Data-Driven Sparse Sensor Placement

Krithika Manohar\textsuperscript{1}, Bingni W. Brunton\textsuperscript{2}, J. Nathan Kutz\textsuperscript{1}, Steven L. Brunton\textsuperscript{3}

\textsuperscript{1} Department of Applied Mathematics, University of Washington, Seattle, WA 98195, United States
\textsuperscript{2} Department of Biology, University of Washington, Seattle, WA 98195, United States
\textsuperscript{3} Department of Mechanical Engineering, University of Washington, Seattle, WA 98195, United States

Abstract

Sensor placement or selection is one of the most challenging unsolved problems in engineering. Optimizing sensor placement to characterize a high-dimensional system is critical for the downstream tasks of prediction, estimation, and control. The central focus of this paper is to select the best and fewest measurement locations for reconstruction in a given basis. This goal implicitly assume that the high-dimensional state has a compact representation in some known basis, and this inherent compressibility enables sparse sensing. We compare two leading perspectives: 1) compressed sensing with random measurements in a universal basis (Fourier, wavelets, etc.), and 2) optimized measurements in a tailored basis, leveraging the ability of machine learning to exploit patterns in data. We synthesize concepts from sensor selection and model reduction to find the fewest sensors for recovery in a data-driven proper orthogonal decomposition (POD) basis. Optimized sensors are efficiently chosen using QR column pivots of the POD basis. Sensors obtained through this method are shown to surpass the performance of random sensors in the reconstruction of high-dimensional states in several compelling real-world datasets. Reducing the number of sensors through principled selection may be critically enabling when sensors are costly, and may also enable faster state estimation for low latency, high bandwidth control. MATLAB code is provided for all examples.

Notation

- **Scalars**
  - $n$: State dimension
  - $m$: Number of snapshots
  - $p$: Number of sensors (measurements)
  - $r$: Intrinsic rank of tailored basis $\Psi_r$
  - $K$: Sparsity of state in universal basis $\Psi$

- **Vectors**
  - $x \in \mathbb{R}^n$: High-dimensional state
  - $y \in \mathbb{R}^p$: Measurements of state
  - $a \in \mathbb{R}^r$: Tailored basis coefficients
  - $s \in \mathbb{R}^n$: $K$-sparse basis coefficients
  - $\psi \in \mathbb{R}^n$: POD modes (columns of $\Psi_r$)

- **Matrices**
  - $C \in \mathbb{R}^{p \times n}$: Measurement matrix
  - $\Psi \in \mathbb{R}^{n \times n}$: Universal basis
  - $\Psi_r \in \mathbb{R}^{n \times r}$: Tailored basis of rank $r$
  - $\Theta = C\Psi$: Product of measurement and basis
  - $X \in \mathbb{R}^{n \times m}$: Data matrix
  - $\Sigma$: Diagonal matrix of singular values
  - $\mathbf{V}$: Matrix of right singular vectors

Optimal sensor and actuator placement is an important unsolved problem in control theory. Nearly every downstream control decision is affected by these sensor/actuator locations, but determining optimal locations amounts to an intractable brute force search among the combinatorial possibilities. Indeed, there are $\binom{n}{p} \approx \frac{n!}{(n-p)!p!}$ possible choices of $p$ point sensors out of an $n$ dimensional state $x$. Determining optimal sensor and actuator placement in general, even for linear feedback control, is an open challenge. Instead, sensor and actuator locations are routinely chosen according to heuristics and intuition. This paper discusses how to design optimal sensor locations for signal reconstruction in a framework that scales to arbitrarily large problems, leveraging modern techniques in machine learning and sparse sampling. Extensions to sensor and actuator optimization for closed-loop control are also discussed. The overarching sensor placement problem is summarized in Panel 1.
Panel 1: Mathematical formulation of sensor selection

Many physical systems are described by a high-dimensional state \( x \in \mathbb{R}^n \), yet the dynamics evolve on a low-dimensional attractor that can be leveraged for prediction and control. Thus, a state \( x \) that evolves according to nonlinear dynamics

\[
\dot{x}(t) = f(x(t))
\]

does not need to be sampled at a high rate. Here \( f \) is a nonlinear map from \( \mathbb{R}^n \) to \( \mathbb{R}^n \). In a universal basis \( \Psi \in \mathbb{R}^{n \times n} \), such as Fourier or wavelet bases, \( x \) may have a sparse representation:

\[
x = \Psi s, \quad s \in \mathbb{R}^n,
\]

where \( s \) is a sparse vector indicating which few modes of \( \Psi \) are active. In a tailored basis \( \Psi_r \in \mathbb{R}^{n \times r} \), such as a proper orthogonal decomposition (POD) basis, \( x \) may have a low-rank representation:

\[
x = \Psi_r a, \quad a \in \mathbb{R}^r.
\]

The central challenge in this work is to design a measurement matrix \( C \in \mathbb{R}^{p \times n} \) consisting of a small number \((p \ll n)\) of optimized measurements

\[
y = Cx, \quad y \in \mathbb{R}^p,
\]

that facilitate accurate reconstruction of either \( s \) or \( a \), and hence \( x \). Combining (1) and (3) yields

\[
y = (C\Psi)s = \Theta s,
\]

which is referred to as the compressed sensing problem, while combining (2) and (3) yields

\[
y = (C\Psi_r)a = \Theta a.
\]

In both cases, effective measurements \( C \) given a basis \( \Psi \) or \( \Psi_r \) are chosen so that the operator \( \Theta \) is well-conditioned for signal reconstruction. Thus, it is possible to solve for the sparse coefficients \( s \) or the low-rank coefficients \( a \) given the measurements \( y \), either by \( \ell_1 \) minimization in (4) or pseudoinverse of \( \Theta \) in (5), respectively. The goal of this work is to optimize the measurements in \( C \). Moreover, in many physical applications, it is desired that \( C \) consists of rows of the identify matrix, corresponding to individual point sensors of individual components of \( x \).

There are myriad complex systems that would benefit from principled, scaleable sensor and actuator placement, including fluid flow control [1], power grid optimization [2, 3], epidemiological modeling and suppression [4, 5], bio-regulatory network monitoring and control [6, 7], and high-performance computing [8, 9], to name only a few. In applications where individual sensors are expensive, reducing the number of sensors through principled design may be critically enabling. In applications where fast decisions are required, such as in high performance computing or feedback control, computations may be accelerated by minimizing the number of sensors required. In other words, low-dimensional computations may be performed directly in the sensor space.

