of the realizations lie below this curve.

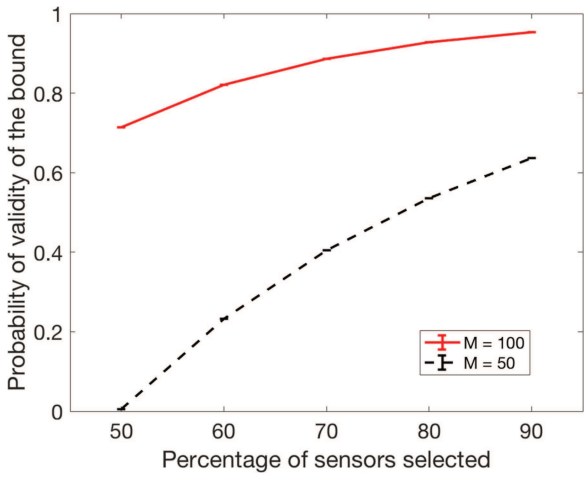

Fig. 6. Probability of validity of the analytic bound using TDOA sensors. The error bars indicate ± 1 standard deviation.

Akin to the previous subsection, we first present the probabilistic guarantees similar to those for the ranging sensors, as summarized in Figures 5 and 6. In this setup, we assumed a circular arrangement for the TDOA sensors (cf. Figure 8), i.e., the sensors are placed on the vertices of a regular polygon of radius equal to 100 units, with one of the sensors identified as the reference sensor. The numerical procedure is identical to that for the range sensors and with $\epsilon = 0.7$. These results also suggest a similar conclusion as in the case of the range sensors that the

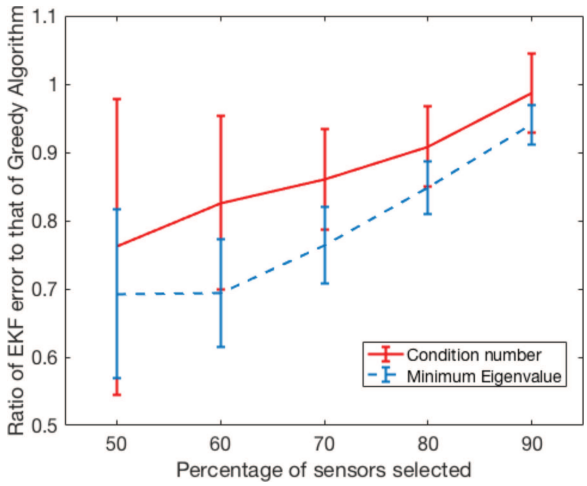

Fig. 7. The ratio, $\frac{\lambda_{\max}(P_{S,T})}{\lambda_{\max}(P_{G,T})}$ for different sizes of S for the TDOA sensors. The error bars indicate ± 1 standard deviation. In this simulation, total number of sensors $M = 100$, the number of steps $T = 10$ and 20 Monte Carlo runs of Algorithm 1 were carried out for every value of x and 20 Monte Carlo runs for every value of N.

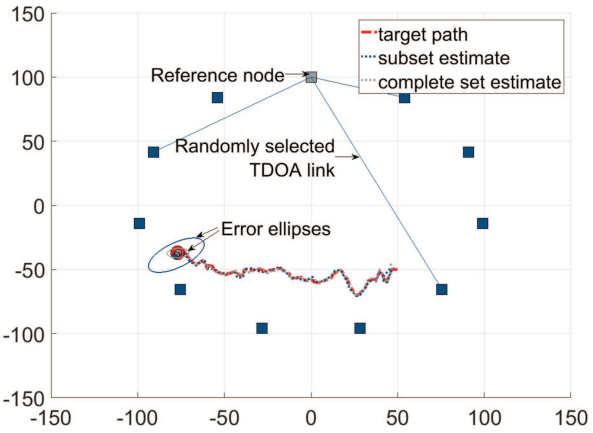

Fig. 8. Simulation setup for evaluation of Algorithm 1 using TDOA measurements. Blue squares represent the available sensors, while the gray square represents the reference node. Straight lines between square indicates that the TDOA measurements is part of the subset selected by Algorithm 1. Error ellipses are plotted around the target (red circle), where the blue ellipse represents the confidence bound using the subset, while the gray ellipse shows the confidence bound using the entire set of measurement.

sampling scheme provides useful high probability guarantees when the density of sensors is high in a region.

We also compared the performance of Algorithm 1 to that of the greedy algorithm outlined in the previous sub-section. The results are summarized in Figure 7 and we see that our approach provides an improvement over the greedy algorithm, in both metrics over a wide range of the fraction of sensors. The choice of the initial target position x is selected uniformly randomly from the environment.

