Blind Calibration of Networks of Sensors: Theory and Algorithms

Laura Balzano
University of California, Los Angeles
sunbeam@ee.ucla.edu

Robert Nowak
University of Wisconsin, Madison
nowak@ece.wisc.edu

May 2007

Abstract

This chapter considers the problem of blindly calibrating sensor response using routine sensor network measurements. We show that as long as the sensors slightly oversample the signals of interest, then unknown sensor gains can be perfectly recovered. Remarkably, neither a controlled stimulus nor a dense deployment is required. We also characterize necessary and sufficient conditions for the identification of unknown sensor offsets. Our results exploit incoherence conditions between the basis for the signals and the canonical or natural basis for the sensor measurements. Practical algorithms for gain and offset identification are proposed based on the singular value decomposition and standard least squares techniques. We investigate the robustness of the proposed algorithms to model mismatch and noise on both simulated data and on data from current sensor network deployments.

1 Introduction

With the wide variety of sensor network applications being envisioned and implemented, it is clear that in certain situations the applications need more accurate measurements than uncalibrated, low-cost sensors provide. Arguably, calibration errors are one of the major obstacles to the practical use of sensor networks [3], because they allow a user to infer a difference between the readings of two spatially separated sensors when in fact that difference

may be due in part to miscalibration. Consequently, automatic methods for jointly calibrating sensor networks in the field, without dependence on controlled stimuli or high-fidelity groundtruth data, is of significant interest. We call this problem *blind calibration*.

One approach to blind sensor network calibration is to begin by assuming that the deployment is very dense, so that neighboring nodes should (in principle) have nearly identical readings [4]. Unfortunately, many existing and envisioned sensor network deployments may not meet the density requirements of such procedures.

The difference in our approach is that it leverages correlation in the collection of sensors without requiring a dense deployment, making it much more suitable for practical applications. We assume this correlation is defined and that the readings from n sensors lie in a subspace of n-dimensional Euclidean space.

We assume a linear model for the sensor calibration functions. This means that the sensor readings are calibrated up to an unknown gain and offset (bias) for each sensor, possibly after applying a suitable and fixed transformation to the raw sensor readings, e.g., taking the logarithm or applying the original factory calibration transformation.

This work makes three main contributions. First, we propose a novel automatic sensor calibration procedure that requires solving a linear system of constraints involving routine sensor measurements. By "routine" we mean that actual signal measured by the sensor network is uncontrolled and unknown. This is why we refer to the problem as blind calibration. The constraint equations are based on mild assumptions that guarantee that the sensor measurements are at least slightly correlated over space, i.e., the network oversamples the underlying signals of interest. Second, we prove the rather surprising fact that these assumptions, which are commonly met in practice, suffice to perfectly recover unknown sensor gains. That is, *it is possible to blindly calibrate the gains using only the routine readings made by the sensors*. Third, we prove that the sensor offsets (biases) can also be partially recovered from routine readings; they can be completely recovered with some additional overhead.

To give a preview of our approach, suppose we are measuring a temperature field with an array of n sensors. Temperature fields tend to vary smoothly, and so they may be considered to be bandlimited. The Nyquist theorem dictates a minimum spacing between sensors in order to adequately sample a bandlimited signal. If sensors are spaced more closely than the minimum requirement, then we are "oversampling" the signal. In this case, the underlying bandlimited signal will lie in a lower dimensional (low frequency)

subspace of the n dimensional measurement space. This condition provides a useful constraint for blind calibration. Correctly calibrated signals must lie in the lower dimensional subspace, and this leads to a system of linear equations which can be used to solve for the gain and offset calibration parameters.

2 Problem Formulation

Consider a network of n sensors. At a given time instant, each sensor makes a measurement, and we denote the vector of n measurements by $\boldsymbol{x} = [x(1), \ldots, x(n)]'$, where $'$ denotes the vector transpose operator (so that $\boldsymbol{x}$ is an $n \times 1$ column vector). We will refer to $\boldsymbol{x}$ as a "snapshot." When necessary, we will distinguish between snapshots taken at different times using a subscript (e.g., $\boldsymbol{x}_s$ and $\boldsymbol{x}_t$ are snapshots at times s and t).

Each sensor has an unknown gain and offset associated with its response, so that instead of measuring $\boldsymbol{x}$ the sensors report

$$y(j) \;=\; \frac{x(j) - \beta(j)}{\alpha(j)},\; j = 1, \ldots, n$$

where $\boldsymbol{\alpha} = [\alpha(1), \ldots, \alpha(n)]'$ are the sensors' gain calibration factors and $\boldsymbol{\beta} = [\beta(1), \ldots, \beta(n)]'$ are the sensors' calibration offsets. It is assumed that $\alpha(j) \neq 0, j = 1, \ldots, n$. With this notation, the sensor measurement $y(j)$ can be calibrated by the linear transformation $x(j) = \alpha(j)y(j) + \beta(j)$. We can summarize this for all n sensors using the vector notation

$$\boldsymbol{x} \;=\; \boldsymbol{Y}\boldsymbol{\alpha} + \boldsymbol{\beta}\;, \tag{1}$$

where $\boldsymbol{Y} = \mathrm{diag}(\boldsymbol{y})$ and the diag operator is defined as

$$\mathrm{diag}(\boldsymbol{y}) \;=\; \begin{bmatrix} y(1) & & \\ & \ddots & \\ & & y(n) \end{bmatrix}.$$

The blind calibration problem entails the recovery of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ from *routine* uncalibrated sensor readings such as $\boldsymbol{y}$.

In general, without further assumptions, blind calibration appears to be an impossible task. However, it turns out that under mild assumptions that may often hold in practice, quite a bit can be learned from raw (uncalibrated) sensor readings like $\boldsymbol{y}$. Assume that the sensor network is slightly "oversampling" the phenomenon being sensed. Mathematically, this means that the

calibrated snapshot $\boldsymbol{x}$ lies in a lower dimensional subspace of n-dimensional Euclidean space. Let $\mathcal{S}$ denote this "signal subspace" and assume that it is r-dimensional, for some integer $0 < r < n$. For example, if the signal being measured is bandlimited and the sensors are spaced closer than required by the Shannon-Nyquist sampling rate, then $\boldsymbol{x}$ will lie in a lower dimensional subspace spanned by frequency basis vectors. If we oversample (relative to Shannon-Nyquist) by a factor of 2, then $r = n/2$. Basis vectors that correspond to smoothness assumptions, such as low-order polynomials, are another potentially relevant example. In general, the signal subspace may be spanned by an arbitrary set of r basis vectors. The calibration coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and the signal subspace $\mathcal{S}$ may change over time, but here we assume they do not change over the course of blind calibration. As we will see, this is a reasonable assumption, since the network may be calibrated from very few snapshots.

Let $\boldsymbol{P}$ denote the orthogonal projection matrix onto the orthogonal complement to the signal subspace $\mathcal{S}$. Then every $\boldsymbol{x} \in \mathcal{S}$ must satisfy the constraint

$$\boldsymbol{P}\boldsymbol{x} \;=\; \boldsymbol{P}(\boldsymbol{Y}\,\boldsymbol{\alpha} + \boldsymbol{\beta}) \;=\; 0 \tag{2}$$

This is the key idea behind our blind calibration method. Because the projection matrix $\boldsymbol{P}$ has rank $n - r$, the constraint above gives us $n - r$ linearly independent equations in $2n$ unknown values ($\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$). If we take snapshots from the sensor network at k distinct times, $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$, then we will have $k(n - r)$ equations in $2n$ unknowns. For $k \geq 2n/(n - r)$ we will have more equations than unknowns, which is a hopeful sign. In fact, as we will show, in many cases it is possible to blindly recover $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ from a sufficient number of uncalibrated sensor snapshots.

3 Initial Observations

3.1 Blind Calibration in No Noise

Given k snapshots at different time instants $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$, the subspace constraint (2) results in the following system of $k(n - r)$ equations:

$$\boldsymbol{P}(\boldsymbol{Y}_i\,\boldsymbol{\alpha} + \boldsymbol{\beta}) \;=\; 0\ ,\ i = 1, \ldots, k \tag{3}$$

The true gains and offsets must satisfy this equation, but in general the equation may be satisfied by other vectors as well. Establishing conditions

that guarantee that the true gains and/or offsets are the only solutions is the main theoretical contribution of this work.

It is easy to verify that the solutions for $\boldsymbol{\beta}$ satisfy

$$\boldsymbol{P}\boldsymbol{\beta} = -\boldsymbol{P}\bar{\boldsymbol{Y}}\boldsymbol{\alpha} \tag{4}$$

where $\bar{\boldsymbol{Y}} = \frac{1}{k}\sum_{i=1}^{k}\boldsymbol{Y}_i$, the time-average of the snapshots. One immediate observation is that the constraints only determine the components of $\boldsymbol{\beta}$ (in terms of the data and $\boldsymbol{\alpha}$) in the signal "nullspace" (the orthogonal complement to $\mathcal{S}$). The component of the offset $\boldsymbol{\beta}$ that lies in the signal subspace is unidentifiable. This is intuitively very easy to understand. Our only assumption is that the signals measured by the network lie in a lower dimensional subspace. The component of the offset in the signal subspace is indistinguishable from the *mean* or average signal. Recovery of this component of the offset requires extra assumptions, such as assuming that the signals have zero mean, or additional calibration resources, such as the non-blind calibration of some of the sensor offsets. We discuss this further in Section 5.

Given this characterization of the $\boldsymbol{\beta}$ solutions, we can re-write the constraints (3) in terms of $\boldsymbol{\alpha}$ alone:

$$\boldsymbol{P}(\boldsymbol{Y}_i - \bar{\boldsymbol{Y}})\boldsymbol{\alpha} \;=\; 0\;,\; i = 1, \ldots, k \tag{5}$$

If $\widehat{\boldsymbol{\alpha}}$ is a solution to this system of equations, then every vector $\boldsymbol{\beta}$ satisfying $\boldsymbol{P}\boldsymbol{\beta} = -\boldsymbol{P}\bar{\boldsymbol{Y}}\widehat{\boldsymbol{\alpha}}$ is a solution for $\boldsymbol{\beta}$ in the original system of equations (3). In other words, for a given $\widehat{\boldsymbol{\alpha}}$, the value of the component of the offset in the nullspace is $\boldsymbol{P}\bar{\boldsymbol{Y}}\widehat{\boldsymbol{\alpha}}$.

Another simple but very important observation is that there is one degree of ambiguity in $\boldsymbol{\alpha}$ that can never be resolved blindly using routine sensor measurements alone. The gain vector $\boldsymbol{\alpha}$ can be multiplied by a scalar c, and it cannot be distinguished whether this scalar multiple is part of the gains or part of the true signal. We call this scalar multiple the global gain factor. A constraint is needed to avoid this ambiguity, and without loss of generality we will assume that $\alpha(1) = 1$. This constraint can be interpreted physically to mean that we will calibrate all other sensors to the gain characteristics of sensor 1. The choice of sensor 1 is arbitrary and is taken here simply for convenience.