Scaleable optimization of sensor and actuator placement is a grand challenge problem, with tremendous potential impact and considerable mathematical depth. With existing mathematical machinery, optimal placement can only be determined in general using a brute-force combinatorial search. Although this approach has been successful in small-scale problems [10], a combinato-
rial search does not scale well to larger problems. Moore’s law of exponentially increasing computer power cannot keep pace with this combinatorial growth in complexity.

Despite the challenges of sensing and actuation in a high-dimensional, possibly nonlinear dynamical system, there are promising indicators that this problem may be tractable with modern techniques. High-dimensional systems, such as are found in fluids, epidemiology, neuroscience, and the power grid, typically exhibit dominant coherent structures that evolve on a low-dimensional attractor. Indeed, much of the success of modern machine learning rests on the ability to identify and take advantage of patterns and features in high-dimensional data. These low-dimensional patterns are often identified using dimensionality reduction techniques [11] such as the proper orthogonal decomposition (POD) [12, 13, 14], which is a variant of principal component analysis (PCA), or more recently via dynamic mode decomposition (DMD) [15, 16, 17, 18], diffusion maps [19, 20, 21, 22], etc. In control theory, balanced truncation [23], balanced proper orthogonal decomposition (BPOD) [24, 25], and the eigensystem realization algorithm (ERA) [26], have been successfully applied to obtain control-oriented reduced-order models for many high-dimensional systems.

In addition to advances in dimensionality reduction, key developments in optimization, compression, and the geometry of sparse vectors in high-dimensional spaces are providing powerful new techniques to obtain approximate solutions to NP-hard, combinatorially difficult problems in scalable convex optimization architectures. For example, compressed sensing [27, 28, 29] provides convex algorithms to solve the combinatorial sparse signal reconstruction problem with high probability. Ideas from compressed sensing have been used to determine the optimal sensor locations for categorical decisions based on high-dimensional data [30], including in fluid dynamics [31, 32]. Recently, compressed sensing, sparsity-promoting algorithms such as the lasso regression [33, 34, 35], and machine learning have been increasingly applied to characterize and control dynamical systems [36, 37, 38, 39, 40, 31, 41, 42, 43, 44, 45, 46, 47]. These techniques have been effective in modeling high dimensional fluid systems using POD [48] and DMD [49, 50, 51, 52]. Information criteria [53, 54, 55] has also been leveraged for the sparse identification of nonlinear dynamics [41], as in [56, 57], and may also be useful for sensor placement.

Thus, key advances in two fields are fundamentally changing our approach to the acquisition and analysis of data from complex dynamical systems: 1) machine learning, which exploits patterns in data for low-dimensional representations, and 2) sparse sampling, where a full signal can be reconstructed from a small subset of measurements. The combination of machine learning and sparse sampling is synergistic, in that underlying low-rank representations facilitate sparse measurements. Exploiting coherent structures underlying a large state space allows us to estimate and control systems with few measurements and sparse actuation. Low-dimensional, data-driven sensing and control is inspired in part by the high performance exhibited by biological organisms, such as an insect that performs robust, high-performance flight control in a turbulent fluid environment with minimal sensing and low-latency control [58]. They provide proof-by-existence that it is possible to assimilate sparse measurements and perform low-dimensional computations to interact with coherent structures in a high-dimensional system (i.e., a turbulent fluid).

Here we explore two competing perspectives on high-dimensional signal reconstruction: 1) the use of compressed sensing based on random measurements in a universal encoding basis, and 2) the use of highly specialized sensors for reconstruction in a tailored basis, such as POD or DMD. These choices are also discussed in the context of feedback control. Many competing factors impact control design, and a chief consideration is the latency in making a control decision, with large latency imposing limitations on robust performance [59, 60]. Thus, for systems with fast dynamics and complex coherent structures, it is important to make control decisions quickly based on efficient low-order models, with sensors and actuators placed strategically to gather information and exploit sensitivities in the dynamics.
1 Compressed sensing: Random measurements in a universal basis

The majority of natural signals, such as images and audio, are highly compressible, meaning that when the signal is written in an appropriate coordinate system, only a few basis modes are active. These few values corresponding to the large mode amplitudes must be stored for accurate reconstruction, providing a significant reduction compared to the original signal size. In other words, in the universal transform basis, the signal may be approximated by a \textit{sparse} vector containing mostly zeros. This inherent sparsity of natural signals is central to the mathematical framework of compressed sensing. Signal compression in the Fourier domain is illustrated on an image example in Panel 2.

The theory of compressed sensing \cite{1,2,3,4,5} inverts this compression paradigm. Instead of collecting high-dimensional measurements just to compress and discard most of the information, it may be possible to collect a low-dimensional subsample or compression of the data and then infer the sparse vector of coefficients in the transformed coordinate system.

1.1 Theory of compressed sensing

Mathematically, a compressible signal $x \in \mathbb{R}^n$ may be written as a sparse vector $s \in \mathbb{R}^n$ in a new basis given by $\Psi \in \mathbb{R}^{n \times n}$:

$$x = \Psi s.$$  \hspace{1cm} (6)

The vector $s$ is called $K$-sparse if there are exactly $K$ nonzero elements. To be able to represent any natural signal, rather than just those from a tailored category, the basis $\Psi$ must be complete.

Consider a set of measurements $y \in \mathbb{R}^p$, obtained via a measurement matrix $C \in \mathbb{R}^{p \times n}$:

$$y = Cx = C\Psi s = \Theta s.$$  \hspace{1cm} (7)

In general, for $p < n$ (7) is underdetermined, and there are infinitely many solutions. The least-squares (minimum $\|s\|_2$) solution is not sparse, and typically yields poor reconstruction. Instead, knowing that natural signals are sparse, we seek the sparsest $s$ consistent with the measurements $y$:

$$s = \arg\min_{s'} \|s'\|_0, \text{ such that } y = C\Psi s',$$  \hspace{1cm} (8)

where $\|s\|_0$ is the $\ell_0$ pseudo-norm corresponding to the number of non-zero entries of $s$. Unfortunately, this optimization problem is intractable, requiring a combinatorial brute-force search across all sparse vectors $s$. A major innovation of compressed sensing is a set of conditions on the measurement matrix $C$ that allow the non-convex $\ell_0$-minimization in (8) to be relaxed to a convex $\ell_1$-minimization:

$$s = \arg\min_{s'} \|s'\|_1, \text{ such that } y = C\Psi s',$$  \hspace{1cm} (9)

where $\|s\|_1 = \sum_{k=1}^n |s_k|$. This formulation is shown schematically in Fig. 1.

