

PAPER • OPEN ACCESS

Application of kernel principal component analysis for optical vector atomic magnetometry

To cite this article: James A McKelvy *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 045048

View the [article online](#) for updates and enhancements.

You may also like

- [Transformer fault diagnosis based on MPA-RF algorithm and LIF technology](#)
Pengcheng Yan, JingBao Wang, Wenchang Wang et al.
- [Solar Flare Predictive Features Derived from Polarity Inversion Line Masks in Active Regions Using an Unsupervised Machine Learning Algorithm](#)
Jingjing Wang, Yuhang Zhang, Shea A. Hess Webber et al.
- [Mixed kernel principal component weighted regression based on just-in-time learning for soft sensor modeling](#)
Shulong Yin, Yonggang Li, Bei Sun et al.



PAPER

OPEN ACCESS

RECEIVED
1 June 2023REVISED
2 October 2023ACCEPTED FOR PUBLICATION
22 November 2023PUBLISHED
8 December 2023

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Application of kernel principal component analysis for optical vector atomic magnetometry

James A McKelvy^{1,*} , Irina Novikova² , Eugeny E Mikhailov² , Mario A Maldonado² , Isaac Fan³ , Yang Li³, Ying-Ju Wang³, John Kitching³ and Andrey B Matsko¹

¹ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, United States of America

² Physics Department, College of William & Mary, Williamsburg, VA 23187, United States of America

³ National Institute of Standards and Technology, Boulder CO 80305, United States of America

* Author to whom any correspondence should be addressed.

E-mail: james.a.mckelvy@jpl.nasa.gov

Keywords: kernel principal component analysis, vector magnetometry, unsupervised machine learning, support vector regression machines

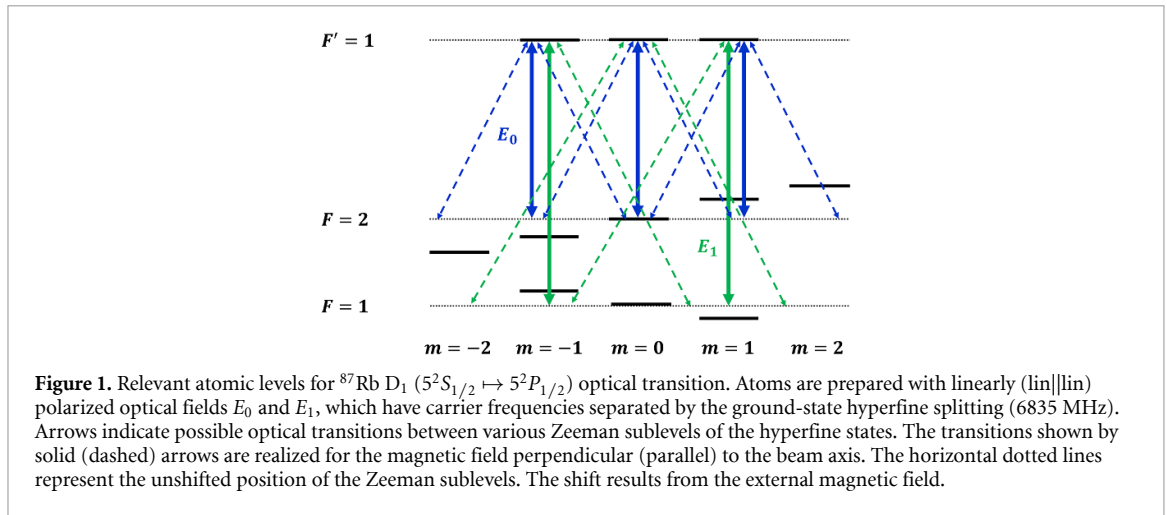
Abstract

Vector atomic magnetometers that incorporate electromagnetically induced transparency (EIT) allow for precision measurements of magnetic fields that are sensitive to the directionality of the observed field by virtue of fundamental physics. However, a practical methodology of accurately recovering the longitudinal angle of the local field through observations of EIT spectra has not been established. In this work, we address this problem of angle determination with an unsupervised machine learning algorithm utilizing nonlinear dimensionality reduction. The proposed algorithm was developed to interface with spectroscopic measurements from an EIT-based atomic rubidium magnetometer and uses kernel principal component analysis (KPCA) as an unsupervised feature extraction tool. The resulting KPCA features allow each EIT spectrum measurement to be represented by a single coordinate in a new reduced dimensional feature space, thereby streamlining the process of angle determination. A supervised support vector regression (SVR) machine was implemented to model the resulting relationship between the KPCA projections and field direction. If the magnetometer is configured so that the azimuthal angle of the field is defined with a polarization lock, the KPCA-SVR algorithm is capable of predicting the longitudinal angle of the local magnetic field within 1 degree of accuracy and the magnitude of the absolute field with a resolution of 70 nT. The combined scalar and angular sensitivity of this method make the KPCA-enabled EIT magnetometer competitive with conventional vector magnetometry methods. © 2023. All rights reserved.

1. Introduction

Emerging methods in machine learning and the continued maturation of embedded system processors has enabled the development of smart sensor technologies within the field of metrology [1–4]. Joint optimization of sensor design and statistical signal processing allows for the efficient extraction of complex relationships within sensor data, resulting in improved sensitivity and speed in sensor performance. Unsupervised machine learning techniques, such as those used for clustering and dimensionality reduction, have allowed for numerical modeling of physical processes without the need for initial assumptions of the underlying physics. In atomic, molecular, and optical (AMO) spectroscopy, machine learning techniques have been used for regression problems such as absorbance measurement [5, 6], signal restoration [7, 8], density estimation [9] and quantum state reconstruction [10]. Furthermore, applications have been found in classification problems for the identification of light sources [11, 12], near infrared spectroscopy [13], and electron microscopy [14].

Optical vector atomic magnetometry is a maturing AMO research area that stands to benefit from these data-driven modeling methods, due to the complexity of the physics involved. Atomic magnetometers are



typically scalar sensors that measure transitions between magneto-sensitive spin sublevels of the ground state of alkali atoms. These instruments boast both high accuracy and sensitivity, and have been found to be valuable in underwater anomaly detection as well as planetary science [15–18]. At the time of writing, the highest performing scalar atomic magnetometers utilize spin-exchange relaxation free (SERF) techniques [19, 20], and can achieve sensitivity on the order of $1 \text{ fT}/\sqrt{\text{Hz}}$. However, SERF magnetometers are limited to scalar measurements at low magnetic fields, and are therefore ill-suited to unshielded, outdoor applications. In addition to the SERF designs, an assortment of other atomic magnetometers have been developed and commercialized [21].

Vector atomic magnetometry is typically performed by taking measurements with properly optimized scalar magnetometers. For instance, introducing low frequency AC fields along a set of coordinate axes that have been defined within a laboratory frame results in vector detection [22]. This approach requires an intricate coil system and fundamentally relies on physically aligning the orthogonality of the laboratory coordinate system. Consequently, any misalignment in the field orientation introduces errors in the determination of the field direction as well as cross-axis sensitivity.

Both the limitations of range and alignment sensitivity can be addressed through the use of all-optical interrogation techniques that utilize electromagnetically induced transparency (EIT). These measurements eliminate the need for an external coil system by probing an atomic resonance structure that is fundamentally sensitive to changes in both the magnitude and direction of the magnetic field vector. The measurements performed with ^{87}Rb atoms, with the energy level structure shown in figure 1, allow for the observation of seven transmission peaks associated with the available two-photon resonances between Zeeman sublevels. Measurements of the separation between these transmission peaks allow for scalar measurements of the absolute magnetic field. Changes in the direction of the magnetic field result in a variation of the coupling strengths between different magnetic sublevels, meaning that the direction of the field can be determined by examining the variation of the relative amplitudes of the transmission peaks. Therefore, measurements of the position and relative amplitudes of the EIT resonance peaks affords a means to perform vector measurements of the local magnetic field. This method for optical vector atomic magnetometry was developed theoretically [23] and demonstrated as a proof of concept [24].

While the vector EIT magnetometer has relatively low maximum achievable sensitivity when compared against state-of-the-art scalar magnetometers, the EIT measurement approach overcomes many of the limitations associated with conventional vector magnetometers. The advantages of the EIT measurement strategy include acceptable accuracy of scalar measurements due to the measurement of the electron spin precession frequency, as well as direction determination based on a coordinate system entirely defined by the direction of the laser wave vector and the direction of the input polarization of light. These parameters are much simpler to align with an external coordinate system than the complex coil systems associated with alternative vector magnetometry methods. Furthermore, the EIT measurement approach allows for sensitivity over a large dynamic range (up to 1 T), overcoming the limited range capabilities of competing magnetometers. A diagram defining the vector coordinate angles is shown in figure 2. Given this cylindrical coordinate system, the magnetic field vector, $\vec{B}$, can be defined with the following expression:

$$\vec{B} = |B| \sin(\theta) \cos(\phi) \mathbf{i} + |B| \sin(\theta) \sin(\phi) \mathbf{j} + |B| \cos(\theta) \mathbf{k} \quad (1)$$

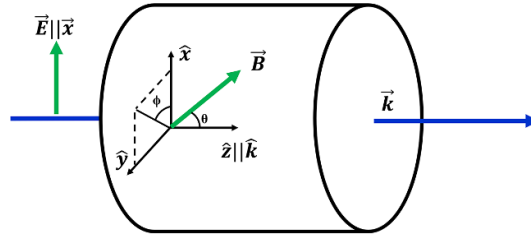


Figure 2. Magnetic field vector angle definitions used for the EIT wave vector $\vec{k}$. The laser coordinate system is defined by the laser beam direction and the polarization of the light. The angle ϕ denotes the azimuthal angle of the magnetic field and θ denotes the longitudinal angle.

