Article

Optical Fiber Ball Resonator Sensor Spectral Interrogation through Undersampled KLT: Application to Refractive Index Sensing and Cancer Biomarker Biosensing

Daniele Tosi 1,2,*, Zhannat Ashikbayeva 1, Aliya Bekmurzayeva 1,2, Zhuldyz Myrkhiyeva 1, Aida Rakhimbekova 1, Takhmina Ayupova 1 and Madina Shaimerdenova 1

1 School of Engineering and Digital Sciences, Nazarbayev University, Nur-Sultan 010000, Kazakhstan; zhashikbayeva@nu.edu.kz (Z.A.); abekmurzayeva@nu.edu.kz (A.B.); zhuldyz.myrkhiyeva@nu.edu.kz (Z.M.); aida.belova@nu.edu.kz (A.R.); takhmina.ayupova@nu.edu.kz (T.A.); madina.shaimerdenova@nu.edu.kz (M.S.)

2 Laboratory of Biosensors and Bioinstruments, National Laboratory Astana, Nur-Sultan 010000, Kazakhstan

* Correspondence: daniele.tosi@nu.edu.kz

Abstract: Optical fiber ball resonators based on single-mode fibers in the infrared range are an emerging technology for refractive index sensing and biosensing. These devices are easy and rapid to fabricate using a CO$_2$ laser splicer and yield a very low finesse reflection spectrum with a quasi-random pattern. In addition, they can be functionalized for biosensing by using a thin-film sputtering method. A common problem of this type of device is that the spectral response is substantially unknown, and poorly correlated with the size and shape of the spherical device. In this work, we propose a detection method based on Karhunen−Loève transform (KLT), applied to the undersampled spectrum measured by an optical backscatter reflectometer. We show that this method correctly detects the response of the ball resonator in any working condition, without prior knowledge of the sensor under interrogation. First, this method for refractive index sensing of a gold-coated resonator is applied, showing 1594 RIU$^{-1}$ sensitivity; then, this concept is extended to a biofunctionalized ball resonator, detecting CD44 cancer biomarker concentration with a picomolar-level limit of detection (19.7 pM) and high specificity (30–41%).

Keywords: optical fiber biosensor; optical fiber sensor; ball resonator; optical fiber spherical tip; digital signal processing; Karhunen−Loève transform (KLT); cancer biomarker diagnostic

1. Introduction

Optical fiber biosensors make use of optical fiber waveguides for detecting biological analytes on the fiber surface, using suitable biofunctionalization methods that allow sensors to have excellent sensitivity and high specificity [1], and find important applications as immunosensors [2] in the diagnostic of cancer [2,3], cardiac [4], and urologic biomarkers [5], as well as in modern wearable devices [6]. Biosensors base their working principle on refractive index (RI) detection [7]; using a suitable functionalization method, it is possible to convert a refractometer into a biosensor and measure the response to a specific analyte as well as the specificity of detection.

An important set of optical fiber biosensors is based on single-mode fibers (SMFs) operating in the infrared range. The use of SMF fibers, in lieu of large-core or microstructured fibers or plastic fibers, yields a great simplicity in the fabrication and interrogation of the devices using telecom grade equipment [8]. Since SMF fibers have been widely standardized in terms of materials, geometrical size, and type of coatings (such as ITU-T G.652.D standard), it is possible to manufacture devices using the same equipment and components used in optical communications (for example, fiber splitters, lasers and LED sources, photodiodes, filters). In return, SMF-based devices working at infrared wavelengths have...
exceptionally low fiber losses that enable remote sensing, great accuracy which results in a low limit of detection and high figures of merit and have a great availability of splicing and interconnection tools.

The working principle of most fiber optic biosensors relies on the detection of the reflection or transmission spectrum of the sensor [7]. Within the spectrum, we can identify and monitor specific spectral features, that depend on the RI. The shape of the spectral envelope we can observe on the analyzer is directly tied to the type of sensor and is often referred to as a spectral “signature” or “fingerprint”. In each spectral signature, we can identify one or more features, such as reflection peaks or transmission dips, and track their wavelength or intensity to obtain an estimate of the RI change.

Fiber Bragg gratings (FBGs) represent the widest approach for RI sensing and biosensing using SMF fibers in the infrared range [9]. While a standard FBG is not RI sensitive, we can modify the geometry of the grating or the shape of the fiber to have highly sensitive biosensors. The most common approach is to use an etched FBG which behaves as a narrowband notch filter with a sensitivity depending on the fiber thickness [10]. Another important approach makes use of a tilted FBG, which encodes the RI into cladding modes propagated out of the fiber [11]; tilted gratings can also increase their sensitivity by etching the fiber waveguide [12] or using plasmonic effects generated in a thin metallic film [13]. Long-period gratings (LPG), on the other hand, are gratings having an extended period, and working as band-reject filters combining co-propagating fiber modes [14].