For the $\ell_1$-minimization in (9) to yield the sparsest solution in (8) with high probability, the
measurements $C$ must be chosen so that $\Theta = C\Psi$ satisfies a restricted isometry property (RIP):

$$(1 - \delta_K)\|s\|_2^2 \leq \|C\Psi s\|_2^2 \leq (1 + \delta_K)\|s\|_2^2,$$  

(10)

where $\delta_K$ is a small positive restricted isometry constant. In particular, there are two conditions on $C$ for a RIP to be satisfied for all $K$-sparse vectors $s$:

1. The measurements $C$ must be incoherent with respect to the basis $\Psi$. This incoherence means that the rows of $C$ are sufficiently uncorrelated with the columns of $\Psi$, as quantified by $\mu$:

$$\mu(C, \Psi) = \sqrt{n} \max_{j,k} \|c_k^\top \psi_j\|.$$  

(11)

Small $\mu$ indicates better incoherent measurements, with an optimal value of $\mu = 1$. Here, $c_k$ denotes the $k$-th row of $C$ and $\psi_j$ the $j$-th column of $\Psi$, both of which are assumed to be normalized.

2. The number of measurements $p$ must satisfy:

$$p \sim O(K \log(n/K)).$$  

(12)

The $K \log(n/K)$ term above is generally multiplied by a small constant multiple of the incoherence. Thus, fewer measurements are required if they are less coherent.

Intuitively, the existence of a RIP implies that the geometry of sparse vectors is preserved through the measurement matrix $C\Psi$. Determining the exact constant $\delta_K$ may be extremely challenging in practice, and it tends to be more desirable to characterize the statistical properties of $\delta_K$, as the measurement matrix $C$ may be randomly chosen. Panel 6 describes why it is not possible to use QR pivot locations as optimized sensors for compressed sensing, since they fail to identify the sparse structure of an unknown signal.

Typically, a generic basis such as Fourier or wavelets may be used to represent the signal sparsely. Spatially localized measurements (i.e., single pixels in the case of an image) are optimally incoherent with respect to the Fourier basis, so that $\mu(C, \Psi) = 1$. Thus, single pixel measurements are ideal because they excite a broadband frequency response. In contrast, a measurement corresponding to a fixed Fourier mode would be uninformative; if the signal is not sparse in this particular frequency, this measurement provides no information about the other Fourier modes. For many engineering applications, spatially localized measurements are desirable, as they correspond to physically realizable sensors, such as buoys placed in the ocean.

One of the major results of compressed sensing is that random projection measurements of the state (i.e., entries of $C$ that are Bernoulli or Gaussian random variables) are incoherent with respect to nearly any generic basis $\Psi$. This result is truly remarkable; however, the incoherence of random projections is not optimal, and typically scales as $\mu \sim \sqrt{2 \log(n)}$. Moreover, it may be difficult to obtain random projections of the full state $x$ in physical applications.

There are many alternative strategies to solve for the sparsest solution to (7). Greedy algorithms
are often used [66, 67, 68, 69, 70, 71, 72, 73, 74], including the compressed sampling matching pursuit (CoSaMP) algorithm [75].

1.2 Example: Compressed sensing

As a simple example, we consider a sparse signal that is constructed as the sum of three distinct cosine waves:

\[
x(t) = \cos(2\pi \times 37t) + \cos(2\pi \times 420t) + \cos(2\pi \times 711t).
\]

The Shannon-Nyquist sampling theorem [76, 77] states that for full signal reconstruction, we must sample at twice the highest frequency present, indicating a theoretical minimum sampling rate of 1422 Hz. However, since the signal is sparse, we may sample at considerably lower than the Nyquist rate, in this case at an average of 256 Hz, shown in Fig. 2. Note that for accurate sparse signal reconstruction, these measurements must be randomly spaced in time, so that the relative spacing of consecutive points may be quite close or quite far apart. Spacing points evenly with a sampling rate of 256 Hz would alias the signal, resulting in poor reconstruction.

```matlab
N = 4096;
t = linspace(0, 1, N);
x = cos(2*pi*37*t) + cos(2*pi*420*t) + cos(2*pi*711*t);
perm = randperm(N,p);
y = x(perm);

%% Randomly sample signal
p = 256; % num. samples, p=N/16
Psi = dct(eye(N, N)); % build Psi
Theta = Psi(perm, :); % Measure Psi

%% Solve compressed sensing problem
Psi = dct(eye(N, N)); % build Psi
Theta = Psi(perm, :); % Measure Psi

xrecon = idct(s); % reconstruct x
```
2 Optimal sparse sensing in a tailored basis

The compressed sensing strategy above is ideal for the recovery of a high-dimensional signal of unknown content using random measurements in a universal basis. However, if information is available about the type of signal (e.g., the signal is a turbulent velocity field or an image of a human face), it is possible to design optimized sensors that are tailored for the particular signals of interest. Dominant features are extracted from a training dataset consisting of representative examples, for example using the proper orthogonal decomposition (POD). These low-rank features, mined from patterns in the data, facilitate the design of specialized sensors that are tailored to a specific problem.

Low-rank embeddings, such as POD, have already been used in the ROM community to select measurements in the state space that are informative for feature space reconstruction. The so-called empirical interpolation methods seek the best interpolation points for a given basis of POD features. These methods have primarily been used to speed up the evaluation of nonlinear terms in high-dimensional, parameterized systems. However, the resulting interpolation points correspond to measurements in state space, and their use for data-driven sensor selection has largely been overlooked. We will focus on this formulation of sensor selection and explore sparse, convex, and greedy optimization methods for solving it.

We begin with brief expositions on POD and our mathematical formulation of sensor selection, followed by an overview of broader sensor selection. We conclude this section with our generalized sensor selection method that connects empirical interpolation methods, such as QR pivoting to optimize condition number, with broader sensor selection techniques. The QR pivoting method, discussed in Sec. 2.3, is particularly favorable, as it is fast, simple to implement, and provides nearly optimal sensors tailored to a specific POD basis. Throughout this section, we make explicit certain connections shared with the broader sensor selection task, in which a subset of measurements is chosen from a larger set of possible measurements.