In Sections 4 and 5, we show that with noiseless measurements and perfect knowledge of the subspace $\boldsymbol{P}$, there is exactly one solution to 5 and it is possible to calibrate the gain calibration factors. We also show that we are able to calibrate the offset calibration factors with further information.

3.2 Blind Calibration in Noise

If noise, mismodeling effects, or other errors are present in the uncalibrated sensor snapshots, then a solution to (5) may not exist. There are many methods for finding the best possible solution, and in this work we employ singular value decomposition and standard least squares techniques.

First, note that the constraints can be expressed as

$$\mathbf{C}\,\boldsymbol{\alpha} \;=\; 0 \tag{6}$$

where the matrix $\mathbf{C}$ is given by

$$\mathbf{C} \;=\; \begin{bmatrix} \boldsymbol{P}(\boldsymbol{Y}_1 - \bar{\boldsymbol{Y}}) \\ \vdots \\ \boldsymbol{P}(\boldsymbol{Y}_k - \bar{\boldsymbol{Y}}) \end{bmatrix} \tag{7}$$

In the ideal case, there is always at least one solution[1] to the constraint $\mathbf{C}\boldsymbol{\alpha} = 0$, since the true gains must satisfy this equation. As we will show in Section 4, under certain conditions there will always be exactly one solution for $\boldsymbol{\alpha}$.

On the other hand, if the sensor measurements contain noise or if the assumed calibration model or signal subspace is inaccurate, then a solution may not exist. That is, the matrix $\mathbf{C}$ may have full column rank and thus will not have a right nullspace. A reasonable *robust* solution in such cases is to find the right singular vector of $\mathbf{C}$ associated with the smallest singular value. This vector is the solution to the following optimization.

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} \;=\; & \arg\min_{\boldsymbol{\alpha}} \|\mathbf{C}\boldsymbol{\alpha}\|_2^2 \\ & \text{subject to } \|\boldsymbol{\alpha}\|_2^2 = 1 \end{aligned} \tag{8}$$

In other words, we find the vector of gains such that $\mathbf{C}\widehat{\boldsymbol{\alpha}}$ is as close to zero as possible. This vector can be efficiently computed in numerical computing environments, such as Matlab, using the economy size singular value decomposition (svd)[2]. Note that in the ideal case (no noise or error) the svd solution satisfies (5). Thus, this is a general-purpose solution method.

Blind calibration of the gains can also be implemented by solving a system of equations in a least squared sense as follows. Recall that we have one constraint on our gain vector, $\boldsymbol{\alpha}(1) = 1$. This can be interpreted as

[1] This does not include the trivial solution $\boldsymbol{\alpha} = 0$. With no noise and perfect knowledge of $\boldsymbol{P}$, there is always one vector in the nullspace of $\mathbf{C}$.

[2] The Matlab command is svd$(\mathbf{C}, 0)$.

knowing the gain coefficient for the first sensor. We can use this knowledge as an additional constraint on the solution.

$$\begin{array}{rcc} \widehat{\boldsymbol{\alpha}} & = & \arg\min_{\boldsymbol{\alpha}} \|\mathbf{C}\boldsymbol{\alpha}\|_2^2 \\ & & \text{subject to } \boldsymbol{\alpha}(1) = 1 \end{array} \tag{9}$$

An equivalent optimization without constraints can be derived as follows. If we let $\mathbf{c_1}, \ldots, \mathbf{c_n}$ be the columns of $\mathbf{C}$, let $\tilde{\boldsymbol{\alpha}}$ be the gain vector with $\alpha(1)$ removed, and let $\tilde{\mathbf{C}}$ be the matrix $\mathbf{C}$ with the first column removed, we can rewrite the system of equations as $\tilde{\mathbf{C}}\,\tilde{\boldsymbol{\alpha}} = -\mathbf{c_1}$. The robust solution is the value of $\tilde{\boldsymbol{\alpha}}$ that minimizes the LS criterion $\|\tilde{\mathbf{C}}\,\tilde{\boldsymbol{\alpha}} + \mathbf{c_1}\|_2^2$.

More generally, we may know several of the gain coefficients for what we call *partially blind calibration*. Let $\mathbf{h}$ be the sum of the $\boldsymbol{\alpha}(i)\mathbf{c_i}$ corresponding to the known gains, let $\tilde{\boldsymbol{\alpha}}$ be the gain vector with the known gains $\boldsymbol{\alpha}(i)$ removed, and let $\tilde{\mathbf{C}}$ be the matrix $\mathbf{C}$ with those columns $\mathbf{c_i}$ removed. Now we have $\tilde{\mathbf{C}}\,\tilde{\boldsymbol{\alpha}} = -\mathbf{h}$ and the robust solution is the minimizer of

$$\|\tilde{\mathbf{C}}\,\tilde{\boldsymbol{\alpha}} + \mathbf{h}\|_2^2 \tag{10}$$

We can solve this optimization in a numerically robust manner by avoiding the squaring of the matrix $\tilde{\mathbf{C}}$ that is implicit in the conventional LS solution, $\tilde{\boldsymbol{\alpha}} = (\tilde{\mathbf{C}}'\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{C}}'(-\mathbf{h})$. This "squaring" effectively worsens the condition number of the problem and can be avoided by using QR decomposition techniques[3].

3.3 Estimation of the Subspace

The key component of the problem formulation proposed in Section 2 is the true signal subspace $\boldsymbol{P}$. In some cases, the phenomenon behavior has enough structure that the subspace can be derived. In other cases, the deviation of the true signals from an assumed subspace is small enough that, with a robust implementation of blind calibration, the results will still be good. Or, if a subspace does not change over time, it can be estimated early in the deployment with calibrated sensors and then used later to identify how the sensor calibration has drifted.

¿From experience, we have seen that there are further cases where blind calibration is sensitive to errors in the subspace. It is beyond the scope of this work to provide extensive tools for identifying the true signal subspace,

[3]We used $\boldsymbol{\alpha} = \tilde{\mathbf{C}}\backslash -\mathbf{h}$ in Matlab.

however here we briefly describe estimation of the subspace rank r from the measurements $y(i)$, $i = 1, \ldots, k$.

Given a particular measurement vector $\boldsymbol{y}$, we can subtract the mean of the k measurement vectors $\bar{\boldsymbol{y}}$ and get

$$\boldsymbol{y} - \bar{\boldsymbol{y}} = \frac{\boldsymbol{x} - \bar{\boldsymbol{x}}}{\boldsymbol{\alpha}} \tag{11}$$

where division of the vector is element-wise. Without loss of generality, we can assume that $\bar{\boldsymbol{x}} = 0$. Letting $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{\alpha})$ we can rewrite this as $\boldsymbol{y} = \boldsymbol{A}^{-1}\boldsymbol{x}$.

Each signal $\boldsymbol{x}$ lies in an r-dimensional subspace $\mathcal{S}$, where $r < n$ but is unknown. Let $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_r$ denote a basis for $\mathcal{S}$. Then $\boldsymbol{x} = \sum_{i=1}^{r} \theta_i \boldsymbol{\phi}_i$, for certain coefficients $\theta_1, \ldots, \theta_r$. Therefore, $\boldsymbol{y} = \sum_{i=1}^{r} \theta_i \boldsymbol{A}^{-1} \boldsymbol{\phi}_i$.

We can find the unknown r by building the covariance matrix of the measurements.

$$\mathbf{Cov}_{\boldsymbol{y}} = \frac{\mathbf{1}}{\mathbf{k}} \sum_{\mathbf{j=1}}^{\mathbf{k}} (\boldsymbol{y}_{\mathbf{j}} - \bar{\boldsymbol{y}})(\boldsymbol{y}_{\mathbf{j}} - \bar{\boldsymbol{y}})^{\mathbf{T}} \tag{12}$$

As long as $k > r$ and the vectors $\boldsymbol{x}_i$, $i = 1, \ldots, k$, are linearly independent, the rank of this covariance matrix will be r.

4 Gain Calibration

This section theoretically characterizes the existence of unique solutions to the gain calibration problem. As pointed out in Section 3.1, the gain calibration problem can be solved independently of the offset calibration task, as shown in (5), which corresponds to simply removing the mean snapshot from each individual snapshot. Therefore, it suffices to consider the case in which the snapshots are zero-mean and to assume that $\bar{\boldsymbol{Y}} = 0$, in which case the gain calibration equations may be written as

$$\boldsymbol{P}\, \boldsymbol{Y}_i\, \boldsymbol{\alpha} \;=\; 0 \;,\; i = 1, \ldots, k \tag{13}$$

The results we present also hold for the general case in which $\bar{\boldsymbol{Y}} \neq 0$. We first consider general conditions guaranteeing the uniqueness of the solution to (13) and then look more closely at the special case of bandlimited subspaces.

4.1 General Conditions

The following conditions are sufficient to guarantee that a unique solution to (13) exists.

A1. Oversampling: Each signal $\boldsymbol{x}$ lies in a known r-dimensional subspace $\mathcal{S}$, $r < n$. Let $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_r$ denote a basis for $\mathcal{S}$. Then $\boldsymbol{x} = \sum_{i=1}^{r} \theta_i \boldsymbol{\phi}_i$, for certain coefficients $\theta_1, \ldots, \theta_r$.

A2. Randomness: Each signal is randomly drawn from $\mathcal{S}$ and has mean zero. This means that the signal coefficients are zero-mean random variables. The joint distribution of these random variables is absolutely continuous with respect to Lebesgue measure (i.e., a joint r-dimensional density function exists). For any collection of signals $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, $k > 1$, the joint distribution of the corresponding kr coefficients is also absolutely continuous with respect to Lebesgue measure (i.e., a joint kr-dimensional density function exists).

A3. Incoherence: Define the $nr \times n$ matrix

$$\mathbf{M}_\Phi = \begin{bmatrix} \boldsymbol{P}\,\mathrm{diag}(\boldsymbol{\phi}_1) \\ \vdots \\ \boldsymbol{P}\,\mathrm{diag}(\boldsymbol{\phi}_r) \end{bmatrix} \tag{14}$$

and assume that $\mathrm{rank}(\mathbf{M}_\Phi) = n-1$. Note that $\mathbf{M}_\Phi$ is a function of the basis of the signal subspace. The matrix $\boldsymbol{P}$, the orthogonal projection matrix onto the orthogonal complement to the signal subspace $\mathcal{S}$, can be written as $\boldsymbol{P} = \mathbf{I} - \boldsymbol{\Phi}\boldsymbol{\Phi}'$, where $\mathbf{I}$ is the $n \times n$ identity matrix and $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_r]$.