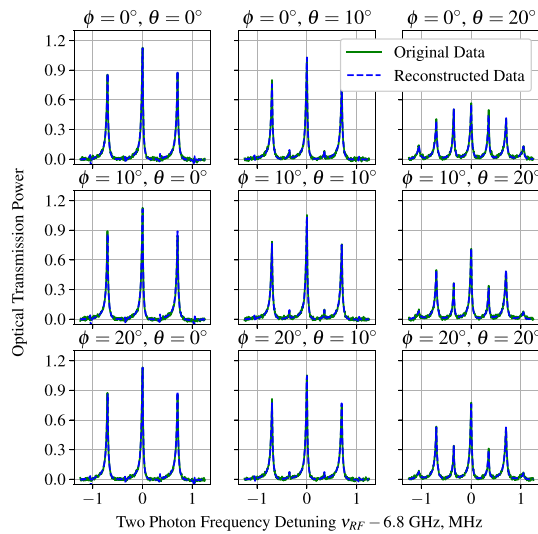


Figure 3. Sample EIT spectra recorded for different values of ϕ and θ . Measurements are performed by passing a laser beam through the Rb vapor cell and using a photodetector to record the transmission amplitude through the cell as a function of the laser detuning frequency. Measurements from experiment are shown with the green trace and examples of reconstructed EIT spectra using ten principal components are shown with the dashed blue trace. These spectra were reconstructed via the PCA technique discussed in the paper.

where $|B|$ is the magnitude, ϕ is the azimuthal angle, and θ is the longitudinal angle for the magnetic field. Examples of EIT spectra measurements recorded for different values of ϕ and θ are shown in figure 3.

While ϕ could be locked to the $\phi = 0^\circ$ position by actuating the polarization of the lasers (it can be done in an experiment using an eclectically actuated liquid crystal polarization rotator), a method for identifying the θ is less trivial. Initially demonstrated in [24], it was shown that the information needed to recover θ is contained within the amplitudes of the seven EIT resonances (figure 3). Theoretically, θ can be determined by evaluating the ratios of specific resonances in the EIT spectrum. However, it was unclear if the measurement can accurately recover θ in a realistic physical system where shapes of the resonances are distorted. Furthermore, while the information about the magnetic field is contained in each of the resonances, it was unknown which ratios of amplitudes could optimally determine θ in atomic cells containing buffer gases. To address this problem, we explored a variety of unsupervised machine learning techniques to capture the empirical behavior of this system. The adoption of unsupervised learning techniques was motivated by a desire to assess the system's behavior without relying on assumptions or detailed theoretical models. We found that dimensionality reduction techniques such as principal component analysis (PCA) were particularly effective at describing changes in the EIT spectra, and kernel PCA (KPCA), a nonlinear adaptation of PCA, was ultimately selected as the optimal procedure for determining θ in this type of EIT-based vapor cell vector magnetometer.

PCA methods can be used to reduce the dimensionality of the EIT spectra such that each particular spectrum can be represented as a coordinate in a new reduced dimensional subspace. The resulting coordinates can then be regressed to capture correlations between different configurations of the EIT spectrum and the direction of the local magnetic field. To this end, dimensionality reduction is not the primary objective in the application of KPCA to vector magnetometry. Rather, KPCA is used for feature

extraction in service of the ultimate goal of recovering labels of ϕ , θ , and $|B|$ from the observations of EIT spectra.

Applications of classical PCA techniques have been validated in quantum optics and used to study and classify phase transitions for muon spectroscopy [25], and the utility of nonlinear PCA methods in regression has also been thoroughly studied [26]. However, the potential of KPCA for regression problems in AMO spectroscopy has not been fully explored. While this paper specifically outlines the application of PCA to the EIT spectroscopy associated with this vector atomic magnetometer, the methods explained herein can be applied to any situation in which a user seeks to efficiently extract some complex relationship from a library of RF spectra.

Our experiments demonstrate that it is possible to create a vector magnetometer without involvement of the theoretical description of the physical process. The machine learning technique allows for measurement of the magnetic field with reasonable accuracy. This paper is organized as follows: in section II, we discuss the mathematics associated with applying variations of PCA to RF spectra; in section III, we apply these PCA methods to spectra collected with a prototype of the vector EIT magnetometer; in section IV, we demonstrate how the principal component projections can be utilized to recover individual scalar components of the magnetic field vector; in section V, we compare the PCA approach against a more conventional spectroscopy method and discuss how the PCA approach can be extended to recover additional components of the magnetic field vector; section VI concludes the paper.

2. Methodology

The central idea of the dimensionality reduction technique is to identify a limited number of eigenfunctions describing the possible variety of the EIT spectra, and then to use these eigenfunctions as a basis for a decomposition of the realizations of the spectra. As such, the decomposition coefficients will characterize modifications of the shape of the spectra. This approach is similar to the usage of Fourier decomposition in the harmonic spectral analysis, with the difference that PCA explicitly identifies the components experiencing maximum change with the variation of external parameters. In what follows we describe the approach in more detail and consider a few of its variations.

2.1. Principal component analysis

Classical PCA is the most widely used method of dimensionality reduction, and it is particularly effective for situations in which data is linearly separable [27]. Many real-world data sets are not linearly separable, and so nonlinear extensions of PCA have long been a source of interest within the study of dimensionality reduction [28, 29]. While this work explores the application of the nonlinear KPCA for magnetometry, the details concerning the implementation of KPCA are best understood first through the lens of classical PCA. A brief outline of PCA is provided below prior to the outline of the KPCA algorithm implemented for the magnetometer.

Let X be an $n \times p$ input data matrix containing n rows of independent observations and p columns of features. For the case of spectroscopic data, let us consider n as the number of spectrum measurement repetitions and p as the transmission values recorded at particular frequencies of the spectrum (number of points per spectrum). Thus X can be treated as a set of n , p -dimensional vectors in $\mathbb{R}^p$. We wish to represent the high dimensional structure of these spectra in a more concise manner, so that X transforms to a set of n , m -dimensional vectors in $\mathbb{R}^m$ (and $m < p$). Let us characterize possible variations of these spectra in terms of deviation from some average shape, the motivation for this step will be made clear later on.

Assuming the spectra have been aligned in frequency, we can define the mean spectrum by [30]

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i. \quad (2)$$

Each observation of the spectrum, X_i , can then be described by the deviation from the mean spectrum $\bar{X}$, giving rise to the centered data matrix

$$X_{0i} = X_i - \bar{X}, \quad i = 1 \dots n \quad (3)$$

where X_{0i} denotes the i^{th} column of the centered data matrix X_0 . The goal of PCA is to create a linear map that transforms X_0 into a new feature space $X_0 \mapsto WX_0$, where W is the compression matrix of size $p \times m$. Note here that the matrix multiplication is read from right to left. Projections of X_0 in the new feature space can be calculated with:

$$F = WX_0 \quad (4)$$

where the feature matrix F , with size $n \times m$, is a representation of each observation in X_0 with the reduced dimensionality of m features. The loss associated with this compression process can be determined with some recovery matrix, R , with size $m \times p$, such that an approximation of the original data matrix, $\tilde{X}_0$, can be calculated with $\tilde{X}_0 = RWX_0$. As such, the central problem in PCA is to determine W and R in a manner that minimizes the total squared Euclidean distance between the original and reconstructed data matrices. This is stated mathematically in equation (5) [31]:

$$\operatorname{argmin}_{W,R} \sum_{i=1}^n \|X_{0i} - RWX_{0i}\|^2. \quad (5)$$

2.1.1. PCA via covariance matrix

Previous studies have demonstrated that the objective of minimizing the reconstruction error is equivalent to maximizing the variance of the projected data [32], and both can be accomplished through eigendecomposition of the sample covariance matrix of X , given by:

$$C = \frac{1}{n} X_0^T X_0. \quad (6)$$

Equation (6) results in a $p \times p$ matrix where the diagonal entries describe the variance between features of X and all other entries describe the covariance between features of X . Therefore, the total variance of X can be calculated by taking the trace of C . The corresponding eigenvalue expression for matrix C is:

$$Cu_k = \lambda_k u_k, \quad k = 1 \dots p \quad (7)$$

where λ_k is an eigenvalue of C corresponding to the eigenvector u_k , and $u_k^T u_k = 1$. The eigenvectors u_k are therefore an orthonormal set of representative EIT spectra into which any measured spectrum can be decomposed. Vectors from the centered data set may be projected onto the orthonormal basis defined by these eigenvectors with

$$X_{\text{proj}} = AX_0 \quad (8)$$

where $A = [u_1 \dots u_p]^T$ is the matrix of eigenvectors of C . The covariance of X_{proj} then is given by:

$$C_{\text{proj}} = \frac{1}{n} X_{\text{proj}}^T X_{\text{proj}} = \frac{1}{n-1} A^T C A \quad (9)$$

Because the vectors of A are orthonormal, this mapping is a similarity transformation and therefore X_{proj} preserves the total variance of X .

It is now possible to reconstruct individual vectors in the centered data set as a linear combination of the projections and eigenvectors:

$$X_{0i} = \sum_{k=1}^p u_k (u_k^T X_{0i}). \quad (10)$$

The eigenvectors of C are called the principal axes (also known as principal directions) of our data set, and the projections of specific observations onto these axes are called the principal components. Note however that this nomenclature can be inconsistent in the literature, and that some sources use ‘principal components’ to refer to the eigenvectors themselves rather than projections onto these eigenvectors. For the purposes of this paper, the term ‘principal components’ refers exclusively to the projections of X into the reduced dimensional subspace. Recall that we calculated the principal components and axes of X using the centered data matrix X_0 . The analysis can also be performed with the original, non-centered data matrix X . However, under this approach, the first principal axis of the data will simply be the average shape of the input data. Ordinarily, the average shape is not particularly insightful in evaluating patterns within the data set, and so it is often discarded by first centering the input data.