2.1 Proper orthogonal decomposition

POD is a widespread data-driven dimensionality reduction technique [13, 14] used in many domains; it is also commonly known as the Karhunen-Loève expansion, principal component analysis (PCA) [80], and empirical orthogonal functions [81]. POD expresses high-dimensional states \( \mathbf{x} \in \mathbb{R}^n \) as linear combinations of a small number of orthonormal eigenmodes \( \psi \) (i.e., POD modes) that define a low-dimensional embedding space. States are projected into this POD subspace, yielding a reduced representation that can be used to streamline tasks that would normally be expensive in the high-dimensional state space. This low-rank embedding does not come for free, but instead requires training data to tailor the POD basis to a specific problem. POD is illustrated on a simple example of extracting coherent features in images of human faces in Panel 3.

A low-dimensional representation of \( \mathbf{x} \) in terms of POD coefficients \( a \) can be lifted back to the full state with a linear combination of POD modes:

\[
\mathbf{x} \approx \sum_{k=1}^{r} a_k \psi_k.
\]

For time-series data \( \mathbf{x}(t) \), the coefficients \( a_k(t) \) vary in time and \( \psi_k \) are purely spatial modes without time dependence, resulting in a space-time separation of variables. Thus, care should be taken applying POD to data from a traveling wave problem.

The eigenmodes \( \psi_k \) and POD coefficients \( a_k \) are easily obtained from the singular value decomposition (SVD). Given a data matrix of state space observations \( \mathbf{X} = [x_1 \ x_2 \ \ldots \ x_m] \), the resulting eigenmodes are the orthonormal left singular vectors \( \mathbf{\Psi} \) of \( \mathbf{X} \) obtained via the SVD:

\[
\mathbf{X} = \mathbf{\Psi} \Sigma \mathbf{V}^T \approx \mathbf{\Psi}_r \Sigma_r \mathbf{V}^T_r.
\]  

(13)

The matrices \( \mathbf{\Psi}_r \) and \( \mathbf{V}_r \) contain the first \( r \) columns of \( \mathbf{\Psi} \) and \( \mathbf{V} \) (left and right singular vectors, respectively), and the diagonal matrix \( \Sigma_r \) contains the first \( r \times r \) block of \( \Sigma \) (singular values). The SVD is the optimal least-squares approximation to the data for a given rank \( r \), as demonstrated by the Eckart-Young theorem [82]:

\[
\mathbf{X}_* = \arg \min_{\mathbf{X}} \| \mathbf{X} - \mathbf{\bar{X}} \|_F \quad \text{s. t. rank}(\mathbf{\bar{X}}) = r,
\]

(14)
Panel 3: Proper orthogonal decomposition and eigenfaces

One of the most visually striking and intuitive applications of POD is the feature extraction of facial images. These POD eigenmodes of these datasets are called eigenfaces due to their resemblance to generic human faces. We demonstrate this application of POD on the extended Yale B dataset \([78, 79]\), consisting of cropped and aligned images of several individuals in different lighting conditions. We obtain a resized version of the dataset in the form of MATLAB data files from http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html. Each image is a 32 × 32 matrix of grayscale pixel values, which is reshaped into a 1024 × 1 vector and stacked as a column of a data matrix \(X\). This example is a benchmark problem for sensor selection in Sec. 3.2.

![Figure 3: Training images from Extended Yale B dataset will be used as inputs to POD.](image)

\[
\text{meanface} = \text{mean}(X, 2);
\]
\[
X = X - \text{repmat}(	ext{meanface}, 1, \text{size}(X, 2)); \quad \% \text{mean centered data}
\]

Figure 4: First ten eigenfaces or POD modes successfully recover important facial information such as the main facial features (eyes, nose, mouth) followed by depth information (brows, ridges, chin).

\[
\begin{align*}
\% \text{proper orthogonal decomposition} \\
[\Psi, \Sigma, V] = \text{svd}(X, 'econ'); \\
eigenfaces = \Psi(:, 1:10);
\end{align*}
\]

where \(X_* = \Psi_r \Sigma_r V_r^T\), and \(\| \cdot \|_F\) is the Frobenius norm. The low-dimensional vector of POD coefficients for a state \(x\) is given by the orthogonal projection \(a = \Psi_r^T x\). Thus, the POD is a widely used dimensionality reduction technique for high-dimensional systems. This reduction allows computational speedup of numerical time-stepping, parameter estimation, and control.

Choosing the intrinsic target rank without magnifying noise in the data is a difficult task. In practice, \(r\) is often chosen by thresholding the singular values to capture some percentage of the variance in the data. Gavish and Donoho \([83]\) derive an optimal hard threshold for singular values based on the singular value distribution and aspect ratio of the data matrix, assuming additive Gaussian white noise of unknown variance. This threshold criterion has been effective in practice, even in cases where the noise is likely not Gaussian.

### 2.2 Sensor selection formulation

Sensor selection identifies optimized measurements for reconstruction of unknown states \(x \in \mathbb{R}^n\) that can be expanded in a known basis \(\Psi_r\)

\[
x_j = \sum_k \Psi_{jk} a_k, \quad (15)
\]

where \(\Psi_{jk}\) is the coordinate form of \(\Psi_r\). Consider the \(p\) linear measurements of \(x\) stored in \(y = Cx\),

\[
y_i = \sum_j C_{ij} x_j = \sum_j C_{ij} \sum_k \Psi_{jk} a_k. \quad (16)
\]

When \(x\) is unknown, it can be reconstructed by approximating the unknown basis coefficients \(a\) with the Moore-Penrose pseudoinverse, \(a = (C \Psi_r)^{\dagger} y\). Note that the above equations are equivalent to (2) and (3) from Panel 1. A schematic of sparse sampling in a tailored basis \(\Psi_r\) is shown in Fig. 5.
The goal of sparse measurement selection is to choose a measurement matrix \( C \) representing \( p \) distinct point measurements in state space, so that \( C \) consists of \( p \) rows of the identity matrix. Although more general linear measurements of \( x \) may be admissible in some problems, point measurements are physically appealing as spatially localized sensors. The row-wise sum \( c^+ = \sum_j c_k \) satisfies \( \|c^+\|_0 = p \), where \( c_k \) is the \( k \)-th row of \( C \). The sensor selection problem can be formulated as choosing \( C \) to make the pseudoinverse of \( \Theta = C\Psi_r \) as well-conditioned as possible, thus making the estimation of the coefficients in a robust to measurement noise on \( y \). The condition number is discussed in more detail in Panel 4.