Finally, we present the results of application of Algorithm 1 to state estimation using an EKF. As before, the performance of an EKF is measured using $\lambda_{\max}(P_{S,T})$. The simulation setup is shown in Figure 8 where the estimation performance using a subset of the measurements is compared to that using the

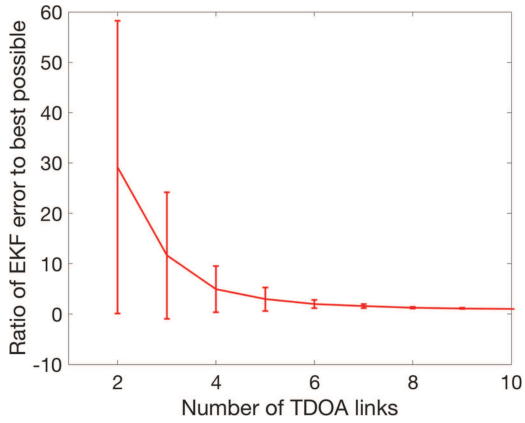

Fig. 9. The ratio, $\frac{\lambda_{\max}(P_{S,T})}{\lambda_{\max}(P_T)}$ for different sizes of S for TDOA measurements. The errorbars indicate ± 1 standard deviation.

full set. The total number of sensors used in the simulation is 11, resulting in 10 range-difference links for the reference node. Algorithm 1 was applied to randomly select a subset of the available range-difference links. Two error ellipses are shown in Figure 8. As expected, the error ellipse for using the complete set of range-differences is tighter than that using a subset of measurements. A video illustrating this approach has been uploaded as an attachment to this submission.

For each size of S, the process was repeated 100 times before increasing the size of the set by one. We compared the performance of the EKF using the set S computed using Algorithm 1 to that of the EKF using all of the M sensors in terms of the ratio $\frac{\lambda_{\max}(P_{S,T})}{\lambda_{\max}(P_T)}$. The trend is summarized in Figure 9, wherein a lower value on the y-axis indicates better performance. We observe that the ratio improves with increasing N as is expected.

Localization using range-difference measurements heavily relies on the relative positions between target and sensors [18], [20]. This explains the large variations in the data in Figure 9 when only 2 or 3 TDOA links are used, as it becomes more likely to select an unfavorable configuration for TDOA localization. Further, we observe that the performance is fairly flat beyond 40 percent of the sensors used suggesting that the marginal improvement in the quality of the estimate decreases with a high confidence with increasing number of TDOA links. We also repeated this simulation with the choice of the trace of the estimation error covariance matrix as metric leading to the ratio, $\mathrm{Tr}(P_{S,T})/\mathrm{Tr}(P_T)$. We obtain qualitatively similar results with a large variance for 2 or 3 TDOA links, but with mean ratio much closer to one than in Figure 9. We skip the details in the interest of space.

V. Conclusion and Future Directions

This letter considered the problem of randomly sampling sensors out of the total set in order to provide probabilistic guarantees on the minimum eigenvalue and inverse condition number of the observability Gramian. We analyzed a simple randomized algorithm based on uniform sampling without replacement and derived novel high-probability lower bounds on the two metrics. These lower bounds are relative to the observability Gramian constructed out of all of the sensors. The utility of our theoretical results lies in the regime of large total number of sensors

thereby providing an alternative to the combinatorially complex formulation of the sensor selection problem. The empirical performance of proposed approach showed a graceful degradation of performance with respect to the use of an EKF with all of the available sensors. We also observe an improvement over a greedy heuristic for specific choice of the metrics, especially in the regime of low number of samples of sensors.

We are presently studying the case of sampling sensors with non-uniform probability. Longer term extensions include problems involving sensor scheduling for nonlinear systems, an improved analysis especially in the regime of low number of sensors, and characterization of the error of the randomized algorithm relative to the best set of a given cardinality. Also of interest are the applications to non-homogeneous sensor networks and applicability to specific sensors such as bearing along with a limited field of view.

REFERENCES

[1] G. E. Hovland, and B. J. McCarragher, "Dynamic sensor selection for robotic systems," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1997, vol. 1, pp. 272–277.

[2] J. R. Spletzer and C. J. Taylor, "Dynamic sensor planning and control for optimally tracking targets," *Int. J. Robot. Res.*, vol. 22, no. 1, pp. 7–20, 2003.

[3] V. Isler and M. Magdon-Ismail, "Sensor selection in arbitrary dimensions," *IEEE Trans. Autom. Sci. Eng.*, vol. 5, no. 4, pp. 651–660, Oct. 2008.

[4] O. Tekdas and V. Isler, "Sensor placement for triangulation-based localization," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 3, pp. 681–685, Jul. 2010.

[5] M. Imani and U. M. Braga-Neto, "Optimal finite-horizon sensor selection for Boolean Kalman filter," in *Proc. IEEE 51st Asilomar Conf. Signals, Syst., Comput.*, 2017, pp. 1481–1485.