The eigenvalues of C correspond to the amount of explained variance described with each eigenvector. If we select only m eigenvectors corresponding to the largest m eigenvalues, we can approximate X_{0i} with equation (11):

$$\tilde{X}_{0i} = \sum_{k=1}^m u_k (u_k^T X_{0i}). \quad (11)$$

This expression is equivalent to $\tilde{X}_0 = D^T D X_0$, where $D = [u_1 \dots u_m]^T$ has been constructed as the matrix of eigenvectors corresponding to the m largest eigenvalues. Therefore, equation (5) is satisfied if we set W to D and R to D^T . Any future observations of the system, X_{new} may also be projected onto the principal axes with:

$$X_{\text{proj}_{\text{new}}} = D(X_{\text{new}} - \bar{X}). \quad (12)$$

2.1.2. PCA via Gramian matrix

The above solution assumes that $n > p$ for the calculation of D . In the event that $p > n$, it is computationally more efficient to compute the Gramian matrix of X in place of the covariance matrix. While the covariance matrix (size $p \times p$) describes the Euclidean distance between the different features of X , the Gramian matrix (size $n \times n$) describes the Euclidean distance between the different observations of X . The Gramian matrix takes the form:

$$G = \frac{1}{n} X_0 X_0^T. \quad (13)$$

Once again, PCA is performed by solving the eigenvalue problem for G and retaining only the m eigenvectors corresponding to the m largest eigenvalues. However, note that under this formulation, the eigenvectors of G do not correspond to the principal axes of X . Instead, eigenvectors of G correspond to the principal component projections of X . The cause of this difference is clear when PCA is considered through the framework of single value decomposition (SVD) [33]. The SVD theorem states that a real matrix can be represented by [34]:

$$X_0 = U \Sigma V^T \quad (14)$$

where U is an $n \times n$ orthonormal matrix where the columns are the eigenvectors of $X_0 X_0^T$ and V is a $p \times p$ orthonormal matrix where the columns are the eigenvectors of $X_0^T X_0$. Σ is an $n \times p$ diagonal matrix of the singular values of X_0 , which are given by the square roots of the eigenvalues of either $X_0 X_0^T$ or $X_0^T X_0$. Note that the eigenvalues of $X_0 X_0^T$ and $X_0^T X_0$ are equivalent. By plugging equation (14) into equation (6), we get:

$$C = V \frac{\Sigma^2}{n} V^T. \quad (15)$$

This expression is simply the eigendecomposition of C , and it affirms that the SVD matrix V coincides with the matrix A . By extension, the $U \Sigma$ term in equation (14) coincides with the principal components of X . Similarly, by plugging equation (14) into equation (13), we have:

$$G = U \frac{\Sigma^2}{n} U^T. \quad (16)$$

So we see that the eigenvectors of G coincide with the matrix U in equation (14). Because Σ is only a diagonal matrix of the singular values of X , we can see that the matrix U also coincides with principal components of X . However, these projections will not be scaled by the singular values of X , as they would be with the covariance method. In order to calculate the principal components of X using the Gramian method such that they coincide with the principal components of the covariance method, the eigenvectors of G must be scaled by the square roots of the eigenvalues (the singular values of X):

$$X_{\text{proj}_G} = U \Sigma. \quad (17)$$

The computational complexity associated with solving PCA via the covariance matrix is $O(np^2)$ for the construction of C plus $O(p^3)$ for calculating the eigenvalues of C . The computational complexity of solving PCA with the Gramian matrix is $O(n^2p)$ for the construction of G plus $O(n^3)$ for calculating the eigenvalues of G . Therefore, the decision of performing PCA through the covariance or Gramian matrix should be determined by the relative size of the observations and features of the data set.

2.2. Kernel principal component analysis

Note that the Gramian matrix method allows us to calculate U without the need to explicitly calculate V . This means that for a given data set, only the inner products between vectors need to be calculated in order to perform PCA. KPCA exploits this attribute by mapping the input data into a higher dimensional space and computing the Gramian matrix implicitly. While this step may seem counter intuitive to the goal of dimensionality reduction, it is worth noting that data that is not linearly separable can often be made linearly separable in a higher dimensional space. KPCA involves projecting the input data into this high-dimensional space and then applying classical PCA to the projected data via the Gramian matrix method. This nonlinear

dimensionality reduction allows us to obtain a reduced dimensional representation of the data that is linearly separable, even when the input data is not.

The process is accomplished by introducing a nonlinear transformation $\phi(x)$ that maps the p -dimensional vectors of X into an ℓ -dimensional feature space where $\ell > p$ (let us assume for the moment that X_0 remains centered under this mapping, see equation (3)). If ℓ is not too large, PCA could still be performed explicitly in the higher dimensional space with the covariance matrix:

$$C = \frac{1}{n} \phi(X_0)^T \phi(X_0). \quad (18)$$

This calculation will now have the computational cost of $O(n\ell^2)$ for the construction of C plus $O(\ell^3)$ for calculating the eigenvalues of C . Fortunately, kernel substitution can be utilized to calculate the Gramian matrix of the ℓ -dimensional data using only the original p -dimensional data [28]. Kernel substitution consists of defining an arbitrary kernel function (such as the nonlinear transformation described above) that is used to calculate an $n \times n$ kernel matrix for X . The entries of the kernel matrix can be expressed in terms of pairs of the p -dimensional input data vectors. The $n \times n$ kernel matrix, expressed in terms of the kernel function $\phi(x)$ is given by:

$$k(X_i, X_j) = \phi(X_i)^T \phi(X_j). \quad (19)$$

While we initially assumed that X_0 remained centered under $\phi(x)$, this cannot be assumed to be true in general. In order to avoid the cost of computing the ℓ -dimensional average shape vector, we can instead determine the centered kernel matrix with the following expression [32]:

$$\tilde{K} = K - \mathbf{1}_n K - K \mathbf{1}_n + \mathbf{1}_n K \mathbf{1}_n \quad (20)$$

where $\mathbf{1}_n$ is an $n \times n$ matrix where each element is $1/n$. $\tilde{K}$ is then equivalent to the Gramian matrix of the ℓ -dimensional data, and the eigenvectors of $\tilde{K}$ denote the nonlinear principal components of X scaled by the singular values of X .

An assortment of kernel functions can be selected for KPCA. Popular kernels include polynomial, sigmoid, and radial basis function (RBF) kernels. For the EIT spectra evaluated within this paper, a Gaussian RBF kernel was chosen, with the associated kernel matrix shown in equation (21). This kernel function was selected because of the flexibility in nonlinear mapping provided by the hyperparameter γ , which is included in equation (21) as a free parameter to be optimized for a specific application

$$k(X_{0_i}, X_{0_j}) = \exp(-\gamma \|X_{0_i} - X_{0_j}\|_2^2). \quad (21)$$

The value of γ can be understood qualitatively as the extent to which the resulting dimensionality reduction is nonlinear, and the selection of γ is typically determined through unsupervised clustering methods or supervised cross validation methods. Guidelines for selecting an appropriate value of γ with both approaches are discussed in sections 3 and 4 of this paper. For the EIT magnetometer, varying γ was found to be an effective method of adjusting the clustering of the EIT spectra projections. The nonlinearity introduced with KPCA allowed the magnetometer model to resolve subtle changes in the EIT spectra, which were found to be associated with subtle changes in the magnetic field. Outside of magnetometry, some excellent examples of KPCA applications can be found in [35–37].

2.3. Regression

While this work is focused on exploring the potential of PCA methods for feature extraction of EIT spectra, creating an operational magnetometer that utilizes PCA for vector measurements requires the implementation of a regression model that correlates these features back to physical quantities associated with the observed magnetic field. The problem becomes a supervised learning task, requiring data sets in which the true values of $\vec{B}$ are known for each EIT spectrum. The appropriate regression algorithm to select depends heavily on the system behavior observed with PCA, but for the EIT magnetometer considered in the study, support vector regression (SVR) machines were found to efficiently fit the observed trend of principal components. The SVR approach was selected because it was flexible and well-suited to the inherent sparsity of the EIT data set (each spectrum is associated with discrete values of the magnetic field, but ϕ and θ are truly continuous values).

The goal of the SVR machine is to identify some function, f , that can approximate the data such that the difference between the predicted values and true values lie within some preset threshold, ϵ . The function is calculated so that the prediction error is within the ϵ limit for as many data points as possible while keeping the function as flat as possible. Similar to KPCA, the SVR method accomplishes this goal by employing a

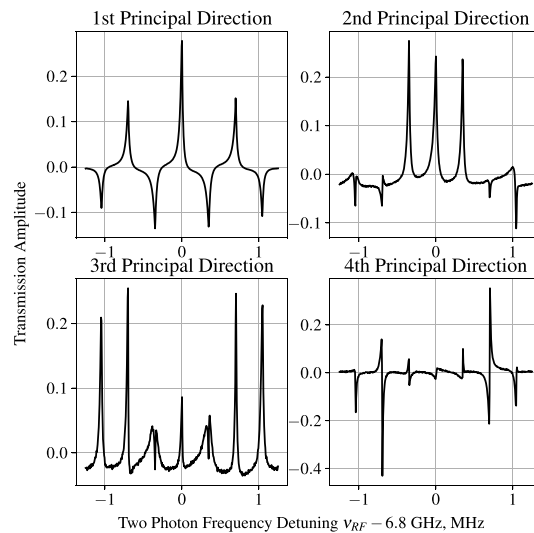


Figure 4. First four principal axes calculated for the EIT data set. Projections of measurement data onto this basis (the principal components) describe the extent to which these shapes are present in the original measurements.