Condition number minimization is the viewpoint taken by the reduced-order modeling community for selecting point measurements within a low-rank POD basis. These empirical interpolation methods (EIM) assume \( p = r \) sensors so \( \Theta \) is square and invertible. EIM and subsequent discrete variants such as DEIM \cite{85} and Q-DEIM \cite{86} are greedy procedures that control the condition number of \( (C\Psi_r)^{-1} \) by minimizing the spectral norm (i.e. largest singular value)

\[
C_* = \arg\min_C \| (C\Psi_r)^{-1} \|_2, \tag{21}
\]

where \( C \) is subject to the same structural constraints as above. Related strategies include selecting measurements at maxima of successive POD modes or iteratively seeking measurements that decrease the condition number of \( C\Psi_r \) \cite{87, 88}.

This strategy is particularly advantageous when only a few modes are required to characterize the data. When singular values decay slowly, we may require \( p > r \) measurements for well-conditioned reconstruction. In Sec. 2.3, we generalize one particular empirical interpolation method, Q-DEIM, which uses QR pivoting to determine sensor locations. In particular, we extend this method to the case of \( p > r \) sensors, making it viable for more general sensor selection and significantly speeding up its computation.

### 2.2.1 General formulation

Sensor selection is more generally framed as follows: given a set of \( n \) possible measurement indices \( V \), select a subset of sensor indices \( S \) to optimize a carefully chosen function evaluating their quality. The problem is mathematically realized as

\[
\max_{S \subseteq V} f(S) \text{ subject to constraints on } S. \tag{22}
\]

This problem belongs to an area of research called submodular function optimization, surveyed in \cite{89}. The so-called submodular function, \( f(S) : 2^V \rightarrow \mathbb{R} \), evaluates these subsets to some scalar value, thus rendering a brute force search over its entire domain (the power set of all possible subsets \( 2^V \)) computationally intractable even for moderate values of \( n \).

Possible choices of \( f \) include mutual information, entropy, estimation error covariance, and spatial coverage of sensors. Often, even a single evaluation of \( f \) is expensive when matrix determinants or factorizations are involved. The present work aims to reconstruct states from measurements with minimal variance of reconstruction error, which requires maximizing the determinant of error covariance matrices. The minimization of various error metrics is an active topic in optimal experiment design, for instance, the determinant or volume (D-optimal), the trace or mean-squared error (A-optimal), and maximum eigenvalue or worst-case variance (e-optimal) of error covariance matrices.

Joshi and Boyd \cite{90} frame sensor selection as selecting \( p \) rows of \( \Psi_r \) that minimize the volume of the estimation error covariance matrix, \( (C\Psi_r)^T C\Psi_r = \Theta^T \Theta \). Since matrix volume is the
**Panel 4: Condition number**

The condition number of a matrix $\Theta$ gives a measure of how sensitive matrix multiplication or inversion is to errors in the input, with larger condition number indicating worse performance. The condition number $\kappa(\Theta)$ is given by the ratio of the maximum and minimum singular values of $\Theta$:

$$\kappa(\Theta) = \frac{\sigma_{\text{max}}(\Theta)}{\sigma_{\text{min}}(\Theta)}.$$  \hspace{1cm} (17)

To see the effect of the condition number on matrix multiplication, consider a square, invertible $\Theta$ and an input signal $x$ that is contaminated by noise $\epsilon_x$. Further, we consider the worst-case scenario where $x$ is aligned with the right singular vector of $\Theta$ corresponding to the minimum singular value $\sigma_{\text{min}}(\Theta)$ and where the error $\epsilon_x$ is aligned with the right singular vector of $\Theta$ corresponding to the maximum singular value $\sigma_{\text{max}}(\Theta)$. Thus, error is scaled by $\sigma_{\text{max}}$ while the signal is scaled by $\sigma_{\text{min}}$:

$$\Theta (x + \epsilon_x) = \sigma_{\text{min}}x + \sigma_{\text{max}}\epsilon_x.$$ \hspace{1cm} (18)

Thus, we see that the signal to noise ratio (SNR) is reduced by a factor equal to the condition number:

$$\text{SNR}_{\text{in}} = \frac{x}{\epsilon_x} \implies \text{SNR}_{\text{out}} = \frac{\sigma_{\text{min}}}{\sigma_{\text{max}}} \frac{x}{\epsilon_x} = \text{SNR}_{\text{in}} / \kappa.$$ \hspace{1cm} (19)

If the condition number is large, then error can be amplified relative to the signal. The ideal condition number is 1, where $\Theta$ has all singular values equal to 1, for example if $\Theta$ is a unitary matrix.

Similarly, the sensitivity of the matrix inverse to error is also related to the condition number:

$$x + \epsilon_x = \Theta^{-1} (y + \epsilon_y).$$ \hspace{1cm} (20)

Again, worst-case scenario errors in the inversion are amplified by the condition number (this time with the error being aligned with the singular vector corresponding to minimum singular value).

In general, a random error $\epsilon_x$ or $\epsilon_y$ will have some component in this worst-case direction, and will be amplified by the maximum singular value of $\Theta$ or $\Theta^{-1}$. Thus, it is desirable to explicitly control the condition number of $\Theta = C\Psi_r$ by choice of the row selection operator $C$. For invertible matrices, the condition numbers of $\Theta$ and $\Theta^{-1}$ are the same. The discussion above generalizes for rectangular $\Theta$.
is given by the pivoted QR factorization of its condition number, the general sensor selection of pivoted QR to minimize matrix volume through Case where desired measurement matrix modes when selecting Case where row selected POD basis, as described below. The pivoted QR yields an orthonormal basis, which is known as Q-DEIM in ROM settings. QR pivoting itself dates back to the 1960s by Businger and Golub to solve least-squares problems [91], and it has found utility in various measurement selection applications [92, 93, 94]. Similar to empirical interpolation methods such as DEIM, pivots from the QR factorization optimally condition the measurement or row selected POD basis, as described below.