Assumption A1 guarantees that the calibrated or true sensor measurements are correlated to some degree. This assumption is crucial since it implies that measurements must satisfy the constraints in (3) and that, in principle, we can solve for the gain vector $\boldsymbol{\alpha}$. Assumption A2 guarantees the signals are not too temporally correlated (e.g., different signal realizations are non-identical with probability 1). Also, the zero-mean assumption can be removed, as long as one subtracts the average from each sensor reading. Assumption A3 essentially guarantees that the basis vectors are sufficiently incoherent with the canonical sensor basis, i.e., the basis that forms the columns of the identity matrix. It is easy to verify that if the signal subspace basis is *coherent* with the canonical basis, then $\mathrm{rank}(\mathbf{M}_\Phi) < n - 1$. Also, note that $\mathbf{M}_\Phi \mathbf{1} = 0$, where $\mathbf{1} = [1, \ldots, 1]'$, which implies that $\mathrm{rank}(\mathbf{M}_\Phi)$ is at most $n - 1$. In general, assumption A3 only depends on the assumed signal subspace and can be easily checked for a given basis. In our experience, the condition is satisfied by most signal subspaces of practical interest, such as lowpass, bandpass or smoothness subspaces.

Theorem 1. *Under assumptions A1, A2 and A3, the gains* $\boldsymbol{\alpha}$ *can be perfectly recovered from any* $k \geq r$ *signal measurements by solving the linear system of equations (5).*

The theorem is proved in the Appendix. The theorem demonstrates that the gains are *identifiable* from routine sensor measurements; that is, in the absence of noise or other errors, the gains are perfectly recovered. In fact, the proof shows that under A1 and A2, the condition A3 is both necessary and sufficient. When noise and errors are present, the estimated gains may not be exactly equal to the true gains. However, as the noise/errors in the measurements tend to zero, the estimated gains tend to the true gains.

4.2 Bandlimited Subspaces

In the special case in which the signal subspace corresponds to a frequency domain subspace, a slightly more precise characterization is possible which shows that even fewer snapshots suffice for blind calibration. As stated above, assumption A3 is often met in practice and can be easily checked given a signal basis $\boldsymbol{\Phi}$. One case where A3 is automatically met is when the signal subspace is spanned by a subset of the Discrete Fourier Transform (DFT) vectors:

$$\boldsymbol{\phi}_m = [1, e^{\frac{-i2\pi m}{n}}, \ldots, e^{\frac{-i2(n-1)\pi m}{n}}]'/\sqrt{n}, \quad m = 0, \ldots, n-1$$

In this case we show that only $\lceil \frac{n-1}{n-r} \rceil + 1$ snapshots are required. This can be significantly less than r, meaning that the time over which we must assume that the subspace and calibration coefficients are unchanging is greatly reduced. The following assumptions and theorem summarize this result.

B1. Oversampling: Assume that each signal $\boldsymbol{x}$ lies in a bandlimited r-dimensional subspace $\mathcal{S}$, $r < n$, spanned by $\boldsymbol{\phi}_{m_1}, \ldots, \boldsymbol{\phi}_{m_r}$, where $m_1, \ldots, m_r$ are distinct integers from the set $\{0, \ldots, n-1\}$. Furthermore, assume that these integers are aperiodic in the following sense. Let $\mathbf{s}$ denote the vector with one at locations $m_1, \ldots, m_r$ and zero otherwise. This vector indicates the *support* set of the signal subspace in the DFT domain. The integers $m_1, \ldots, m_r$ are called aperiodic if every circular (mod n) shift of $\mathbf{s}$ is distinct. It is easy to check this condition and, in fact, most bandlimited subspaces have aperiodic support sets.

B2. Randomness: Note that each signal $\boldsymbol{x} \in \mathcal{S}$ can be written as $\boldsymbol{x} = \sum_{j=1}^{r} \theta_j \boldsymbol{\phi}_{m_j}$, for certain coefficients $\theta_1, \ldots, \theta_r$. Each signal is randomly drawn from $\mathcal{S}$ and has mean zero. This means that the

signal coefficients are zero-mean random variables. The joint distribution of these random variables is absolutely continuous with respect to Lebesgue measure (i.e., a joint r-dimensional density function exists). Also assume that multiple signal observations are statistically uncorrelated.

Assumption B1 guarantees that the calibrated or true sensor measurements are spatially correlated to some degree. As before, this assumption is crucial since it implies that measurements must satisfy the constraints in (3) and that, in principle, we can solve for the gain vector $\boldsymbol{\alpha}$. The rationale behind the assumption that the frequency support set is aperiodic is less obvious, but its necessity is due to the 2π-periodicity of the DFT (see [1] for further details). Assumption B2 guarantees the signals are not temporally correlated (analogous to A2, above). The following theorem characterizes the identifiability of the sensor gains in this situation.

Theorem 2. *It is necessary to make at least $k \geq \lceil \frac{n-1}{n-r} \rceil$ signal measurements in order to determine the gains, where $\lceil z \rceil$ denotes the smallest integer greater than or equal to z. Moreover, under assumptions B1 and B2, the gains $\boldsymbol{\alpha}$ can be perfectly recovered from any $k = \lceil \frac{n-1}{n-r} \rceil + 1$ signal measurements by solving the linear system of equations (5).*

The theorem is proved in the Appendix. The proof takes advantage of the special structure of the DFT basis. Alternatively, one could apply Theorem 1 in this case to obtain a slightly weaker result; namely that under B1 and B2 $k \geq r$ observations suffice to perfectly recover the gains.

5 Offset Calibration

The component of the offset in the signal subspace is generally unidentifiable, but in special cases it can be determined. For example, if it is known that the phenomenon of interest fluctuates symmetrically about zero (or some other known value), then the average of many measurements will tend to zero (or the known mean value). In this situation, the average

$$\frac{1}{k}\sum_{i=1}^{k} \boldsymbol{y}_i \;\; = \;\; \left(\frac{1}{k}\sum_{i=1}^{k} \boldsymbol{x}_i - \boldsymbol{\beta}\right) / \boldsymbol{\alpha} \;\approx\; -\boldsymbol{\beta}/\boldsymbol{\alpha}$$

where the the division operation is taken element-by-element. This follows since $\frac{1}{k}\sum_{i=1}^{k} \boldsymbol{x}_i \approx 0$ for large enough k. Thus we can identify the offset simply by calculating the average of our measurements. More precisely,

we can identify $\widetilde{\boldsymbol{\beta}} = \boldsymbol{\beta}/\boldsymbol{\alpha}$, which suffices since we can equivalently express the basic relationship (1) between calibrated and uncalibrated snapshots as $\boldsymbol{x} = (\boldsymbol{Y} + \widetilde{\boldsymbol{\beta}})\,\boldsymbol{\alpha}$.

Another situation in which we can determine (or partially determine) the component of the offset in the signal subspace is when we have knowledge of the correct offsets for a subset of the sensors. We call this *partially blind* offset calibration. Suppose that we are able to directly measure the offsets at $m < n$ sensors, indexed by m distinct integers $1 \leq \ell_1, \ldots, \ell_m \leq n$. Let $\boldsymbol{\beta}_m$ denote an $m \times 1$ vector these offsets. Let $\boldsymbol{T}$ be an $m \times n$ "selection" matrix that when applied to an arbitrary $n \times 1$ vector produces an $m \times 1$ vector of the elements at locations $\ell_1, \ldots, \ell_m$ from the original vector. With this notation, we can write $\boldsymbol{\beta}_m = \boldsymbol{T}\boldsymbol{\beta}$. Also note that

$$\boldsymbol{\beta}_m = \boldsymbol{T}\boldsymbol{\beta} = \boldsymbol{T}(\boldsymbol{P}\boldsymbol{\beta} + (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\beta}) = \boldsymbol{T}\boldsymbol{P}\boldsymbol{\beta} + \boldsymbol{T}(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\beta},$$

where $(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\beta}$ is the offset component in the signal subspace and $\boldsymbol{P}\boldsymbol{\beta}$ is the offset component in the orthogonal complement to the signal subspace.

As pointed out in Section 3.1, we can determine the component of the offset in the nullspace using $\boldsymbol{P}\boldsymbol{\beta} = -\boldsymbol{P}\bar{\boldsymbol{Y}}\widehat{\boldsymbol{\alpha}}$. Let us assume that this component is known (from the estimated calibration gains), and define $\boldsymbol{\beta}_\Delta = \boldsymbol{T}(\boldsymbol{I} - \boldsymbol{P})\boldsymbol{\beta}$, the signal subspace component of the offset at sensors $\ell_1, \ldots, \ell_m$. This component satisfies the relation

$$\boldsymbol{\beta}_\Delta \;=\; \boldsymbol{\beta}_m - \boldsymbol{T}\boldsymbol{P}\boldsymbol{\beta} \;=\; \boldsymbol{\beta}_m + \boldsymbol{T}\boldsymbol{P}\bar{\boldsymbol{Y}}\widehat{\boldsymbol{\alpha}}. \tag{15}$$

The projection matrix corresponding to the signal subspace, $(\boldsymbol{I} - \boldsymbol{P})$, can be written in terms of a set of orthonormal column vectors, $\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_r$, that span the signal subspace. Let $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 \cdots \boldsymbol{\phi}_r]$ denote an $n \times r$ matrix whose columns are the basis vectors. Then $(\boldsymbol{I} - \boldsymbol{P}) = \boldsymbol{\Phi}\boldsymbol{\Phi}'$ and so we can also write $\boldsymbol{\beta}_\Delta \;=\; \boldsymbol{T}\boldsymbol{\Phi}\boldsymbol{\Phi}'\boldsymbol{\beta}$. Note that the offset component in the signal subspace is completely determined by the r parameters $\boldsymbol{\theta} = \boldsymbol{\Phi}'\boldsymbol{\beta}$. Defining $\boldsymbol{\Phi}_T = \boldsymbol{T}\boldsymbol{\Phi}$, we can write $\boldsymbol{\beta}_\Delta \;=\; \boldsymbol{\Phi}_T\boldsymbol{\theta}$. If $\boldsymbol{\Phi}_T$ is invertible, then using (15) the parameters $\boldsymbol{\theta}$ can be uniquely determined by $\boldsymbol{\theta} = \boldsymbol{\Phi}_T^{-1}(\boldsymbol{\beta}_m + \boldsymbol{T}\boldsymbol{P}\bar{\boldsymbol{Y}}\widehat{\boldsymbol{\alpha}})$. Thus, if $\boldsymbol{\Phi}_T$ is invertible, then the complete offset vector $\boldsymbol{\beta}$ can be determined from the subset of offsets $\boldsymbol{\beta}_m$ and the estimated gains $\widehat{\boldsymbol{\alpha}}$. If $\boldsymbol{\Phi}_T$ has rank $q < r$, then we can determine the signal subspace offset component up to a remaining unidentifiable component in a smaller $r - q$ dimensional subspace of the signal subspace.