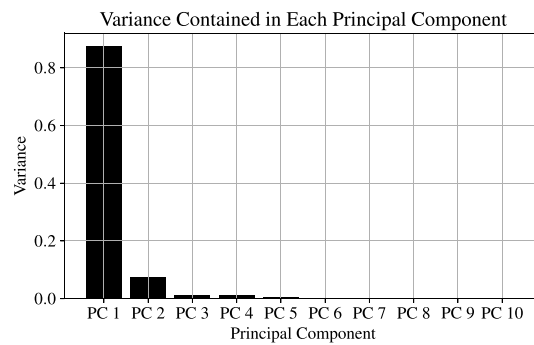


Figure 5. Eigenvalues calculated from covariance matrix of the EIT spectra. Over 98% of the total variance of the data set can be explained using only the first ten principal components.

kernel trick to project the input data into a higher dimensional feature space. In the new feature space, the projected data is fit to a maximum-margin hyperplane in which linear regression can be performed. A detailed explanation on the implementation of the SVR machine is beyond the scope of this paper, but an excellent overview can be found with [39–41]. Additionally, examples of SVR incorporated with KPCA can be found in [42, 43]. For the EIT magnetometer, the SVR machine was trained to receive principal components calculated from the EIT spectra as an input, and would then make a continuous-valued prediction for the scalar quantities of θ , ϕ , or B associated with the measured spectrum.

3. Application

To explore the relationship between EIT spectra and the direction of the magnetic field, the EIT Rb magnetometer was tested in a magnetically shielded environment. Manipulating the current in the Helmholtz coils, placed inside the shields, adjusted the direction of the applied magnetic field. For the initial testing, the applied field was maintained with an approximate magnitude of 500 mG (Earth's field). The EIT spectrum was recorded with 10° increments in the azimuthal angle, ϕ , and 5° increments in the longitudinal angle, θ , with multiple spectra recorded for each direction. Each spectrum consisted of 500 transmission amplitudes measured as a function of the laser detuning frequency, and approximately 1500 spectra were collected for the initial evaluation. The EIT measurements were prepared into the format listed in section II, creating a 1500×500 data matrix on which PCA could be performed. After calculating the mean spectrum and centering the data, the 500×500 covariance matrix was assembled to perform classical PCA on the EIT data set. The first four principal axes of the EIT data set, corresponding to the eigenvectors of the covariance matrix with the four largest eigenvalues, are shown in figure 4. The relative explained variance contained in the first 10 eigenvectors is shown in figure 5.

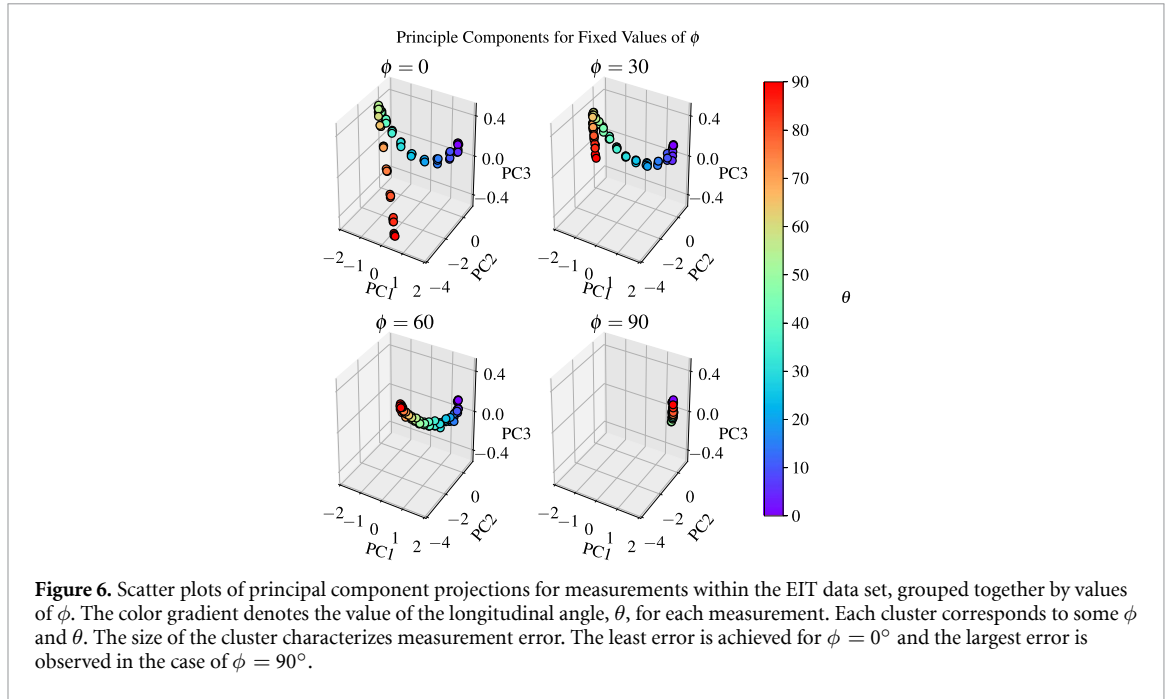


Figure 6. Scatter plots of principal component projections for measurements within the EIT data set, grouped together by values of ϕ . The color gradient denotes the value of the longitudinal angle, θ , for each measurement. Each cluster corresponds to some ϕ and θ . The size of the cluster characterizes measurement error. The least error is achieved for $\phi = 0^\circ$ and the largest error is observed in the case of $\phi = 90^\circ$.

An approximation of the original data set can be calculated using a linear combination of the principal components and axes. Examples of the reconstructed EIT spectra are shown in figure 3. While this reconstruction process can be useful for denoising a signal, the true utility of PCA is found in its ability to represent each measurement as a coordinate in the reduced dimensional space. Figure 6 shows the first three projections of each measurement for the EIT data set in a three dimensional scatter plot for different values of the azimuthal angle ϕ .

Classical PCA provides a means to quantify the sensitivity of this measurement system with respect to the determination of θ as a function of ϕ . As figure 6 shows, the principal components of the EIT spectra all fall along a smooth curve in the reduced dimensional space. Note that in the top left plot of figure 6, we observe clustering at specific values of θ . This trend continues for each value of ϕ ; however, at higher values of ϕ , the spacing between clusters decreases. Fortunately, ϕ can be independently controlled with the polarization of the laser. Therefore, for optimal performance, the laser polarization must be locked to $\phi = 0^\circ$ in order to utilize the full range of possible amplitude variations in the EIT spectra.

For KPCA, a similar trend can be observed, but the RBF kernel function introduces the tunable parameter γ that can adjust the contour of the projections curve. The influence of γ on the projections curve is demonstrated with figure 7. For low values of γ , the RBF kernel converges to a linear kernel, and so the classical and kernel principal component projections coincide. As γ increases, both the broader projections curve and the clustering behavior change. Thus, KPCA provides some flexibility in arranging how different measurements may be grouped together. This flexibility is particularly valuable in resolving small values of θ , in which the classical PCA projections are not well separated. Note that the trends shown here extend beyond the first three principal components, and the two and three dimensional scatter plots are only included to give the reader a qualitative understanding of the sort of system behavior that PCA methods can uncover.

For many applications, γ is optimized with cross-validation approaches, in which a regression model is created to evaluate the predictive power of the KPCA projections as a function of γ . Naturally, this method necessitates a supervised learning algorithm that requires assumptions about how different observations of the data may be correlated. Such assumptions potentially defeat the purpose of the unsupervised analysis, because the achievable performance of the model will be dependent on these underlying assumptions. Fortunately, there are clustering evaluation metrics that can be incorporated to optimize γ without compromising the unsupervised nature of the analysis. For the EIT dimensionality reduction, the silhouette score method was used to optimize γ as it pertains to resolving values of θ . The silhouette score was calculated by fitting the KPCA projections into k clusters with a k -means clustering algorithm and evaluating the size of clusters relative to the separation between neighboring clusters [38]. For a given set of KPCA projections, the silhouette score, C_S , was calculated using:

$$C_S = \frac{1}{n} \sum_i^n \frac{b_i - a_i}{\max(a_i, b_i)} \quad (22)$$

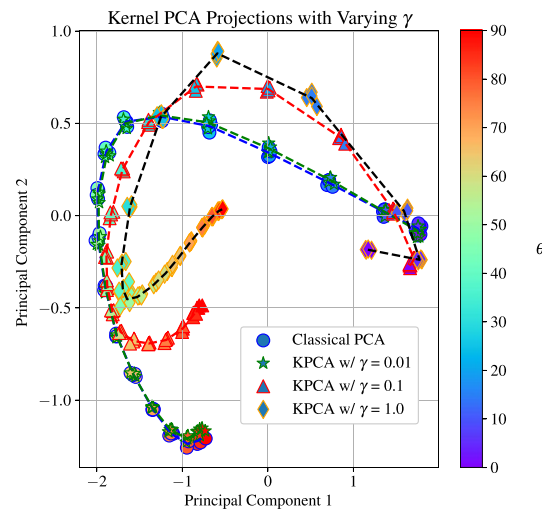


Figure 7. Scatter plot showing the first two principal component projections for each measurements where $\phi = 0$ for different values of γ for the KPCA RBF kernel. In this plot, all of the projections have been scaled by the singular values of the EIT data set. As γ increases, the clusters corresponding to low values of θ become well resolved. For classical PCA, these same clusters are not well separated, and it is difficult to resolve small changes in the spectrum at low values of θ . This relationship illustrates the utility of the KPCA technique.