**Case where** \(p = r\): In empirical interpolation, when selecting \(r\) interpolation points within \(r\) basis modes \(\Psi_r\), the pivoted QR yields an orthonormal matrix \(Q\), an upper triangular matrix \(R\), and the desired measurement pivot matrix \(C\) such that

\[
\Psi_r^T C^T = QR. \tag{24}
\]

\[
\text{\$ QR sensor selection, } p>K \text{ \{Q,R,pivot\} = qr(Psi_r'*Psi_r', 'vector'); pivot = pivot(1:p)}
\]

This requires an expensive QR factorization of an \(n \times n\) matrix, whose storage requirements scale quadratically with state dimension. However, this operation may be advantageous for several reasons. The row selection given by the first \(p\) QR pivots increase the leading \(r\) singular values of \(\Theta^T\), hence increasing \(\det(\Theta^T\Theta)\) – the same objective maximized by the convex optimization of Joshi and Boyd [90]; their method performs one Cholesky factorization on an \(n \times n\) matrix with runtime \(O(n^3)\) per Newton iteration, and this must be recomputed for each new choice of \(p\). Our proposed method only requires a single \(O(n^3)\) QR factorization and results in a hierarchical list of all \(n\) total pivots, with the first \(p\) pivots optimized for reconstruction in \(\Psi_r\) for any \(p > r\). Thus, additional sensors may be leveraged if available. Panel 5 summarizes the methods discussed.

The QR factorization is implemented and optimized in most standard scientific computing packages and libraries, including MATLAB, LAPACK, NumPy, among many others. In addition to software-enabled acceleration, QR runtime can be

<table>
<thead>
<tr>
<th>Method</th>
<th>Objective</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\ell_0) combinatorial search</td>
<td>(C_0 = \text{argmax} \log det(C\Psi_r)^T C\Psi_r) subject to (</td>
<td>c^+_i</td>
</tr>
<tr>
<td>(\ell_1) convex optimization</td>
<td>(C_1 = \text{argmax} \log det(C\Psi_r)^T C\Psi_r) subject to (</td>
<td>c^+_i</td>
</tr>
<tr>
<td>SSPOR greedy</td>
<td>Case (p = r) : (\Psi_r^T C^T = QR) Case (p &gt; r) : ((\Psi_r \Psi_r^T) C^T = QR)</td>
<td>(O(n^3))</td>
</tr>
</tbody>
</table>
Panel 6: Why QR pivoting does not work for compressed sensing

The QR pivoting algorithm can be used to efficiently find a sparse solution of an underdetermined linear system \( y = \Theta s \). In fact, in MATLAB, this is as simple as using the `\` command:

\[
\| s = \text{Theta} \backslash y; \|
\]

However, QR pivoting does not always provide the sparsest solution. In addition, the pivot locations, and hence the nonzero entries in \( s \), are determined entirely by \( \Theta \). Therefore, the structure of the nonzero coefficients in \( s \) have nothing to do with the specific signal or measurements. Given \( p > K \) measurements of an unknown signal in a universal basis, the QR pivot algorithm will fail in two ways: 1) it will return a sparse vector with \( p > K \) nonzero elements (instead of the desired \( K \) nonzero elements), and 2) the nonzero elements will not depend on the frequency content of the actual signal. This observation is illustrated on two test images below. Notice that the sparse coefficients for both the mountain and cappuccino are the same, resulting in poor image reconstruction.

significantly reduced by terminating the procedure after the first \( p \) pivots are obtained. The operation can be accelerated further using randomized techniques, for instance, by the random selection of the next pivot [86] or by using random projections to select blocks of pivots [95, 96, 97]. Panel 6 shows why QR pivoting does not find the sparsest vector in an universal basis for compressed sensing.

Related sensor selection concepts generalize to data-driven classification. For image classification using linear discriminant analysis (LDA), sparse sensors may be selected that map via POD modes into the discriminating subspace [30]. Moreover, sparse classification within libraries of POD modes [98, 99, 40] can be improved by augmenting DEIM samples with a genetic algorithm [100] or adapting QR pivots for classification [58].
3 Comparison of methods

In this section, sensor selection and signal reconstruction algorithms are implemented and compared on data from fluid dynamics, facial images, and ocean surface temperatures. The examples span a wide range of complexity, exhibit both rapid and slow singular value decay, and come from both static and dynamic systems.

In each example, optimized sensors obtained in a tailored basis using the SSPOR/QR pivoting algorithm (Sec. 2.3) outperform random measurements in a universal basis using compressed sensing (Sec. 1.1) for signal reconstruction. Moreover, for the same reconstruction performance, many fewer SSPOR/QR tailored sensors are required, decreasing the cost associated with purchasing, placing, and maintaining sensors, as well as reducing the latency in computations. Thus, for a well-scoped reconstruction task with sufficient training data, we advocate principled sensor selection rather than compressed sensing. When the structure of the underlying signal is unknown, then compressed sensing provides more flexibility with an associated increase in the number of sensors.

3.1 Flow past a cylinder

Fluid flow past a stationary cylinder is a canonical example in fluid dynamics that is high-dimensional yet reveals strongly periodic, low-rank phenomena. It is included here as an ideal system for reduction via POD and hence, minimal sensor placement. The data is generated by numerical simulation of the linearized Navier-Stokes equations using the immersed boundary projection method (IBPM) based on a fast multi-domain method [101, 102]. The computational domain consists of four nested grids so that the finest grid covers a domain of $9 \times 4$ cylinder diameters and the largest grid covers a domain of $72 \times 32$. Each grid has resolution $450 \times 200$, and the simulation consists of 151 timesteps with $\delta t = 0.02$. The Reynolds number is 100, and the flow is characterized by laminar periodic vortex shedding [103]. Vorticity field snapshots are shown in Fig. 7.

The POD modes of this data reflect oscillatory dynamics characterized by periodic vortex shedding. The data is low-rank, and the singular values decay rapidly, as shown in Fig. 8a. The singular values occur in pairs corresponding to harmonics of the dominant vortex shedding frequency. Most of the spectral energy in the dataset is captured by the first 42 POD modes. Thus the intrinsic rank of the dataset in POD feature space is $r = 42$, and the minimal number of sensors given by the SSPOR/QR method is $p = 42$. In this manner we bypass the expensive QR factorization of $\Psi \Psi^T$, a matrix of dimensions 89351 × 89351.

Reconstruction from QR sensors (Fig. 8b) successfully captures modal content with only $p = r$ sensors when fitting to the first 42 POD modes. The first 42 POD modes characterize nearly 100% of the system’s energy, the normalized sum of the singular values. Using modes beyond $r > 42$ results in overfitting, and QR pivoting selects sensors based on uninformative modes. Thus, accuracy stops improving beyond $r = 42$ target modes, whereupon sensor data amplifies noise. However, these tailored sensors perform significantly better than random sensors due to the favorable conditioning properties of QR interpolation points. Panel 7 illustrates the ability of optimized sensing to significantly reduce the number of sensors required for a given performance.