The rank of $\boldsymbol{\Phi}_T$ cannot be greater than m, the number of known sensor offsets, which shows that to completely determine the offset component in the signal subspace we require at least $m = r$ known offsets. In general,

knowing the offsets for an arbitrary subset of m sensors may not be sufficient (i.e., $\mathbf{\Phi}_T$ may not be invertible), but there are important special cases when it is. First note the $\mathbf{\Phi}$, by construction, has full rank r. Also note that the selection matrix $\boldsymbol{T}$ selects the m rows corresponding to the known calibration offsets and eliminates the remaining $n - m$ rows. So, we require that the elimination of any subset of $n - m$ rows of $\mathbf{\Phi}$ does not lead to a linearly dependent set of $(m \times 1)$ columns. This requirement is known as an *incoherence condition*, and it is satisfied as long as the signal basis vectors all have small inner products with the natural or canonical sensor basis ($n \times 1$ vectors that are all zero except for a single non-zero entry). For example, frequency vectors (e.g., Discrete Fourier Transform vectors) are known to satisfy this type of incoherence condition [5]. This implies that for subspaces of bandlimited signals, $\mathbf{\Phi}_T$ is invertible provided $m \geq r$.

6 Evaluation

In order to evaluate whether this theory of blind calibration is possible in practice, we explore its performance in simulation under both measurement noise and the mis-characterization of the projection matrix $\boldsymbol{P}$. Additionally, we show the performance of the algorithm on two temperature sensor datasets, one dataset from a controlled experiment where the sensors are measuring all the same phenomenon and thus lie in a 1-dimensional subspace, and the other from a deployment in a valley at a nature preserve called the James Reserve[4], where the true dimension of the spatial signal is unknown.

6.1 Simulations

To test the blind calibration methods on simulated data, we simulated both a field and snapshots of that field. We generated gain and offset coefficients, measurement noise, and most importantly, a projection matrix $\boldsymbol{P}$.

We simulated a smooth field by generating an 256×256 array of pseudorandom Gaussian noises (i.e., a white noise field) and then convolving it with the smooth impulse response function $h(i,j) = e^{(-s((i-l/2)^2+(j-l/2)^2)}$, $s > 0$. Figure 1 shows an example field with the smoothing parameter $s = 1$, which could represent a smoothly varying temperature field, for example. We simulated sensor measurements by sampling the field on a uniform 8×8 grid of $n = 64$ sensors. For gains, we drew uniformly from $\boldsymbol{\alpha} \in [0.5, 1.5]$ and

[4] `http://www.jamesreserve.edu`

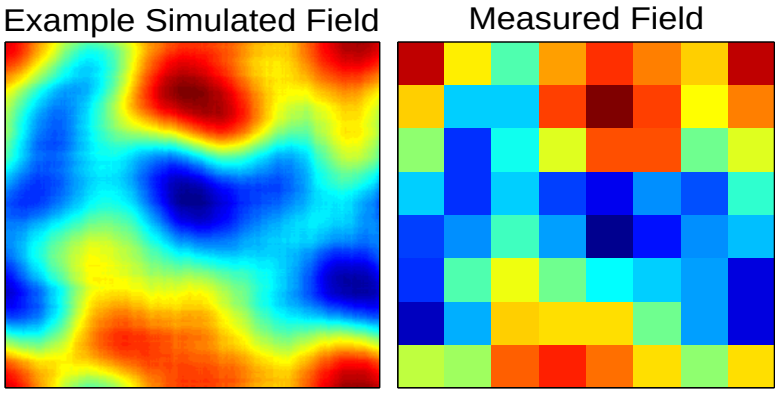

Figure 1: Two example simulated square fields. On the left, a 256×256 field generated with a basic smoothing kernel, which represents a true continuous field. On the right, an 8×8 grid of measurements of the same field. The fields can be quite dynamic and still meet the assumptions for blind calibration. The fields are shown in pseudocolor, with red denoting the maximum valued regions and blue denoting the minimum valued regions.

for offsets from $\boldsymbol{\beta} \in [-.5, .5]$. After applying $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to the measurements, we then added Gaussian noise, with mean zero and variance σ.

Separately, we created $\boldsymbol{P}$ to be a low-pass DFT matrix. We kept 3 frequencies in $2d$, which means with symmetries we have an $r = 49$-dimensional subspace[5]. With this setup, we can adjust the parameters of the smoothing kernel, while keeping $\boldsymbol{P}$ constant, to test robustness of blind calibration to an assumed subspace model that may over- or under-estimate the dimension of the subspace of the true field. The smoothing kernel and projection $\boldsymbol{P}$ both characterize lowpass effects, but the smoothing operator is only approximately described by the projection operator, even in the best case. We can also create our field by projecting the random field onto the r-dimensional subspace using $\boldsymbol{P}$; this represents the case where the true subspace is known exactly.

Estimates of the gains and offsets were calculated using the methods discussed above and described in more detail below. The graphs show error in gain, or equivalently in offset, as a fraction of the vector magnitude as follows:

[5] If the 2-dimensional signal has p frequencies, then the subspace is of rank $r = (2p+1)^2$.

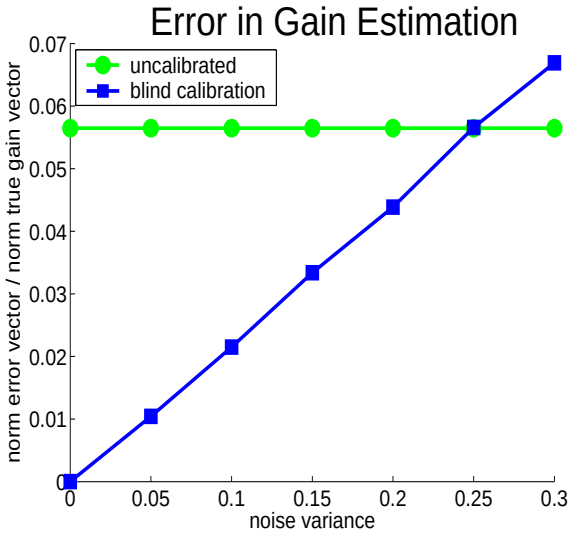

Figure 2: Gain error performance with exact knowledge of $\boldsymbol{P}$ and increasing measurement noise. Error averaged over 100 simulation runs.

$$err_\alpha \quad = \quad \frac{\|\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}\|_2}{\|\boldsymbol{\alpha}\|_2} \tag{16}$$

Error in the uncalibrated signal was calculated in the same way, assuming the estimate for gain, $\hat{\boldsymbol{\alpha}}$, is the all-ones vector and the estimate for offset, $\hat{\boldsymbol{\beta}}$, is the all-zeros vector. Error in the predicted signal was calculated by first applying the gain and offset factors to each measured signal $\boldsymbol{y}_i$ ($\hat{\boldsymbol{x}}_i = \boldsymbol{y}_i \circ \hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}}$ where $\circ$ is the hadamard product) and then calculating error as:

$$err_{\boldsymbol{x}_i} \quad = \quad \frac{\|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2}{\|\boldsymbol{x}_i\|_2}, \qquad i = 1, \ldots, k \tag{17}$$

Finally, the error was then averaged over the k measurements.

6.1.1 Error Results using SVD

We simulated blind calibration with the described simulation set-up. We first generated mean-zero fields using our smoothing kernel and took snapshot measurements of each field. We used $k = 5r$ snapshots (slightly more than the theoretical minimum of $k = r$) in order to improve our zero-mean

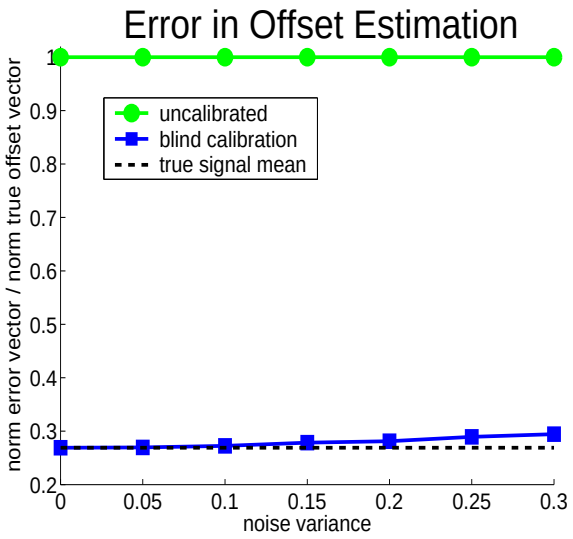

Figure 3: Offset error performance with exact knowledge of $\boldsymbol{P}$ and increasing measurement noise. Error averaged over 100 simulation runs.

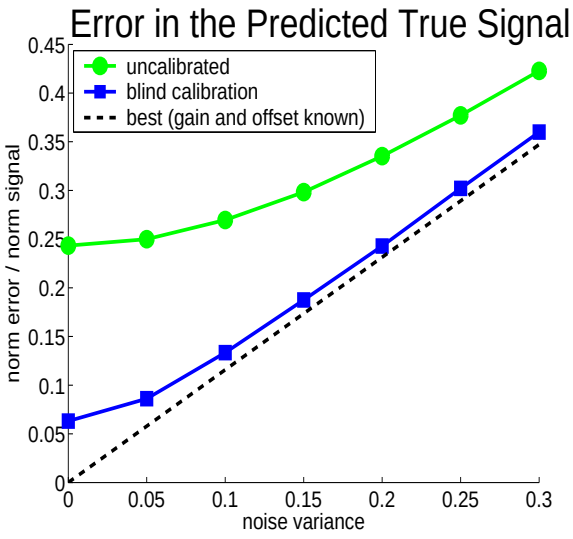

Figure 4: Error in the estimate of the signal x with exact knowledge of $\boldsymbol{P}$ and increasing measurement noise. Error averaged over 100 simulation runs.

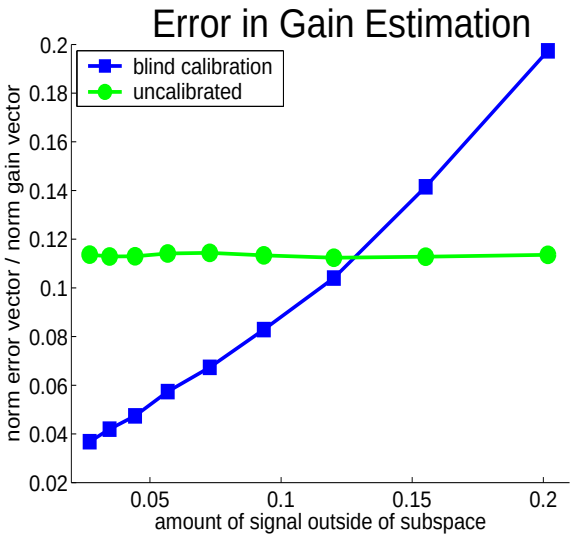

Figure 5: Gain error performance with uncertainty in the subspace $\boldsymbol{P}$. Error averaged over 100 simulation runs.