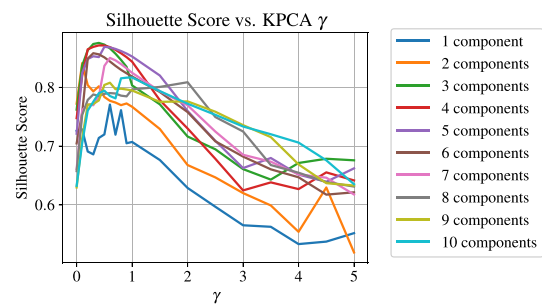


Figure 8. Influence of KPCA γ on the silhouette score, C_S for the EIT data set.

where a_i is a given coordinate in the reduced dimensional feature space and b_i is the closest coordinate from the neighboring cluster. For the EIT measurement where $\phi = 0$, there were 19 distinct values of θ measured. Thus, the k -means clustering algorithm was used to fit the KPCA projections into 19 clusters, and C_S was evaluated as a function of both γ and the number of kernel principal projections used in the dimensionality reduction. The results of this analysis are shown in figure 8.

KPCA achieves improved clustering over classical PCA for the higher order principal components at optimal values of γ . This can be seen in figure 8, as the low values of γ correspond to the clustering performance of classical PCA. This analysis suggests that the optimal performance of the model is found using 3–4 principle components and a $\gamma \approx 0.3$. The improved performance at higher values of γ demonstrates that some manifestations of the EIT spectra are not linearly separable, and so the kernel function is needed to resolve values of θ corresponding to these spectra.

4. Results

The above analysis highlights the value of unsupervised machine learning for gaining insight into the inner workings of the magnetometer system, particularly with the identification that the greatest variation in the EIT spectra occurred when the polarization ϕ was locked to 0° . In response to this finding, two additional data sets were collected to assess the performance of the EIT magnetometer using the PCA/KPCA and the supervised SVR algorithm. The first data set consisted of 2037 measurements of EIT spectra associated with a magnetic field where $\phi = 0^\circ$ and the absolute field B was maintained at 500 mG. This data set was used to characterize the ultimate sensitivity of the PCA method for directionality measurements. The second data set consisted of 1415 measurements of EIT spectra where $\phi = 0^\circ$ and the absolute field was varying from 50 to

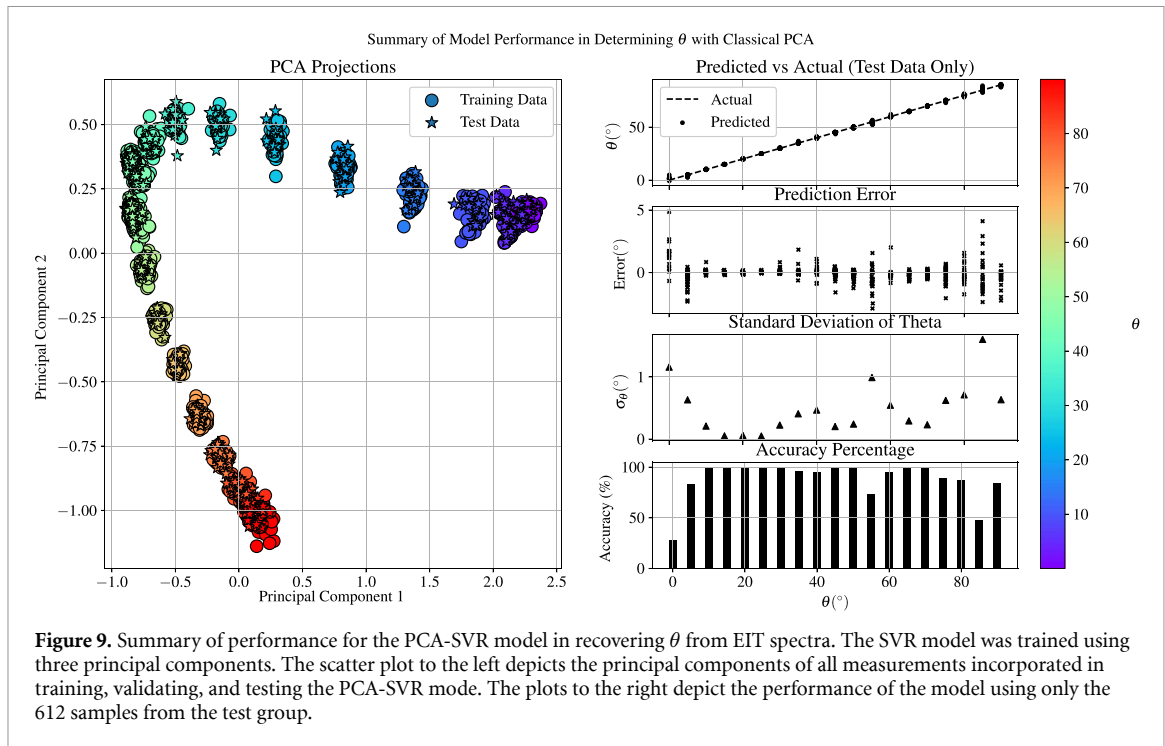


Figure 9. Summary of performance for the PCA-SVR model in recovering θ from EIT spectra. The SVR model was trained using three principal components. The scatter plot to the left depicts the principal components of all measurements incorporated in training, validating, and testing the PCA-SVR mode. The plots to the right depict the performance of the model using only the 612 samples from the test group.

56 μT , and this data set was used to characterize sensitivity of the PCA method for scalar measurements. Results from both modes of measurement are discussed in the following section.

4.1. Angular measurements

The 2037 measurements of EIT spectra were separated into a training group of 1425 samples and a test group of 612 samples. The training samples were used to calculate the principal components of the data set, and the SVR model was trained to recover specific values of θ using five-fold cross validation on the principal components of the 1425 training samples. Hyperparameters of the SVR model were optimized using the coefficient of determination, r^2 , given by:

$$r_{\theta}^2 = 1 - \frac{\sum_i^n (\theta_i - \tilde{\theta}_i)^2}{\sum_i^n (\theta_i - \bar{\theta})^2} \quad (23)$$

where θ_i is the observed (true) value of θ associated with the i th spectrum (controlled in experiment by applying fixed currents to a set of three-dimensional Helmholtz coils), $\tilde{\theta}_i$ is the predicted value of θ from the PCA-SVR model, and $\bar{\theta}$ is the average value of θ . The SVR machine also incorporated an RBF kernel for fitting the principal component curve, and it was implemented and optimized using various libraries from Scikit-Learn in Python [44].

The above training method was first used with classical PCA, and a summary of the PCA-SVR model performance is provided in figure 9. Note that here we define accuracy as the percentage of measurements in which spectra from the test data set were used to correctly predict the value of θ within one degree of the true value from experiment. Clearly the classical PCA method works reasonably well in predicting values of θ . However, this summary shows that the model struggles to resolve both lower and higher values of θ . There are no tuning parameters to be optimized with classical PCA, and so this is the best performance that this method can achieve.

The training method was then repeated using the KPCA approach, with the additional step of optimizing γ for the KPCA RBF kernel alongside the hyperparameters of the SVR machine. The accuracy of the KPCA-SVR model evaluated as a function of KPCA γ is shown in figure 10. This result agrees well with the unsupervised analysis from figure 8, and once again we find that the KPCA model is optimized with four principal components and a $\gamma \approx 0.3$. Note that this plot also shows the accuracy of classical PCA as a function of the number of principal components for low values of γ .

Once optimized, the KPCA-SVR model was able to predict the value of θ for a given EIT measurement with a standard deviation better than 1 degree and an R^2 score of 0.9996 on the test data set. A summary of the performance of the KPCA-SVR model is shown in figure 11. Here we see that the nonlinear

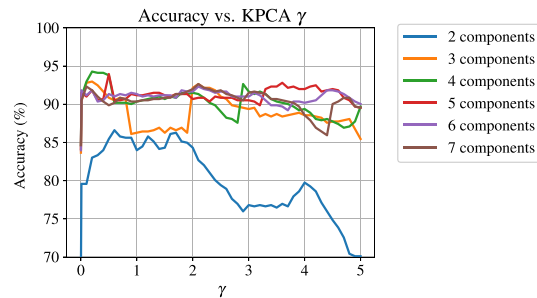


Figure 10. Influence of KPCA γ on the accuracy of the SVR model, with accuracy expressed in terms of the percentage of predictions within one degree of the true value of θ .

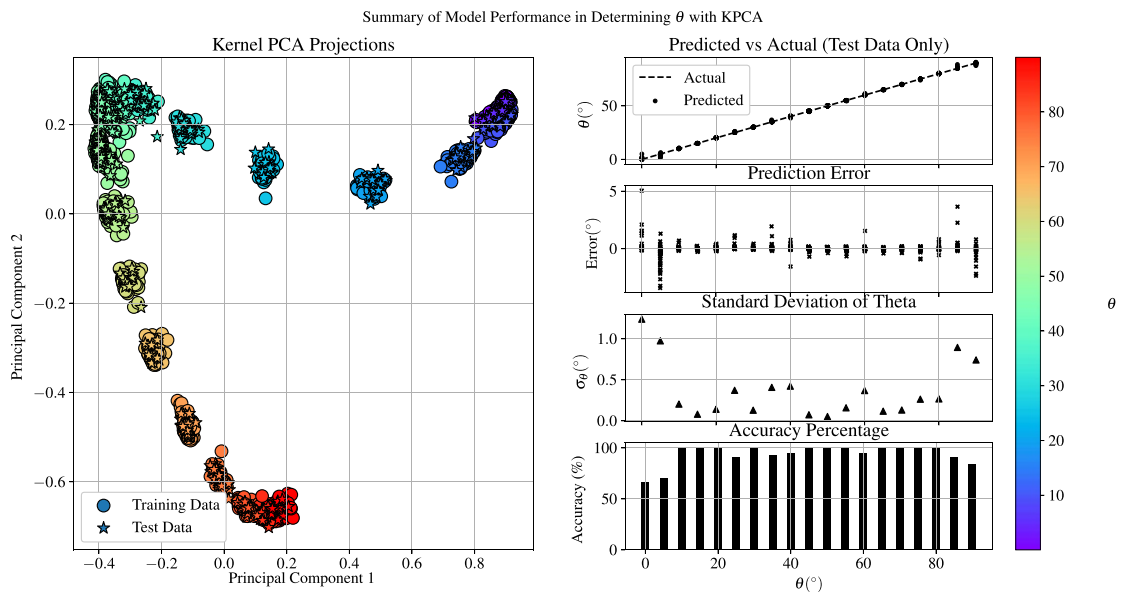


Figure 11. Summary of performance for the KPCA-SVR model in recovering θ from EIT spectra. The SVR model was trained using four kernel principal components calculated with an RBF kernel ($\gamma = 0.35$).

dimensionality reduction introduced with KPCA is able to resolve the the the extremities of θ , affirming that some manifestations of the EIT spectra are not linearly separable.