3.2 Extended Yale B eigenfaces

Image processing and computer vision commonly involve high-resolution data with dimension determined by the number of pixels. Cameras and
Panel 7: Selecting number of sensors \( p \) and rank \( r \) for flow past a cylinder

An overarching goal of optimized sensor placement is choosing the fewest \( p \) sensors for reconstruction. This sparse sensor optimization is facilitated by low-rank structure in the data (Fig. 8a) and inherently involves a trade-off between the number of sensors and reconstruction accuracy. As seen in Fig. 8b, effective sensor optimization moves the elbow of the reconstruction error curve down and to the left, indicating accurate reconstruction with few sensors. Reducing the number of sensors may be critically enabling when sensors are expensive or when low computational latency is desired.

![Figure 8](image-url)

Figure 8: Illustration of singular value decay and sparse signal reconstruction in the fluid flow past a cylinder example. (a) The exponential decay of singular values indicates low-rank dynamics with a modal truncation at \( r = 42 \) modes, the last modes containing any meaningful spatial structure. (b) Recovery with a minimal number of QR sensors with increasing POD basis rank is nearly as good as a snapshot’s approximation in the same POD basis, and overfitting occurs beyond \( r > 42 \). Optimized sensing remains orders of magnitude more accurate than reconstruction with random sensors.

Recording devices capture massive images with rapidly increasing pixel and temporal resolution. However, most pixel information in an image can be discarded for subject identification and automated decision-making tasks.

The extended Yale B face database [78, 79] is a canonical dataset used for facial recognition, and it is an ideal test bed for recovering low-rank structure from high-dimensional pixel space. The data consists of 64 aligned facial images each of 38 stationary individuals in different lighting conditions. We validate our sensor (pixel) selection strategy by recovering missing pixel data in a validation image using POD modes or *eigenfaces* trained on 32 randomly chosen images of each individual.

![Figure 9](image-url)

The normalized singular values are shown in Fig. 9, and the Gavish-Donoho criterion [83] determines a singular value threshold at \( r = 166 \), indicating the intrinsic rank of the training dataset. Indeed, selected eigenfaces are also shown to reveal no meaningful facial structure beyond eigenface 166. QR pixel selection is performed on the first 50 and first 166 eigenfaces, and selected pixels shown in Fig. 10 cluster around important facial features – eyes, nose and mouth.

The image reconstructions in Fig. 10 are estimated from the same number of selected pixels as the number of modes used for reconstruction. For instance, the 50 eigenface reconstruction is uniquely constructed from 50 selected pixels out
Figure 9: Normalized singular values and selected eigenfaces. Facial information progressively decreases across selected eigenfaces, and no facial features can be readily discerned beyond eigenface $r = 166$, the optimal Gavish-Donoho modal truncation value.

Figure 10: Comparison of reconstruction and sensor selection methods. QR sensors emphasize facial features such as the eyes, nose and mouth, and hence achieve adequate reconstruction with as few as 166 sensors (16% of all pixels). $\ell_2$ reconstruction with 50 QR sensors (5% of pixels) surpasses the performance of compressed sensing with 300 random pixels. In comparison, compressed sensing requires 600 sampled pixels for comparable recovery, while $\ell_2$ reconstruction with random sensors is consistently poor.

Of 1024 total – 5% of available pixels. Even at lower pixel selection rates, least-squares reconstruction from QR selected pixels is more successful at filling in missing data and recovering the subject’s face. For comparison, reconstruction of the same face from random pixels using compressed sensing is shown in Fig. 10. Compressed sensing in a universal Fourier basis demonstrates progressively better global feature recovery. However, more than triple the pixels are required for the same quality of reconstruction as in QR selection. Moreover, the convex $\ell_1$ optimization procedure is extremely expensive compared to the single $\ell_2$ regression on subsampled eigenfaces. Therefore data-driven feature selection and structured measurement selection are of significant computational
and predictive benefit, and occur at the small training expense of one SVD and pivoted QR operation.

The convergence of reconstruction with SSPOR sensors is shown in Fig. 11. More sensors than modes are used in reconstruction for this example. The expected error dropoff is observed with increasing number of modes and sensors, although the dropoff is slower than for the cylinder flow (Fig. 8b) due to slower decay of singular values.

### 3.3 Sea surface temperature (SST)

Next we consider the NOAA_OLSST_V2 global ocean surface temperature dataset spanning the duration 1990–2016. The data is publicly available at https://www.esrl.noaa.gov/psd/. Unlike eigenfaces, this dataset is a time series, for which a snapshot is recorded every week. Sensor selection must then track energetic temporal signatures. Sensors and features are trained on the first 16 years (832 snapshots), and a test snapshot is selected from the excluded validation set. The singular value distribution is shown in Fig. 12.

Like the eigenfaces, localized convective phenomena have energetic contributions to otherwise globally uninformative eigenssts. This is best seen in the POD snapshot projections, in which the 100 eigensst projection already sufficiently recovers dynamics, while increasing the number of eigenssts in the projection further refines convect-
tive phenomena. These lower-energy modes containing convective effects contribute to some degree of overfitting in $\ell_2$ reconstruction (Fig. 13). The most interesting of these is the El Niño southern oscillation (ENSO) feature that is clearly identified from QR selected sensors. El Niño is any increase of a-b degrees lasting greater than six months in this highlighted region of the South Pacific. It has been implicated in global weather patterns, warming and climate change.

Remark: Modal separation of intermittent phenomena such as the El Niño is difficult from a time-invariant POD analysis. Separation of isolated, low-energy temporal events cannot be done from a variance-characterizing decomposition such as the POD – reordering the snapshots will yield the same dominant modes. On the other hand, tensor decompositions and temporal-frequency analyses such as multi-resolution dynamic mode decomposition have succeeded at identifying El Niño where POD has failed. Sensor selection using non-normal modes arising from such decompositions remains an open problem and the focus of ongoing work.

4 Discussion

The efficient sensing of complex systems is an important challenge across the physical, biological, social, and engineering sciences, with significant implications for nearly all downstream tasks. In this work, we have demonstrated the practical implementation of several sparse sensing algorithms on a number of relevant real-world examples. As discussed throughout, there is no all-purpose strategy for the sparse sensing of a high-dimensional system. Instead, the choice depends on key factors such as the amount of training data available, the scope and focus of the desired estimation task, cost constraints on the sensors themselves, and the required latency of computations on sensor data. Thus, we partition the sparse sensing algorithms into two fundamental categories: 1) optimized sensing in a data-driven tailored basis, and 2) random sensing in a universal basis.