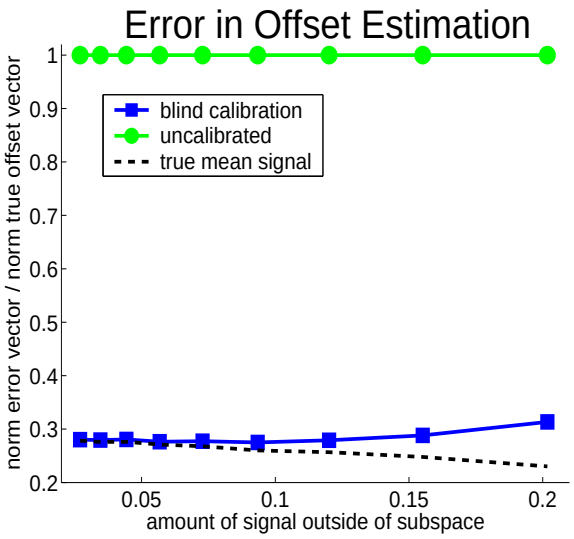

Figure 6: Offset error performance with uncertainty in the subspace $\boldsymbol{P}$. Error averaged over 100 simulation runs.

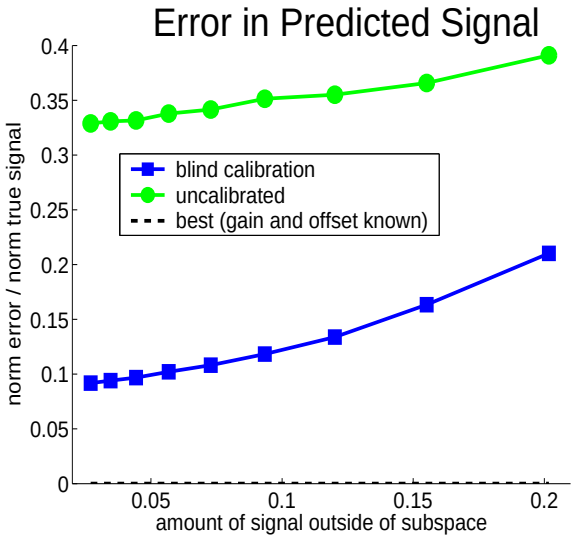

Figure 7: Error in the estimate of the signal x with with uncertainty in the subspace $\boldsymbol{P}$. Error averaged over 100 simulation runs.

signal assumption and to provide added robustness to noise and modeling errors. Then we constructed the matrix $\mathbf{C}$ from equation (7) and took the minimum right singular vector as the estimate of the gains $\boldsymbol{\alpha}$ as described in Section 3.2. We then estimated $\boldsymbol{\beta} = -\bar{\boldsymbol{Y}}\boldsymbol{\alpha}$.

Results from totally blind calibration in simulation, under the burden of increasing noise variance using exact knowledge of the subspace defined by $\boldsymbol{P}$, are shown in Figures 2, 3, and 4. That is, the fields in these simulations were created by projecting random signals into the space defined by projection matrix $\boldsymbol{P}$. The maximum value in the signals was 1, and therefore the noise variance can be taken as a percentage; i.e., variance of 10^{-2} represents 1% noise in the signal. Figure 2 shows the error in gain estimation as compared to the error in the uncalibrated signal. Figure 3 shows the error in offset estimation, along with the uncalibrated error and the energy of the mean true signal. Figure 4 shows error in the resulting estimate of the signals $\boldsymbol{x}$, after applying the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to $\boldsymbol{y}$ in order to recover $\hat{\boldsymbol{x}}$.

Blind calibration did very well in this scenario. Not until the noise in the signal was 25% did the uncalibrated gains outperform the blindly calibrated gains. As for offset, the blindly calibrated estimate was always much better than the uncalibrated, and in fact stayed very close to the

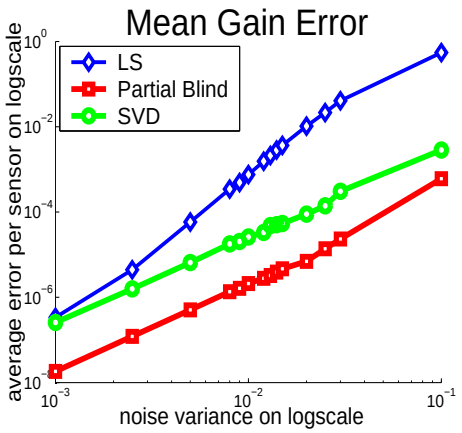

Figure 8: Gain error performance for SVD, blind LS, and partially blind LS. Results show mean error over 50 simulation runs.

theoretical minimum of the mean true signal. The combination is illustrated nicely in Figure 4, where we can see that the error in the blindly calibrated signal remains very close to the case where we know the corect gain and offset. The figures show mean error over 100 simulation runs.

Knowing the true subspace exactly is possible in practice only when performing blind calibration in a very well-known environment, such as an indoor factory. Even in this case, there will be some component of the true signals which is outside of the subspace defined by the chosen $\boldsymbol{P}$. Figure 7 shows how gain and offset error are affected by out-of-subspace components in the true signals. We used a basic smoothing kernel to control smoothness of the true field and kept $\boldsymbol{P}$ constant as described above with $r = 49$. The smoothing kernel and the projection operator are both low-pass operators, but even in the best case, some of the smoothed field will be outside of the space defined by the projection matrix $\boldsymbol{P}$. We defined the error in $\boldsymbol{P}$ as $\|\boldsymbol{x} - \boldsymbol{P}\boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2$. The x-axis value in the figure is the average error in $\boldsymbol{P}$ over 100 random fields smoothed with a given smoothness parameter. The figure shows mean and median error in gain and offset estimates over these 100 simulation runs. Again the results are compelling. The gain estimation error was around 10^{-2} even when 10% of the signal was outside of the subspace. The offset estimation as well was still very accurate, below 7×10^{-3} even when 20% of the signal was outside of the subspace.

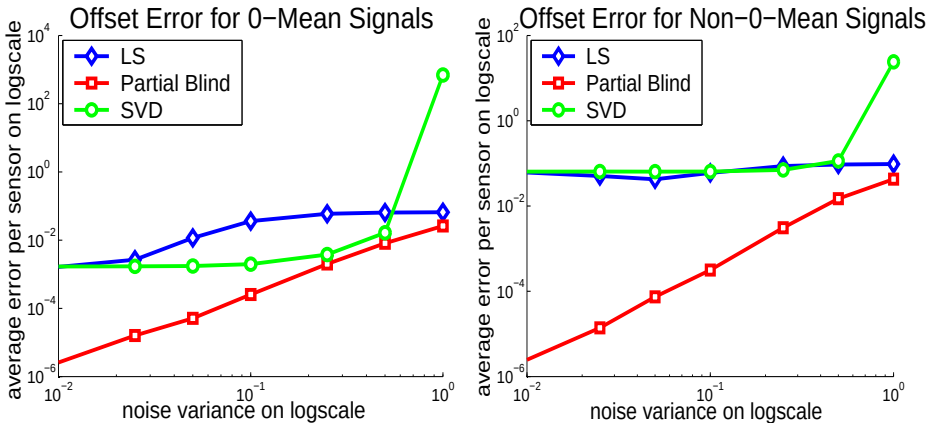

Figure 9: Offset error performance for SVD, blind LS, and partially blind LS. The top graph shows offset error for zero-mean signals, and the bottom graph is for non-zero-mean signals. Results show mean error over 50 simulation runs.

6.1.2 Comparison of Techniques

Here we compare the SVD technique to the LS technique and the totally blind calibration to partially blind calibration, where we know some of the calibration coefficients ahead of time.

The methods each work as follows. *SVD* performs gain estimation using the minimum right singular vector of the svd, i.e. by finding the solution to Equation (8), and normalizes assuming $\boldsymbol{\alpha}(1) = 1$. Offsets are then estimated using $\boldsymbol{\beta} = -\bar{\boldsymbol{Y}}\boldsymbol{\alpha}$. *Totally blind LS* performs gain estimation by solving Equation (9) in a least-squares sense and assuming knowledge only of $\boldsymbol{\alpha}(1) = 1$. Offsets are estimated as in SVD. *Partially blind LS* performs gain estimation by solving equation (10) in the least-squared sense, but now assuming we know some number of true gains and offsets. Offsets are then estimated as described in Section 5 for non-zero mean signals, i.e. using $\boldsymbol{\beta}_\Delta = \boldsymbol{T}\boldsymbol{\Phi}\boldsymbol{\Phi}'\boldsymbol{\beta}$ to solve for $\theta = \boldsymbol{\Phi}'\boldsymbol{\beta}$ and thus $\boldsymbol{\beta}$.

For partially blind LS we use enough of the true offsets such that we can solve for the complete component of $\boldsymbol{\beta}$ in the signal subspace. The fields we simulated are nearly bandlimited subspaces, and so the theory would imply that r true offsets are enough to estimate $\boldsymbol{\beta}$. In order to be robust to noise, we used knowledge of the offsets of $r + 5$ sensors, again slightly more than the bare minimum suggested by the theory.

A comparison of the techniques is quite interesting. First, as we expect,

the partially blind estimation does better than the other two methods in all cases; this follows from the fact that it is using more information. In Figure 8 you can see in the gain estimation, the SVD method out-performs totally blind LS, but partially blind LS has the lowest error of all the methods.

In the case of offset error, the SVD and totally blind LS techniques out-perform one another depending on the noise variance and whether or not the signals are zero-mean.

Figure 9 shows offset error for all three techniques. The partially blind LS method is unaffected by non-zero mean signals, which follows because method for estimating the offsets does not change with a zero-mean assumption. The other methods, on the other hand, capture the mean signal as part of their offset estimates, and as we can see, estimation error using the non-zero-mean signals is higher than using zero-mean signals.

The most intriguing part of these results is that totally blind LS performs slightly better than SVD for the offset estimate in non-zero-mean signals, despite the fact that it is using a gain estimate with more error from the first step in order to estimate the offsets. This implies that if calibration offset is the most important for calibration of your system, and you have non-zero-mean signals, you might prefer the totally blind LS method over the SVD.

6.2 Evaluation on Sensor Datasets

We evaluate blind calibration on two sensor network datasets, which we call the *calibration dataset* and the *cold air drainage transect dataset*.

6.2.1 Calibration Dataset

The *calibration dataset* was collected in September 2005 [2] along with data from a reference-caliber instrument in order to characterize the calibration of the thermistors used for environmental temperature measurement at the James Reserve. From the experiment, the conclusion was drawn that after the factory-supplied calibration was applied to the raw sensor measurements, the sensors differed from the reference thermocouple linearly, i.e. by only a gain and offset. Thus these sensors are suitable for evaluating the work we have done thus far on blind calibration. The data is available in the NESL CVS repository[6].