4.2. Scalar measurements

Similar to the directionality measurement analysis, the 2537 measurements containing variations in θ and $|B|$ were divided into 990 training samples and 425 test samples. These training samples were used fit the PCA model and train a new SVR model to predict the magnitude of $\vec{B}$ from the EIT spectra. Variations in the absolute field do not influence the line shape or the relative amplitudes of the EIT spectra, but they do change the separation between the transmission peaks. The principal axes calculated with this data set reflect this behavior, and we can see in figure 12 that the line shape of the first principal axis has been broadened when compared against that of the first principal axis shown in figure 4. Furthermore, the second principal axis closely resembles the first derivative of the EIT spectra, which is precisely the feature that would be used to make this scalar measurement in a classical laboratory environment.

The trained PCA-SVR model was used to predict $|B|$ for the 425 test samples. This proved to be a very effective method of determining $|B|$, producing an average standard deviation of 70 nT across all values of $|B|$. The performance of this method is summarized in figure 13. This analysis was performed using only the classical PCA approach, assuming that a linear model would be sufficient in describing variations in the position of the EIT resonances as a function of $|B|$. While 70 nT will certainly not compete with conventional atomic magnetometer measurement methods, it is worth noting that the information needed for this measurement is included for free in the same principal components that can be used for the angular measurement of $\vec{B}$. Furthermore, this magnetometer was not configured to optimize long-term scalar measurement stability, primarily due to environmental fluctuations in the laboratory environment. The

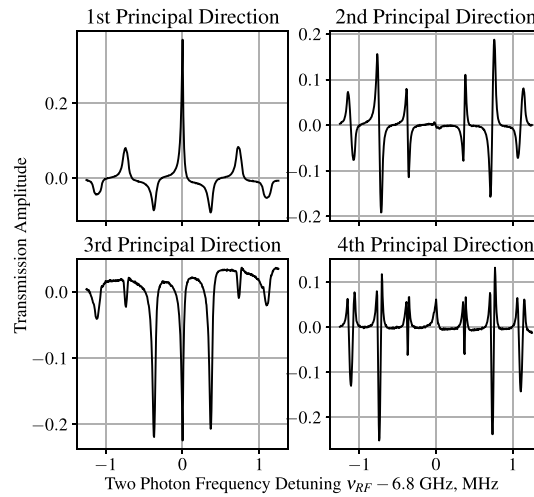


Figure 12. First four principal axes calculated for the EIT data set with variations in $|B|$ from 50 to 56 μT . Note the line broadening of the first principal axis, resulting from variations of the position of the EIT resonances caused by different magnitudes of the absolute field.

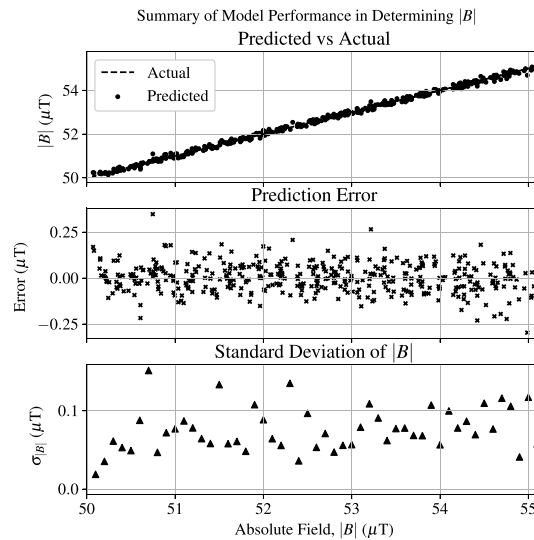


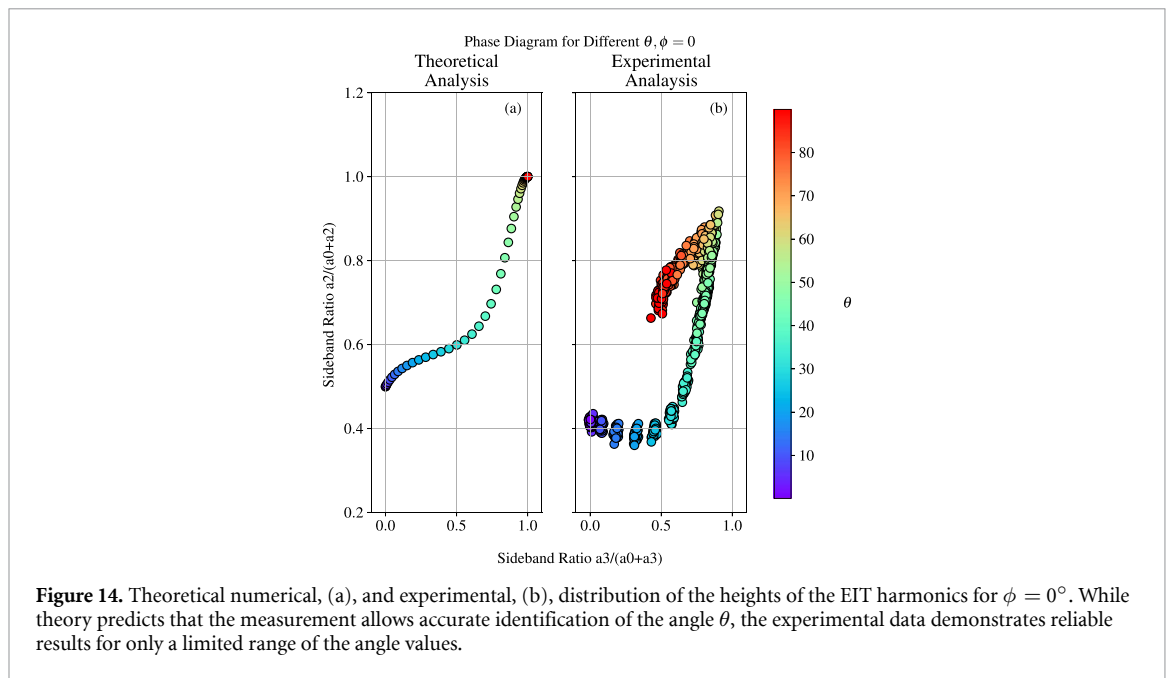
Figure 13. Performance summary of the PCA-SVR model in recovering $|B|$ from EIT spectra.

observed 70 nT stability may simply reflect the stability of the magnetometer rather than the stability of the PCA post-processing. Regardless, this finding asserts the value in using PCA for vector measurements of $\vec{B}$, given its ability to simultaneously perform magnitude and directionality measurements with EIT spectra.

It is known that magnetometers of this type are capable of achieving pT sensitivity [45]. With phase sensitive detection of the frequency separation of the EIT resonances, we achieved sensitivity better than $10 \text{ pT}/\sqrt{\text{Hz}}$ in the 1 Hz–1 kHz frequency range. Preliminary theoretical analysis shows that this result can be improved by an order of magnitude via optimization of the interrogation scheme. The results of this study will be published elsewhere.

5. Discussion

The proposed machine learning approach for vector magnetometry does not provide very high accuracy and sensitivity for magnetic field values, but rather is intended for practical use over a large dynamic range of measurements. It is feasible that higher sensitivity could be achieved by incorporating a multistage algorithm in which a coarse prediction of θ is made and then the spectrum is passed on to an assortment of regression models specially optimized for specific ranges of θ . The development of such an algorithm may be the subject of follow-on work in the future development of this technology.



Furthermore, the reported results involve data sets collected over comparably long periods of time with a tabletop breadboard prototype that has not been optimized for low-noise performance. As such, the system is contaminated with many classical noise sources and intrinsic drifts that contribute to a sensitivity reduction of the proposed technique. While multiple data sets were collected over the course of the study, modifications were made to the test setup between the collection of data sets (replacing optical components, changing vapor cells, etc). Because of these changes to the test setup, the principal axes of the EIT spectra are slightly different between the different data sets. The KPCA-SVR model trained with one data set can be used to make accurate predictions for magnetic field values in another data set, but because of the adjustments to the system, there is a small degradation in performance ($\pm 2.5^\circ$). We believe that this degradation does not reflect a limitation in the proposed methodology, but is simply an artifact from the testbed adjustments. When implementing this protocol in an operational magnetometer, the principal axes of the system should be calculated when other experimental parameters have been optimized and fixed (laser power, RF power, cell temperature, etc).