A critical comparison of the two approaches highlights a number of relative strengths and weaknesses. The first strategy results in a highly optimized set of sensors that are suitable for tightly scoped reconstruction problems where sufficient training data is available. The second strategy requires more sensors for accurate reconstruction but also makes fewer assumptions about the underlying signal, making it more general. We emphasize that optimized sensing in a tailored basis typically provides more accurate signal reconstruction than random measurements, facilitating a reduction in the number of sensors by about a factor of two.
Further, sensor selection and signal reconstruction in the tailored basis is computationally efficient and simple to implement, while compressed sensing generally requires a costly iterative algorithm. In problems where sensors are expensive, or when low-latency decisions are required, the reduction in the number of sensors and the associated speed-up of optimized sensing can be significant. Thus, when the reconstruction task is well-scoped and a sufficient volume of training data is available, we advocate principled sensor selection rather than compressed sensing.

4.1 Potential applied impact

Many fields in science and engineering rely on sensing and imaging. Moreover, any application involving feedback control for stabilization, performance enhancement, or disturbance rejection relies critically on the choice of sensors. We may roughly categorize these sensor-critical problems into two broad categories: 1) problems where sensors are expensive and few (ocean sampling, disease monitoring, espionage, etc.), and 2) problems where sensors are cheap and abundant (cameras,
high-performance computation, etc.).

In the first category, where sensors come at a high cost, the need for optimized sparse sensors is clear. However, it is not always obvious how to collect the training data required to optimize these sensors. In some applications, high-fidelity simulations may provide insight into coherent structures, whereas in other cases a large-scale survey may be required. It has recently been shown that it may be possible to optimize sensors based on heavily subsampled data, as long as coherent structures are non-localized [30].

In the second category, where sensors are readily available, it may still be advantageous to identify key sensors for fast control decisions. For example, in mobile applications, such as vision-based control of a quad-rotor or underwater monitoring of an energy harvesting site using an autonomous underwater vehicle, computational and battery resources may be limited. Restricting high-dimensional measurements to a small subset of key pixels speeds up computation and reduces power consumption. Similar performance enhancements are already exploited in high-performance computing, where expensive function evaluations are avoided by sampling at key interpolation points [104, 85]. Finally, it may also be the case that if measurements are corrupted by noise, reconstruction may improve if uninformative sensors are selectively ignored.

4.2 Extensions to dynamics, control, and multiscale physics

Data-driven sensor selection is generally used for instantaneous full-state reconstruction, despite the fact that many signals are generated by a dynamical system [105, 14]. Even in reduced-order models, sensors are typically used to estimate nonlinear terms instantaneously without taking advantage of the underlying dynamics. However, it is well known that for linear control systems [60, 59], the high-dimensional state may be reconstructed with few sensors, if not a single sensor, by leveraging the time history in conjunction with a model of the dynamics, as exemplified by the Kalman filter [106, 107]. In dynamic estimation and control, prior placement of sensors and actuators is generally assumed. Extending the sensor placement optimization to the model reduction [23, 24, 25] and system identification [26, 108, 109, 110, 111, 112] of linear control systems is an important avenue of ongoing work. In particular, sensors and actuators may be chosen to increase the volume of the controllability and observability Gramians, related to the original balanced truncation literature [23]. More generally, sensor and actuator placement may be optimized for robustness [113, 114, 115, 116], or for network control and consensus problems [117, 118, 119, 120, 121, 122, 123, 124].

The sensor placement algorithms discussed above are rooted firmly in linear algebra, making them readily extensible to linear control systems. Recent advances in dynamical systems are providing techniques to embed nonlinear systems in a linear framework through a suitable choice of measurement functions of the state, opening up the possibility of optimized sensing for nonlinear systems. As early as the 1930s, Koopman demonstrated that a nonlinear system can be rewritten as an infinite-dimensional linear operator on the Hilbert space of measurement functions [125, 126]. This perspective did not gain traction until modern computation and data collection capabilities enabled the analysis of large volumes of measurement data.

Mezic et al. [127, 128, 129, 130, 131, 132] have developed much of the modern Koopman operator theory, and it has been shown that under certain conditions DMD approximates the Koopman operator [15, 16, 17, 18]; sensor fusion is also possible in the Koopman framework [133]. Recently, Koopman analysis has been used to develop nonlinear estimators [134, 135] and controllers [136], although establishing rigorous connections to control theory is an ongoing effort [137, 138, 139]. Koopman theory has also been used to analyze chaotic dynamical systems from time-series data [140, 141], relying on the Takens embedding [142], which is related to sensor selection.

Modern Koopman theory may drive sensor placement and the selection of nonlinear measurement functions on the sensors to embed nonlinear dynamics in a linear framework for optimal nonlinear estimation and control. This approach is consistent with neural control systems, where biological sensor networks (e.g., strain sensors on an
insect wing) are processed through nonlinear neural filters before being used for feedback control.

Beyond extending sensor selection to nonlinear systems and control, there is a significant opportunity to apply principled sensor selection to multiscale systems. Turbulence is an important high-dimensional system that exhibits multiscale phenomena [14, 143, 144, 145]. Data-driven approaches have been used to characterize turbulent systems [1], including clustering [146], network theory [147, 148], DMD-based model reduction [149, 150], and local POD subspaces [151], to name a few. Recently, a multisresolution DMD has been proposed [152], where a low-dimensional subspace may locally characterize the attractor, despite a high-dimensional global attractor. This approach may significantly reduce the number of sensors needed for multiscale problems.

Reproducible Research

A MATLAB code supplement is available at https://github.com/kmanohar/SSPOR_pub [153] for reproducing results in this manuscript, including:

1. Datasets in MATLAB file formats, or links to data that are publicly available online;
2. MATLAB scripts to recreate figures of results.

Acknowledgments

The authors thank Eurika Kaiser, Joshua Proctor, Bernd Noack, Tom Hogan, and Joel Tropp for valuable discussions. SLB and JNK acknowledge support from the Defense Advanced Research Projects Agency (DARPA HR0011-16-C-0016). SLB and KM acknowledge support from the Air Force Research Labs (FA8651-16-1-0003).

References