[6]This data is available at `http://www.ee.ucla.edu/~sunbeam/bc/`

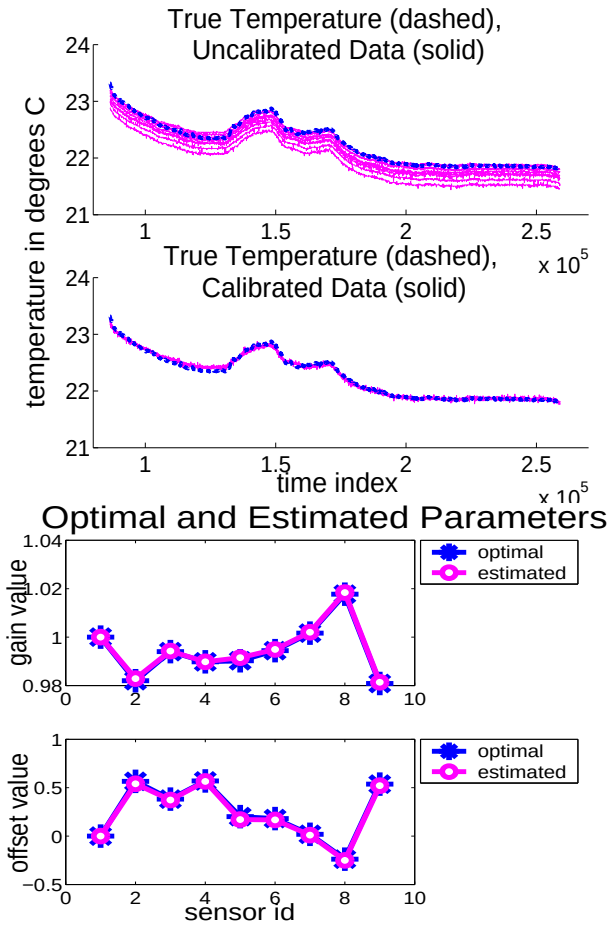

Figure 10: Results of blind calibration on the *calibration dataset*.

The setup of this experiment consisted of nine[7] temperature sensors. These sensors were placed in a styrofoam box along with a thermocouple attached to a datalogger, providing ground truth temperature readings. Therefore, all sensors were sensing the same phenomenon, and so the subspace spanned by the nine measurements is rank one. Thus, for $\boldsymbol{P}$ we used a lowpass dct matrix which kept only the dc frequency space. To illustrate, we used the following commands in Matlab:

```
r = 1; n = 9;
I = eye(n);
U = dct(I);
U(r+1:n,:) = 0;
P = idct(U);
```

We calibrated these data using snapshots from the dataset and the SVD method. Figure 10 shows the calibration coefficient estimates and reconstructed signals for the sensors in the experiment. The gains and offsets were recovered with very little error. The uppermost plot shows the data before and after calibration, along with the ground truth measurement in blue. The lower plot shows the true and estimated gains and offsets. This clearly demonstrates the utility of blind calibration.

6.2.2 Cold Air Drainage Dataset

The *cold air drainage transect dataset* consists of data from an ongoing deployment at the James Reserve. The deployment measures air temperature and humidity in a valley in order to characterize the predawn cold air drainage. The sensors used are the same as the sensors in the *calibration dataset*, and thus again the factory calibration brings them within an offset and gain of one another. The data we used for evaluation is from November 2, 2006, and it is available in the sensor data repository called SensorBase[8]. On this same day, we visited the James Reserve with a reference-caliber sensor and took measurements over the course of the day in order to get the true calibration parameters for comparison.

The deployment consists of 26 mica2 motes which run from one side of a valley to the other (Figure 11) across a streambed and in various regions of tree and mountain shade. Each mote has one temperature and one humidity sensor. For our purposes, we collected calibration coefficients from 10 of the temperature sensors.

[7] The experiment had ten sensors, one of which was faulty. In this analysis we used data from the nine functional sensors.

[8] `http://sensorbase.org`

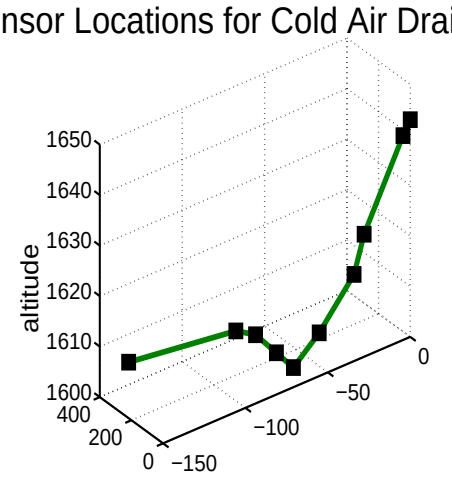

Figure 11: The mica2 motes in the cold air drainage transect run down the side of a hill and across a valley. The mote locations pictured are those that we used in the evaluation of blind calibration.

The signal subspace in this application does not correspond to a simple lowpass or smooth subspace, since sensors at similar elevations may have similar readings, but can be quite distant from each other. In principle, the signal subspace could be constructed based on the geographic positions and elevations of the sensor deployment. However, since we have the calibrated sensor data in this experiment, we can use these data directly to infer an approximate signal subspace. We constructed the projection $\boldsymbol{P}$ using the subspace associated with the four largest singular values of the calibrated signal data matrix.

We performed totally blind calibration using SVD. We constructed $\mathbf{C}$ using 64 snapshots taken over the course of the morning along with $\boldsymbol{P}$ as described. Figure 12 shows the results. The gain error was very small, only .0053 average per sensor, whereas if we were to assume the gain was 1 and not calibrate the sensors at all, the error would be .0180 average per sensor. On the other hand, the offset error was only slightly better with blind calibration than it would have been without: we saw .3953 average error per sensor as compared to 0.4610 error if the offsets were assumed to be zero. We believe that the offset estimation did not perform well due primarily to the fact that the mean signal is not zero in this case (e.g., the average sensor readings depend on elevation). Better offset estimates could be obtained using knowledge of one or more of the true sensor offset values.

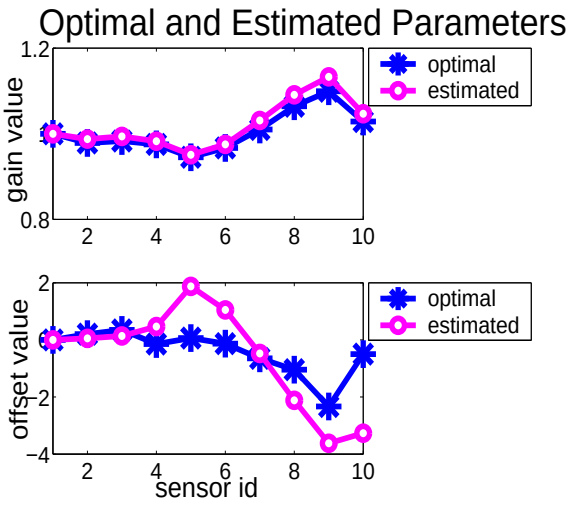

Figure 12: True and estimated gains and offsets for the *cold air drainage transect dataset*.

7 Related Work

The most straightforward approach to calibration is to apply a *known* stimulus $\boldsymbol{x}$ to the sensor network and measure the response $\boldsymbol{y}$. Then using the groundtruth input $\boldsymbol{x}$ we can adjust the calibration parameters so that (1) is achieved. We call this *non-blind* calibration, since the true signal $\boldsymbol{x}$ is known. This problem is called inverse linear regression; mathematical details can be found at [9]. Non-blind calibration is used routinely in sensor networks [11, 14], but may be difficult or impossible in many applications.

As for blind calibration in sensor networks, the problem of relating measurements such as received signal strength or time delay to distance for localization purposes has been studied extensively [10, 13]. This problem is quite different from the blind calibration problem considered in this chapter, which assumes that the measurements arise from external signals (e.g., temperature) and not from range measurements between sensors. In [15], the problem of calibrating sensor range measurements by enforcing geometric constraints in a system-wide optimization is considered. Calibration using geometric and physical constraints on the behavior of a point light source is considered in [6]. The constraint that proximal sensors in dense deployments make very similar measurements is leveraged in [4]. In this work, our constraint is simply that the phenomenon of interest lies in a subspace. This is a much more general constraint and hopefully therefore it can be widely

applicable.

Blind equalization and blind deconvolution [12] are related problems in signal processing. In these problems, the observation model is of the form $\boldsymbol{y} = \mathbf{h} * \boldsymbol{x}$, where $*$ is the convolution operator, and both $\mathbf{h}$ and $\boldsymbol{x}$ must be recovered from $\boldsymbol{y}$. Due to the difference between the calibration and convolution models, there are significant differences between blind deconvolution and blind calibration. However, in certain circumstances, blind calibration is mathematically equivalent to multi-channel blind deconvolution [7, 8]. Blind calibration involves observing multiple unknown signals through one unknown calibration function. Multi-channel blind deconvolution involves observing one unknown signal through multiple unknown channels. In the deconconvolution set-up, a common assumption is that the multiple channels have finite impulse response functions. This places the channels in a lower dimensional subspace (of the space of all impulse responses), and singular value decomposition techniques, similar to those proposed for blind calibration in this chapter, have been devised for solving the blind deconvolution problem in this setting [7, 8]. The blind calibration problem is mathematically equivalent to the multi-channel blind deconvolution problem, in the following special case. First observe that the multiplicative calibration gain can be expressed as a convolutional operation in the frequency domain. Now suppose that the signal subspace is a lowpass frequency subspace spanned by low-frequency DFT vectors. In this case, blind gain calibration is exactly equivalent to multi-FIR-channel blind deconvolution (the two problems are related by the DFT). In general, however, the signal subspace may not be a low-frequency subspace, and the two problems are quite different. In this sense, blind calibration is much more general than multi-channel blind deconvolution, but the relation between the two problems suggests that more sophisticated solution methods, such as IQML [8], might be applicable to blind calibration.

8 Extensions and Future Work

There are many issues in blind calibration that could be explored further. The two main areas ripe for study are the choice of the subspace $\boldsymbol{P}$ and the implementation of blind calibration. There are many possible choices for a suitable subspace, including frequency subspaces and smoothness subspaces. How to choose the subspace when faced with a sensor deployment where the true signals are unknown is an extremely important question for blind calibration. Methodologies for creating a $\boldsymbol{P}$ would be extremely useful to

the more general application of blind calibration, especially ones which could incorporate trusted measurements or the users' knowledge of the physical space where the sensors are deployed. At the same time, implementations of blind calibration that are robust to model error in the subspace would allow users to be more liberal in the choice of $\boldsymbol{P}$.

The theoretical analysis shown here is done under noiseless conditions and with a perfect model. Future work includes both noisy analysis to find analytical bounds that can be compared to simulation results and sensitivity analysis for our system of linear equations. Our experience is that solutions are robust to noise and mismodeling in some cases, and sensitive in others; we do not have a good understanding of the robustness of the methodology at this time.

Extending the formulation to handle non-linear calibration functions would be useful in cases where a raw non-linear sensor response must be calibrated. We believe that many of the techniques developed in this work can be extended to more general polynomial-form calibration functions. Other interesting topics include distributed blind calibration and blind calibration in the presence of faulty sensors.