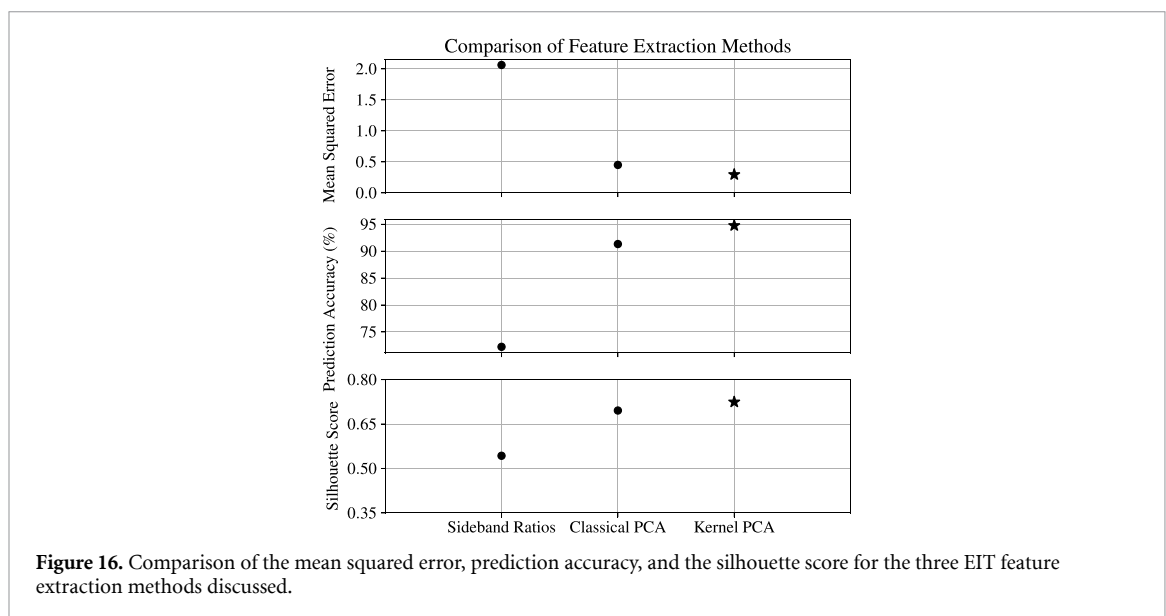
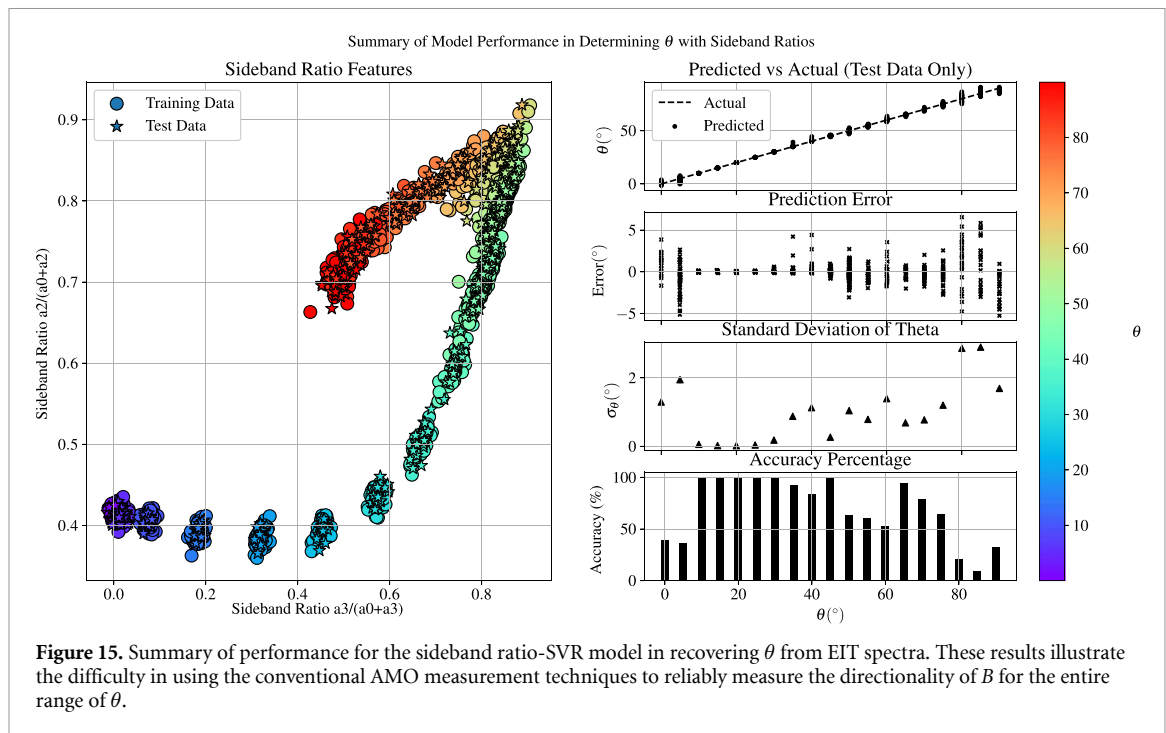
Despite these limitations, the results indicate that machine learning methods such as the proposed KPCA-SVR technique provide a viable alternative to conventional AMO measurement techniques. In what follows, we consider how this technique compares to the conventional methods in spectroscopy, and how the technique might be extended for simultaneous detection of θ and ϕ .

5.1. Comparison with conventional methods

To measure the direction of the magnetic field in a more conventional manner, a sensor must identify the relative height of the observed EIT resonances and compare the measurement results with the theoretical predictions [23]. The observation accuracy is hindered by a few technical limitations resulting from a misalignment of the optical beams, inhomogeneity of the magnetic field, deformation of the resonance curves as the result of a buffer gas presence, residual elliptical polarization of the beams, and other reasons. The technical factors leading to asymmetry of the optical spectrum make accurate theoretical evaluation of the field direction difficult.

This statement is illustrated by figure 14, in which the dependence of the spectral peak heights is found both theoretically and experimentally for the case of fixed azimuthal angle $\phi = 0$. For this analysis, the height of each EIT resonance is measured relative to the background of the measurement, and side band ratios are calculated using the a_0 , a_2 , and a_3 resonances. In figure 14, the vertical feature is calculated with $a_2/(a_0 + a_2)$ and the horizontal feature is calculated with $a_3/(a_0 + a_3)$.

While these curves do not coincide, they do produce features of the same magnitude, and it is clear that these features possess value in recovering θ . In order to evaluate the predictive power of this method, the side band ratios were also used to train an SVR model using the same procedure and data set outline in section 4(A). The summary for the performance of this method is shown in figure 15. While this model is quite competitive for the $0^\circ < \theta < 30^\circ$ range (as expected with the clustering observed in figure 14(b)), it is clear that this method fails to accurately predict θ within one degree for most configurations of $\vec{B}$.



Machine learning methods, like the PCA techniques outlined in this paper, reduce the effect of these systematic errors because they do not rely on a theoretical model of the setup. Any technical imperfections of the system are inherently taken into account when developing data-driven models. By comparing figure 15 against figure 9 and 11, it is clear that the PCA methods outperform the theoretical approach and make predictions for the entire range of directions ($0^\circ < \theta < 90^\circ$) feasible. A performance summary for all three EIT feature extraction methods outlined in this paper is provided in figure 16. Here we see that for all of the figures of merit considered, the PCA methods surpass the theoretical method, and KPCA offers a slight improvement over the classical PCA method. It is worth noting that the time complexity associated with identifying the extrema of each spectrum is approximately $O(pn)$, which is certainly more efficient than any of the PCA techniques. However, the sensitivity to variations in the background and aforementioned technical limitations make this approach less practical for use in uncontrolled environments.

5.2. Simultaneous measurement of longitudinal and azimuthal angles

This work has focused on the utility of KPCA in determining only the longitudinal angle, θ , from EIT spectra collected with the magnetometer. Though slightly more challenging, it is possible to extend the methodology

to determine the azimuthal angle, ϕ , as well. The measurement is not straight forward and involves observing multiple manifestations of the EIT spectra. In figure 6, we observe that the principal components of the EIT spectra all fall along a smooth curve in the new subspace. However, not only do the projections align on a single curve when $\phi = 0^\circ$, but they fall along a single curve for all values of ϕ . This relationship indicates that single manifestations of the EIT spectrum are not unique. As such, if ϕ and θ are unknown, it is impossible to determine the orientation of the magnetic field with only a single trace of the spectrum.

In order to accurately determine θ and ϕ simultaneously, the measurement protocol requires rotating the polarization and observing multiple spectra across variations in ϕ . As ϕ approaches 0° , rotations in the polarization will create significant variations in the EIT spectrum. Conversely, as ϕ approaches 90° , rotations in the polarization will cause minimal variations in the EIT spectrum. Changes in the amount of variation can be quantified with fluctuations in the principal components. This can be seen visually with figure 6, as the separation between clusters for $\theta = 0^\circ$ and $\theta = 90^\circ$ decreases as ϕ increases from $\phi = 0^\circ$ to $\phi = 90^\circ$. The ‘amplitude’ of the spectrum variation can be used to determine the approximate value of ϕ . When ϕ is known, a look-up table or similar tool can be used to determine the value of θ corresponding to the measured spectrum and known ϕ .

While controlling the polarization of the laser and observing multiple spectra is feasible, it increases the time duration of the measurement, and intrinsic drifts within the measurement apparatus begin to degrade the angular sensitivity. With initial tests incorporating this extended measurement approach, we found that the sensitivity of the technique became an order of magnitude worse, resulting in an accuracy on the order of $\pm 5^\circ$ for recovering ϕ and θ from multiple EIT spectra. In practice, it is more effective to simply rotate the polarization such that ϕ is locked to 0° , measure a single trace of the EIT spectrum, and predict the associated value of θ from the principal components of the measured trace. For this reason, this work has focused only on recovering θ from measurements of the EIT spectrum.