[6] M. Anvaripour, M. Saif, and M. Ahmadi, "Novel sensor selection method for tracking objects and providing safe workplace," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1162–1167, 2018.

[7] M. Shimosaka, O. Saisho, T. Sunakawa, H. Koyasu, K. Maeda, and R. Kawajiri, "ZigBee based wireless indoor localization with sensor placement optimization towards practical home sensing," *Adv. Robot.*, vol. 30, no. 5, pp. 315–325, 2016.

[8] M. Van De Wal and B. De Jager, "A review of methods for input/output selection," *Automatica*, vol. 37, no. 4, pp. 487–510, 2001.

[9] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Trans. Autom. Control*, vol. 3, no. 1, pp. 91–101, Mar. 2016.

[10] S. Pequito, S. Kar, and A. P. Aguiar, "Minimum cost input/output design for large-scale linear structural systems," *Automatica*, vol. 68, pp. 384–391, Jun. 2016.

[11] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms," in *Proc. IEEE Amer. Control Conf.*, 2016, pp. 191–196.

[12] S. T. Jawaid and S. L. Smith, "Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems," *Automatica*, vol. 61, pp. 282–288, Nov. 2015.

[13] H. Zhang, R. Ayoub, and S. Sundaram, "Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms," *Automatica*, vol. 78, pp. 202–210, Apr. 2017.

[14] A. Hashemi, M. Ghasemi, H. Vikalo, and U. Topcu, "A randomized greedy algorithm for near-optimal sensor scheduling in large-scale sensor networks," in *Proc. IEEE Annu. Amer. Control Conf.*, 2018, pp. 1027–1032.

[15] S. D. Bopardikar, "Sensor selection via randomized sampling," 2018. [Online]. Available: https://arxiv.org/abs/1712.06511

[16] L. Zhou and P. Tokekar, "Sensor assignment algorithms to improve observability while tracking targets," in *IEEE Trans. Robot.*, to be published, 2019, doi: 10.1109/TRO.2019.2920749.

[17] R. K. Williams and G. S. Sukhatme, "Observability in topology-constrained multi-robot target tracking," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2015, pp. 1795–1801.

[18] B. Yang and J. Scheuing, "Cramer-Rao bound and optimum sensor array for source localization from time differences of arrival," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2005, vol. 4, pp. iv/961–iv/964.

[19] O. Ennasr, G. Xing, and X. Tan, "Distributed time-difference-of-arrival (TDOA)-based localization of a moving target," in *Proc. IEEE 55th Conf. Decis. Control*, 2016, pp. 2652–2658.

[20] O. N. Ennasr and X. Tan, "Distributed estimation and tracking using time-difference-of-arrival (TDOA) measurements," in *Proc. ASME Dyn. Syst. Control Conf.*, 2017, Paper no. V002T14A011.

[21] S. Sinha, U. Vaidya, and R. Rajaram, "Operator theoretic framework for optimal placement of sensors and actuators for control of nonequilibrium dynamics," *J. Math. Anal. Appl.*, vol. 440, no. 2, pp. 750–772, 2016.

[22] S. Lall, J. E. Marsden, and S. Glavaški, "A subspace approach to balanced truncation for model reduction of nonlinear control systems," *Int. J. Robust Nonlinear Control, IFAC-Affiliated J.*, vol. 12, no. 6, pp. 519–535, 2002.

[23] N. D. Powel and K. A. Morgansen, "Empirical observability Gramian rank condition for weak observability of nonlinear systems with control," in *Proc. IEEE 54th Annu. Conf. Decis. Control*, 2015, pp. 6342–6348.

[24] J. Qi, K. Sun, and W. Kang, "Optimal PMU placement for power system dynamic state estimation by using empirical observability Gramian," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2041–2054, Jul. 2015.

[25] A. Clark, Q. Hou, L. Bushnell, and R. Poovendran, "Maximizing the smallest eigenvalue of a symmetric matrix: A submodular optimization approach," *Automatica*, vol. 95, pp. 446–454, 2018.

[26] R. Tempo, E. W. Bai, and F. Dabbene, "Probabilistic robustness analysis: Explicit bounds for the minimum number of samples," *Syst. Control Lett.*, vol. 30, no. 5, pp. 237–242, 1997.

[27] M. Vidyasagar, "Statistical learning theory and randomized algorithms for control," *IEEE Control Syst. Mag.*, vol. 18, no. 6, pp. 69–85, Dec. 1998.

[28] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 742–753, May 2006.

[29] J. A. Tropp, "User-friendly tail bounds for sums of random matrices," *Found. Comput. Math.*, vol. 12, no. 4, pp. 389–434, Jan. 2011.

[30] J. A. Tropp, "Improved analysis of the subsampled randomized Hadamard transform," *Adv. Adaptive Data Anal.*, vol. 3, pp. 115–126, 2011.