9 Conclusions

The problem of sensor calibration is central to the practical use of sensor networks. The blind calibration formulation and methods developed in this work use only routine sensor measurements, and thus give an extremely promising formulation for the mass calibration of sensors. We have shown that calibration gains are identifiable. We have proved how many measurements are necessary and sufficient to estimate the gain factors, and we have shown necessary and sufficient conditions to estimate the offsets. We have demonstrated a working implementation on simulated and real data, which uncovered interesting relationships between implementation and blind calibration performance. Overall, we have demonstrated that blind calibration has great potential to be possible in practice, and we feel that the proposed formulation merits further investigation.

References

[1] L. Balzano and R. Nowak. Blind calibration for signals with bandlimited subspaces. Technical report, Information Sciences Laboratory at the University of Wisconsin-Madison, February 2007.

[2] L. Balzano, N. Ramanathan, E. Graham, M. Hansen, and M. B. Srivastava. An investigation of sensor integrity. Technical Report UCLA-NESL-200510-01, Networked and Embedded Systems Laboratory, 2005.

[3] P. Buonadonna, D. Gay, J. Hellerstein, W. Hong, and S. Madden. Task: Sensor network in a box. Technical Report IRB-TR-04-021, Intel Research Berkeley, January 2005.

[4] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak. A collaborative approach to in-place sensor calibration. *Lecture Notes in Computer Science*, 2634:301–316, 2003.

[5] E. Candes and J. Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 2006.

[6] J. Feng, S. Megerian, and M. Potkonjak. Model-based calibration for sensor networks. *Sensors*, pages 737 – 742, October 2003.

[7] M. Gurelli and C. Nikias. Evam: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals. *IEEE Transactions on Signal Processing*, 43:134–149, January 1995.

[8] G. Harikumar and Y. Bresler. Perfect blind restoration of images blurred by multiple filters: Theory and efficient algorithms. *IEEE Transactions on Image Processing*, 8(2):202 – 219, February 1999.

[9] B. Hoadley. A bayesian look at inverse linear regression. *Journal of the American Statistical Association*, 65(329):356 – 369, March 1970.

[10] A. Ihler, J. Fisher, R. Moses, and A. Willsky. Nonparametric belief propagation for self-calibration in sensor networks. In *Proceedings of the Third International Symposium on Information Processing in Sensor Networks*, 2004.

[11] N. Ramanathan, L. Balzano, M. Burt, D. Estrin, T. Harmon, C. Harvey, J. Jay, E. Kohler, S. Rothenberg, and M.Srivastava. Rapid deployment with confidence: Calibration and fault detection in environmental sensor networks. Technical Report CENS TR 62, Center for Embedded Networked Sensing, 2006.

[12] O. Shalvi and E. Weinstein. New criteria for blind deconvolution of nonmimimum phase systems (channels). *IEEE Trans. on Information Theory*, IT-36(2):312 – 321, March 1990.

[13] C. Taylor, A. Rahimi, J. Bachrach, H. Shrobe, and A. Grue. Simultaneous localization, calibration, and tracking in an ad hoc sensor network. In *IPSN '06: Proceedings of the Fifth International Conference on Information Processing in Sensor Networks*, pages 27–33, 2006.

[14] G. Tolle, J. Polastre, R. Szewczyk, D. Culler, N. Turner, K. Tu, S. Burgess, T. Dawson, P. Buonadonna, D. Gay, and W. Hong. A macroscope in the redwoods. In *Proceedings of Sensys*, 2005.

[15] K. Whitehouse and D. Culler. Calibration as parameter estimation in sensor networks. In *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pages 59–67, 2002.

[16] W. M. Wonham. *Linear Multivariable Control*. Springer-Verlag, New York, 1979.

Theorem 1:

Proof. First note that the case where the signal subspace is one-dimensional ($r = 1$) is trivial. In this case there is one degree of freedom in the signal, and hence one measurement coupled with the constraint that $\alpha(1) = 1$ suffices to calibrate the system. For the rest of the proof we assume that $1 < r < n$ and thus $2 \leq k < n$.

Given k signal observations $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$, and letting $\widehat{\boldsymbol{\alpha}}$ represent our estimated gain vector, we need to show that the system of equations

$$\begin{bmatrix} \boldsymbol{P}\boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{P}\boldsymbol{Y}_k \end{bmatrix} \widehat{\boldsymbol{\alpha}} = 0 \tag{18}$$

has rank $n-1$, and hence may be solved for the $n-1$ degrees of freedom in $\widehat{\boldsymbol{\alpha}}$. Note each subsystem of equations, $\boldsymbol{P}\boldsymbol{Y}_j$, has rank less than or equal to $n-r$ (since $\boldsymbol{P}$ is rank $n-r$). Therefore, if $k < \frac{n-1}{n-r}$, then the system of equations certainly has rank less than $n-1$. This implies that it is necessary that $k \geq \frac{n-1}{n-r}$. Next note that $\boldsymbol{Y}_j = \boldsymbol{X}_j\boldsymbol{A}$, where $\boldsymbol{X}_j = \mathrm{diag}(\boldsymbol{x}_j)$ and $\boldsymbol{A} = \mathrm{diag}([1, 1/\alpha(2), \ldots, 1/\alpha(n)]')$. Then write

$$\begin{bmatrix} \boldsymbol{P}\boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{P}\boldsymbol{X}_k \end{bmatrix} \boldsymbol{d} = 0 \tag{19}$$

where $\boldsymbol{d} = \boldsymbol{A}\widehat{\boldsymbol{\alpha}}$. The key observation is that satisfaction of these equations requires that $\boldsymbol{X}_j\boldsymbol{d} \in \mathcal{S}$, for $j = 1, \ldots, k$. Any $\boldsymbol{d}$ that satisfies this relationship

will imply a particular solution for $\widehat{\boldsymbol{\alpha}}$, and thus $\boldsymbol{d}$ must not be any vector other than the all-ones vector for blind calibration to be possible.

Recall that by definition $\boldsymbol{X}_j = \text{diag}(\boldsymbol{x}_j)$. Also note that $\text{diag}(\boldsymbol{x}_j)\boldsymbol{d} = \text{diag}(\boldsymbol{d})\boldsymbol{x}_j$. So we can equivalently state the requirement as

$$\text{diag}(\boldsymbol{d})\boldsymbol{x}_j \in \mathcal{S},\ j = 1, \ldots, k\ . \tag{20}$$

The proof proceeds in two steps. First, A2 implies that $k \geq r$ signal observations will span the signal subspace with probability 1. This allows us to re-cast the question in terms of a basis for the signal subspace, rather than particular realizations of signals. Second, it is shown that A3 (in terms of the basis) suffices to guarantee that the system of equations has rank $n-1$.

Step 1: We will show that all solutions to (20) are contained in the set

$$\mathcal{D}\ =\ \{\boldsymbol{d}\ :\text{diag}(\boldsymbol{d})\boldsymbol{\phi}_i \in \mathcal{S},\ i = 1, \ldots, r\}.$$

We proceed by contradiction. Suppose that there exists a vector $\widetilde{\boldsymbol{d}}$ that satisfies (20) but does not belong to $\mathcal{D}$. Since $\widetilde{\boldsymbol{d}}$ satisfies (20), we know that there exists an $\boldsymbol{x} \in \mathcal{S}$ such that $\text{diag}(\widetilde{\boldsymbol{d}})\boldsymbol{x} \in \mathcal{S}$. We can write $\boldsymbol{x}$ in terms of the basis, as $\boldsymbol{x} = \sum_{i=1}^{r} \theta_i \boldsymbol{\phi}_i$, and $\text{diag}(\widetilde{\boldsymbol{d}})\boldsymbol{x} = \sum_{i=1}^{r} \theta_i \,\text{diag}(\widetilde{\boldsymbol{d}})\boldsymbol{\phi}_i$. Since by assumption $\widetilde{\boldsymbol{d}}$ does not satisfy $\text{diag}(\widetilde{\boldsymbol{d}})\boldsymbol{\phi}_i \in \mathcal{S}$, $i = 1, \ldots, r$, it follows that the coefficients $\theta_1, \ldots, \theta_r$ must weight the components outside of the signal subspace so that they cancel out. In other words, the set of signals $\boldsymbol{x} \in \mathcal{S}$ that satisfy $\text{diag}(\widetilde{\boldsymbol{d}})\boldsymbol{x} \in \mathcal{S}$ is a proper subspace (of dimension less than r) of the signal subspace $\mathcal{S}$. However, if we make $k \geq r$ signal observations, then with probability 1 they collectively span the entire signal subspace (since they are jointly continuously distributed). In other words, the probability that all k measurements lie in a lower dimensional subspace of $\mathcal{S}$ is zero. Thus, $\widetilde{\boldsymbol{d}}$ cannot be a solution to (20).

Step 2: Now we characterize the set $\mathcal{D}$. First, observe that the vectors $\boldsymbol{d} \propto \mathbf{1}$, the constant vector, are contained in $\mathcal{D}$, and those correspond to the global gain factor ambiguity discussed earlier. Second, note that every $\boldsymbol{d} \in \mathcal{D}$ must satisfy $\boldsymbol{P}\,\text{diag}(\boldsymbol{d})\boldsymbol{\phi}_i = \boldsymbol{P}\,\text{diag}(\boldsymbol{\phi}_i)\boldsymbol{d} = 0$, $i = 1, \ldots, r$, where $\boldsymbol{P}$ denote the projection matrix onto the orthogonal complement to the signal subspace $\mathcal{S}$. Using the definition of $\mathbf{M}_\Phi$ given in (14), we have the following equivalent condition: every $\boldsymbol{d} \in \mathcal{D}$ must satisfy $\mathbf{M}_\Phi \boldsymbol{d} = 0$. We know that the vectors $\boldsymbol{d} \propto \mathbf{1}$ satisfy this condition. The condition $\text{rank}(\mathbf{M}_\Phi) = n-1$ guarantees that these are the only solutions. This completes the proof. □

Theorem 2:

Proof. First note again that the theorem is trivial if the signal subspace is one-dimensional ($r = 1$), since in this case there is one degree of freedom in the signal, and hence one measurement (coupled with the constraint that $\alpha(1) = 1$) suffices to calibrate the system. For the rest of the proof we assume that $1 < r < n$ and thus $2 \leq k < n$.