6. Conclusion

This paper explored the utility of classical and nonlinear PCA techniques within the context of AMO spectroscopy. PCA provides a means of exploring internal patterns within a collection of data without requiring any assumptions on the underlying physics involved. For the case of optical vector atomic magnetometry, PCA was used to identify the optimal testing configuration for the determination of the angle of the local magnetic field, and KPCA was utilized to train an SVR machine to correlate EIT spectra with specific angles. The demonstrated measurement accuracy was better than one degree for θ if $\phi = 0^\circ$, and better than three degrees if ϕ and θ are unknown. PCA primarily models the change in the contrast of the EIT resonances for different angles, but further sensitivity improvement is expected if additional fitting parameters are introduced to account for changes in the absolute field. The clustering behavior observed with this EIT magnetometer data set demonstrates that PCA techniques have great potential for addressing regression problems in AMO spectroscopy. Future work may consider a more rigorous uncertainty quantification for angle predictions, or an active shape model that considers only small portions of each spectrum when performing PCA.

Data availability statement

The data cannot be made publicly available upon publication due to legal restrictions preventing unrestricted public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

The reported research performed by JAM and ABM was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004) and funded by DARPA contract (81-110107). The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

Conflict of interest

The authors declare no conflicts of interest.

ORCID iDs

James A McKelvy  <https://orcid.org/0009-0007-0873-522X>
 Irina Novikova  <https://orcid.org/0000-0001-8433-9092>
 Eugeni E Mikhailov  <https://orcid.org/0000-0001-5555-5183>
 Mario A Maldonado  <https://orcid.org/0000-0001-9336-5823>
 Isaac Fan  <https://orcid.org/0000-0003-4414-7920>
 John Kitching  <https://orcid.org/0000-0002-4540-1954>
 Andrey B Matsko  <https://orcid.org/0000-0001-9939-9311>

References

- [1] Shirmohammadli V and Bahreyni B 2021 Machine learning for sensing applications: a tutorial *IEEE Sens. J.* **22** 10183–95
- [2] Long V J, Wang X, Zhou W, Zhang J, Dai D and Zhu G 2021 A comprehensive review of signal processing and machine learning technologies for UHF PD detection and diagnosis (I): preprocessing and localization approaches *IEEE Access* **9** 69876–904
- [3] Zhao W, Bhushan A, Santamaria A D, Simon M G and Davis C E 2008 Machine learning: a crucial tool for sensor design *Algorithms* **1** 130–52
- [4] Ha N, Xu K, Ren G, Mitchell A and Ou J Z 2020 Machine learning-enabled smart sensor systems *Adv. Intell. Syst.* **2** 2000063
- [5] Zhang W, Zhang R, Fu Y, Enemali G, Si J and Liu C 2021 Machine Learning Based Wavelength Modulation Spectroscopy for Rapid Gas Sensing *Proc. IEEE Int. Instrumentation and Measurement Technology Conf. (I2MTC)* pp 1–5
- [6] Goldenstein C S, Strand C L, Schultz I A, Sun K, Jeffries J B and Hanson R K 2014 Fitting of calibration-free scanned-wavelength-modulation spectroscopy spectra for determination of gas properties and absorption lineshapes *Appl. Opt.* **53** 356–67
- [7] Kong S G and Wu D H 2008 Signal restoration from atmospheric degradation in terahertz spectroscopy *J. Appl. Phys.* **103** 113105
- [8] Abdel-Aal R E 1998 Automatic fitting of Gaussian peaks using abductive machine learning *IEEE Trans. Nucl. Sci.* **45** 1–16
- [9] Hasegawa T, Matsuda S, Kumamoto A, Tsuchiya F, Kasahara Y, Miyoshi Y, Kasaba Y, Matsuoka A and Shinohara I 2019 Automatic electron density determination by using a convolutional neural network *IEEE Access* **7** 163384–94
- [10] Lohani S, Kirby B T, Brodsky M, Danaci O and Glasser R T 2020 Machine learning assisted quantum state estimation *Mach. Learn.: Sci. Technol.* **1** 035007
- [11] Kudyshev Z A, Shalaev V M and Boltasseva A 2020 Machine learning for integrated quantum photonics *ACS Photonics* **8** 34–46
- [12] You C, Quiroz-Juárez M A, Lambert A, Bhusal N, Dong C, Perez-Leija A, Javadi A, León-Montiel R de J and Magaña-Loaiza O S 2020 Identification of light sources using machine learning *Appl. Phys. Rev.* **7** 021404
- [13] Li L, Pan X, Yang H, Zhang T and Liu Z 2019 Supervised dictionary learning with regularization for near-infrared spectroscopy classification *IEEE Access* **7** 100923–32
- [14] Ede J M 2021 Deep learning in electron microscopy *Mach. Learn.: Sci. Technol.* **2** 011004
- [15] Frandsen A M A, Connor B V, Amersfoort J V and Smith E J 1978 The ISEE-C vector Helium magnetometer *IEEE Trans. Geosci. Electron.* **16** 195–8
- [16] Glassmeier K-H *et al* 2010 The fluxgate magnetometer of the BepiColombo Mercury planetary orbiter *Planet. Space Sci.* **58** 287–99
- [17] Hoeksema J T *et al* 2014 The helioseismic and magnetic imager (HMI) vector magnetic field pipeline: overview and performance *Sol. Phys.* **289** 3483–530
- [18] Connerney J E P, Espley J, Lawton P, Murphy S, Odom J, Oliverson R and Sheppard D 2015 The MAVEN magnetic field investigation *Space Sci. Rev.* **195** 257–91
- [19] Happer W and Tang H 1973 Spin-exchange shift and narrowing of magnetic resonance lines in optically pumped Alkali vapors *Phys. Rev. Lett.* **31** 273–6
- [20] Allred J C, Lyman R N, Kornack T W and Romalis M V 2002 High-sensitivity atomic magnetometer unaffected by spin-exchange relaxation *Phys. Rev. Lett.* **89** 130801
- [21] Budker D and M Romalis 2007 Optical magnetometry *Nat. Phys.* **3** 227–34
- [22] Gravrand O, Khokhlov A, Le Mouél J L and Léger J M 2001 On the calibration of a vectorial ⁴He pumped magnetometer *Earth, Planets Space* **53** 949–58
- [23] Yudin V I, Taichenachev A V, Dudin Y O, Velichansky V L, Zibrov A S and Zibrov S A 2010 Vector Magnetometry Based on Electromagnetically Induced Transparency in Linearly Polarized Light *Phys. Rev. A* **82** 033807
- [24] Cox K, Yudin V I, Taichenachev A V, Novikova I and Mikhailov E E 2011 Measurements of the magnetic field vector using multiple electromagnetically induced transparency resonances in Rb vapor *Phys. Rev. A* **83** 015801
- [25] Tula T *et al* 2021 Machine learning approach to muon spectroscopy analysis *J. Phys.: Condens. Matter* **33** 194002
- [26] Helfrecht B A, Cersonsky R K, Fraux G and Ceriotti M 2020 Structure-property maps with kernel principal covariates regression *Mach. Learn.: Sci. Technol.* **1** 045021
- [27] Pearson K 1901 LIII. On Lines and Planes of Closest Fit to Systems of Points in Space *London, Edinburgh Dublin Phil. Mag. J. Sci.* **2** 559–72
- [28] Scholkopf B, Smola A and Muller K R 1997 Kernel principal component analysis *Artif. Neural Netw.* **1327** 583–8
- [29] Scholkopf B, Smola A and Muller K-R 1998 Nonlinear component analysis as a kernel eigenvalue problem *Neural Comput.* **10** 1299–319
- [30] Wang Q, 2014 Kernel principal component analysis and its applications in face recognition and active shape models (arXiv:1207.3538)
- [31] Shalev-Shwartz S and Ben-David S 2014 *Understanding Machine Learning—From Theory to Algorithms* (Cambridge University Press)
- [32] Bishop C M 2006 *Pattern Recognition and Machine Learning* (Springer)
- [33] Gerbrands J J 1981 On the relationships between SVD, KLT and PCA *Pattern Recognit.* **14** 375–38130
- [34] Good I J 1969 Some applications of the singular decomposition of a matrix *Technometrics* **11** 823–31
- [35] Kim K I, Franz M O and Scholkopf B 2005 Iterative kernel principal component analysis for image modeling *IEEE Trans. Pattern Anal. Mach. Intell.* **27** 1351–66

- [36] Honeine P 2012 Online kernel principal component analysis: a reduced-order model *IEEE Trans. Pattern Anal. Mach. Intell.* **34** 1814–26
- [37] Zhou T and Peng Y 2020 Kernel principal component analysis-based gaussian process regression modelling for high-dimensional reliability analysis *Comput. Struct.* **241** 106358
- [38] Rousseeuw P J 1987 Silhouettes: a graphical aid to the interpretation and validation of cluster analysis *J. Comput. Appl. Math.* **20** 53–65
- [39] Angiulli G, Cacciola M and Versaci M 2007 Microwave devices and antennas modelling by support vector regression machines *IEEE Trans. Magn.* **43** 1589–92
- [40] Vapnik V N 1995 *The Nature of Statistical Learning Theory* (Springer)
- [41] Gunn S R 1998 Support vector machines for classification and regression, *ISIS Technical Report*, 14.
- [42] Ni J, Zhang C, Ren L and Yang S X 2012 Abrupt event monitoring for water environment system based on KPCA and SVM *IEEE Trans. Instrum. Meas.* **61** 980–9
- [43] Xuemei L, Lixing D, Jinhu L, Gang X and Jibin L 2010 A novel hybrid approach of KPCA and SVM for building cooling load prediction *Proc. 3rd Int. Conf. on Knowledge Discovery and Data Mining* pp 522–6
- [44] Pedregosa F et al 2011 Scikit-learn: machine learning in python *J. Mach. Learn. Res.* **12** 2825–30
- [45] Fleischhauer M and Scully M O 1994 Quantum sensitivity limits of an optical magnetometer based on atomic phase coherence *Phys. Rev. A* **49** 1973–86