As in the proof of Theorem 1, solutions must satisfy (19), or equivalently the equations $\boldsymbol{x}_j \bullet \boldsymbol{d} \in \mathcal{S}$, for $j = 1, \ldots, k$. Let $\boldsymbol{x}$ denote an arbitrary signal vector, and let $\boldsymbol{z} = \boldsymbol{x} \bullet \boldsymbol{d}$. We can express $\boldsymbol{z}$ in terms of the representation $\boldsymbol{x}$ in the basis of $\mathcal{S}$ as

$$\boldsymbol{z} = \sum_{j=1}^{r} \theta_j (\boldsymbol{\phi}_{m_j} \bullet \boldsymbol{d}).$$

Recall that multiplication in the time domain is equivalent to (circular) convolution in the DFT domain. Let $\boldsymbol{Z}$, $\boldsymbol{X}$, and $\boldsymbol{D}$ be $n \times 1$ vectors denoting the DFTs of $\boldsymbol{z}$, $\boldsymbol{x}$ and $\boldsymbol{d}$, respectively (e.g., $\boldsymbol{Z}(\ell) = \frac{1}{\sqrt{n}} \sum_{q=1}^{n} \boldsymbol{z}(q) e^{-j\frac{2\pi}{n}q\ell}$). Note that $\boldsymbol{X}(\ell) = \sum_{j=1}^{r} \theta_j \delta(\ell - m_j)$, $\ell = 0, \ldots, n-1$, where $\delta(k) = 1$ if $k = 0$ and 0 otherwise. Then $\boldsymbol{Z}$ is the circular convolution of $\boldsymbol{D}$ and $\boldsymbol{X}$; i.e., the ℓth element of $\boldsymbol{Z}$ is given by

$$\boldsymbol{Z}(\ell) \;\; = \;\; \sum_{q=1}^{n} \boldsymbol{D}(q) \boldsymbol{X}([\ell - q]_n)$$

where $[\ell]_n$ is equal to ℓ mod n. Hence, $\boldsymbol{z} \in \mathcal{S}$ if and only if the support of $\boldsymbol{Z}$ is on the set of frequencies $m_1, \ldots, m_r$.

For each $\ell = 1, \ldots, n$, let $\boldsymbol{X}_\ell$ denote the $n \times 1$ vector with entries $\boldsymbol{X}_\ell(q) = \boldsymbol{X}([\ell - q]_n)$, $q = 0, \ldots, 1$ (i.e., $\boldsymbol{X}_\ell$ is obtained by reversing $\boldsymbol{X}$ and circularly shifting the result by ℓ). Then we can write $\boldsymbol{Z}(\ell) = \boldsymbol{X}_\ell' \boldsymbol{D}$. Thus, we can express the constraint on the support of $\boldsymbol{Z}$ as follows:

$$\boldsymbol{X}_\ell' \boldsymbol{D} = 0, \quad \text{for all } \ell \neq m_1, \ldots, m_r \tag{21}$$

Notice that this places $n - r$ constraints on the vector $\boldsymbol{D}$. Also observe that the $n - r$ row vectors $\boldsymbol{X}_\ell'$, $\ell \neq m_1, \ldots, m_r$, correspond to an $(n - r) \times n$ submatrix of the circulant matrix

$$\boldsymbol{\Xi} = [\boldsymbol{X}_0'; \boldsymbol{X}_1'; \cdots; \boldsymbol{X}_{n-1}'] \tag{22}$$

(circulant because each row $\boldsymbol{X}_\ell$ is a circularly shifted version of the others). Furthermore, because signal coefficients $\theta_1, \ldots, \theta_r$ are randomly distributed according to B2, $\boldsymbol{\Xi}$ has full rank. This follows by recalling that circulant matrices are diagonalized by the DFT, and the eigenvalues of a circulant

matrix are equal to the DFT of the first row. The first row of $\boldsymbol{\Xi}$ is $\boldsymbol{X}_0$ (indexed-reversed $\boldsymbol{X}$). It is a simple exercise to see that the DFT of $\boldsymbol{X}_0$ reproduces the original signal $\boldsymbol{x}$. Since the non-zero DFT coefficients of $\boldsymbol{x}$ are randomly distributed according to a continuous density, the elements of the $\boldsymbol{x}$ are non-zero with probability 1. This implies that the eigenvalues of $\boldsymbol{\Xi}$ are non-zero with probability 1, and thus $\boldsymbol{\Xi}$ is full-rank. Consequently, the $n-r$ constraint equations in (21) are linearly independent. Also note that $\boldsymbol{D} \propto [1, 0, \ldots, 0]'$ (DFT of the constant vector) satisfies (21), so in addition to the one degree of freedom due to the intrinsic ambiguity of the global gain factor, there are $r-1$ other degrees of freedom remaining in the solutions to (21).

Now suppose that we make k signal observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, randomly drawn according to B2. Each signal produces a system of constraints of the form in (21). Let $\boldsymbol{X}_{j,0}$, denote the indexed-reversed DFT of $\boldsymbol{x}_j$, $j = 1, \ldots, k$. These vectors generate the first row of k matrices denoted $\boldsymbol{\Xi}_j$, $j = 1, \ldots, k$ (each defined analagously to $\boldsymbol{\Xi}$ above). Note that each vector $\boldsymbol{X}_{j,0}$ displays the same sparsity pattern (since all signals are assumed to lie in an r-dimensional DFT subspace and each vector has at most r non-zero entries). Since B2 assumes that the coefficients of each signal are uncorrelated, it follows that any subset of no more than r of the vectors $\{\boldsymbol{X}_{j,0}\}_{j=1}^{k}$ is a linearly independent set. Now consider the collective constraints generated by the k signal measurements:

$$\boldsymbol{X}'_{j,\ell}\boldsymbol{D} = 0, \quad \text{for all } \ell \neq m_1, \ldots, m_r \ \text{ and } j = 1 \ldots, k \tag{23}$$

These constraints can be expressed in matrix notation by letting $\widetilde{\boldsymbol{\Xi}_j}$ be the $(n-r) \times n$ submatrix obtained by retaining the $n-r$ rows of $\boldsymbol{\Xi}_j$ satisfying $\ell \neq m_1, \ldots, m_r$. Then let $\widetilde{\boldsymbol{\Xi}} = [\widetilde{\boldsymbol{\Xi}_1}; \cdots; \widetilde{\boldsymbol{\Xi}_k}]$. Then (23) can be written as

$$\widetilde{\boldsymbol{\Xi}}\,\boldsymbol{D} = \boldsymbol{0} \tag{24}$$

We know that $\boldsymbol{D} \propto [1, 0, \ldots, 0]'$ satisfies (23), so the number of linearly independent equations above can be at most $n-1$. In follows that the first column of $\widetilde{\boldsymbol{\Xi}}$ is zero, and thus we may eliminate the first element of the vectors $\boldsymbol{D}$ and the first column of $\widetilde{\boldsymbol{\Xi}}$. Define $\bar{\boldsymbol{D}}$ by removing the first element of $\boldsymbol{D}$, $\bar{\boldsymbol{\Xi}}_j$ by removing the first column from $\widetilde{\boldsymbol{\Xi}_j}$, and $\bar{\boldsymbol{\Xi}} = [\bar{\boldsymbol{\Xi}}_1; \cdots; \bar{\boldsymbol{\Xi}}_k]$. The constraints can be written as

$$\bar{\boldsymbol{\Xi}}\,\bar{\boldsymbol{D}} = \boldsymbol{0} \tag{25}$$

and we wish to show that $\bar{\boldsymbol{D}} = \boldsymbol{0}$ is the only solution (i.e., $\bar{\boldsymbol{\Xi}}$ is full rank). The matrix dimensions imply that $\text{rank}(\bar{\boldsymbol{\Xi}}) \leq \min\{k(n-r), n-1\}$, so choosing

$k \geq (n-1)/(n-r)$ is a necessary condition. The necessity of the condition that the integers (frequencies) $m_1, \ldots, m_r$ are aperiodic (see B1) can also be seen at this point. Suppose for the sake of contradiction that the frequencies were not aperiodic. Then, because the support set of one row can align with another (at a different circular shift), one of the columns of $\bar{\boldsymbol{\Xi}}$ is the zero vector, and thus rank($\bar{\boldsymbol{\Xi}}$) would be less than $n-1$.

Now we show that $k = \lceil (n-1)/(n-r) \rceil + 1$ signal measurements suffice to recover the gains. To prove that $\bar{\boldsymbol{\Xi}}$ is full rank in this case, it suffices to show that the nullspaces of $\bar{\boldsymbol{\Xi}}_j$ are disjoint. Without loss of generality, we consider the case of a null vector of $\bar{\boldsymbol{\Xi}}_1$. Let us denote this vector by $\boldsymbol{v}$. We will show that this vector is not in the nullspace of the other submatrices $\bar{\boldsymbol{\Xi}}_j$, $j = 2, \ldots, k$. Define $\bar{\boldsymbol{\Xi}}_{/1} = [\bar{\boldsymbol{\Xi}}_2; \cdots; \bar{\boldsymbol{\Xi}}_k]$. The non-zero entries of the matrix $\bar{\boldsymbol{\Xi}}_{/1}$ are the (random) DFT coefficients from the $k-1$ signal observations $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$, and these are independent of $\boldsymbol{v}$, which depends only on $\boldsymbol{x}_1$. By assumption B2, these coefficients are continuous random variables. Consider the random variable $p = \boldsymbol{v}' \widetilde{\boldsymbol{\Xi}}'_{/1} \widetilde{\boldsymbol{\Xi}}_{/1} \boldsymbol{v}$. Treating $\boldsymbol{v}$ as a fixed vector, the variable p is a quadratic polynomial function of the random DFT coefficients of $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$. There are two distinct possibilities. Either p is the zero function, or p is a non-zero polynomial function. Suppose p is the zero function, then the conditional expectation of p given $\boldsymbol{v}$ satisfies $E[p|\boldsymbol{v}] = 0$. However, note that $E[p|\boldsymbol{v}] = \boldsymbol{v}' E[\bar{\boldsymbol{\Xi}}'_{/1} \bar{\boldsymbol{\Xi}}_{/1}] \boldsymbol{v}$, and it is easy to verify that the matrix $E[\bar{\boldsymbol{\Xi}}'_{/1} \bar{\boldsymbol{\Xi}}_{/1}]$ is full rank as follows. The matrix has a block structure, with each block of having the form $E[\bar{\boldsymbol{\Xi}}'_i \bar{\boldsymbol{\Xi}}_j]$. The blocks corresponding to the same signal observation (i.e., $i = j$) are full rank because of the circulant matrix property discussed above. The blocks corresponding to two different signal observations ($i \neq j$) are exactly zero since the signals are assumed to be uncorrelated with each other and zero mean. Together these to observations show that $E[\bar{\boldsymbol{\Xi}}'_{/1} \bar{\boldsymbol{\Xi}}_{/1}]$ is full rank. Therefore, $E[p|\boldsymbol{v}] > 0$ for every non-zero vector $\boldsymbol{v}$, and it follows that p cannot be equal to the zero function. However, if p is a non-zero polynomial function, then the probability that $p = 0$ is 0, implying that $\bar{\boldsymbol{\Xi}}_{/1} \boldsymbol{v} \neq 0$. This last argument follows from the well-known fact that the probability measure of the set of zeros of a polynomial function of continuous random variables is exactly zero [16]. Thus, we have shown that with probability 1, $\bar{\boldsymbol{\Xi}} \boldsymbol{v} \neq 0$ for every $\boldsymbol{v} \neq 0$, concluding the proof. □