Interferometers are also a common type of RI sensors and biosensors, making use of structures that allow waves exposed to the external RI to interfere with different optical path differentials [15]. Fabry–Perot interferometers fabricated with SMFs use different fiber spans as mirrors, usually encapsulating an air gap [16]; structures based on a metallic mirror for high reflectivity [17] as well as low-finesse all-glass interferometers [18] have been reported.

The spectral fingerprints of gratings and interferometers have high fringe visibility and allow simple detection using signal processing methods common in data acquisition and analysis. FBGs have a sharp reflection peak with reflectivity that is tens of dB higher than the noise floor, and the RI is detected by measuring the wavelength shift. For tilted FBGs, the RI is encoded in the cladding modes, which are visualized as a spectral comb; the spectral features under investigation are usually the cut-off mode [19], very common in the sensing of large RI changes, or the amplitude of a resonant mode, which is more suitable for plasmonic gratings [13]. LPGs have a much wider spectrum, encompassing hundreds of nanometers of bandwidth, and a wavelength shift is observed as the RI increases [7]. Conversely, interferometers have a spectrum with a periodic pattern, with a fringe visibility that depends on the finesse of the cavity; the detection is performed by either measuring a spectral peak or dip, or performing the differential between multiple spectral features, such as in the case of dual RI/temperature sensors [4].

Recently, the use of fiber-optic ball resonators has been introduced [20,21]. This type of sensor operates on the same SMF fiber and infrared wavelength range as gratings and interferometers, but has a much faster and easier fabrication process, as we can reuse the same process developed for the manufacturing of diffractive tip lenses for microscopy [22]. Using a CO₂ laser splicer, this process has a high throughput and is implemented as a single step with a few seconds duration.

Ball resonators do not have a well-defined spectrum, and are rather low-reflectivity devices with fringe visibility in the order of ~1% or below; most importantly, it appears that each ball resonator is characterized by a different spectral fingerprint that is unique for the device and appears as a quasi-random reflection spectrum. Hence, unlike gratings or interferometers that have a well-defined spectrum that entirely depends on the parameters of the device (such as the Bragg wavelength for an FBG, or the cavity length for a Fabry–Perot interferometer), a ball resonator has a spectral fingerprint that we can approximate as a random signal, and changes in wavelength and amplitude in a similar fashion to the other RI sensors and biosensors.
The works presented so far propose a demodulation method for ball resonators that is based on feature extraction, by measuring the most significant spectral feature and tracking its amplitude [20] or pitch wavelength [21]. This method has been demonstrated both by analyzing the return loss [20], or each individual polarization for an enhanced sensitivity [23]. Feature extraction is, however, a sub-optimal method as it requires knowledge of the spectral response of the sensor, which cannot be achieved before the fabrication. A more ideal scenario for the detection of a ball resonator would rely on a method for demodulation that processes the reflection spectrum without having a predefined spectral feature, but can rather work with any spectral fingerprint.

In this work, we introduce a new demodulation method for ball resonators that is based on the KLT (Karhunen–Loève transform) [24] of the undersampled ball resonator spectrum. We show that this simple yet effective method can return an output that depends on the RI, and we show that we can use this method for RI sensing; subsequently, we extend this method to a typical case scenario of a fiber-optic biosensor, by detecting various concentrations of CD44 proteins. The increased CD44 protein level in serum can be an indicator of tumor burden and metastasis in several types of cancer, such as breast, gastric, and colon, and hence it can be used as a biomarker [25,26]. We show that the KLT can work regardless of the shape of the spectral fingerprint.

The paper is arranged as follows: Section 2 describes the fabrication process, the biofunctionalization, and the interrogation setup for the sensors used in this work; Section 3 shows the principle of operation of the KLT applied to the ball resonator detection; Section 4 shows the results of the work, by first analyzing the performances of the KLT and then showing the evidence of RI sensing and CD44 biosensing case scenarios; Section 5 briefly discusses the impact of KLT in sensor detection; finally, Section 6 draws the conclusions.

2. Materials and Methods

2.1. Fabrication of the Ball Resonator Sensor

A sensor based on optical fiber with a spherical tip at the end-point was fabricated on a standard single-mode fiber (SMF-28) using a CO\textsubscript{2} laser splicer (Fujikura L2M-100). This was done by the following process: aligning two fibers, splicing, and then subjecting the produced structure to high laser power which is specific to the fabrication equipment. As a result, a spherical tip at the end of the fiber was formed while undergoing breaking close to the splicing point. Fabrication parameters are shown in Table S1.

The photograph of the sensor fabricated and subsequently biofunctionalized for CD44 detection is shown in Figure 1. The cross-sections show that the diameters are 508.2 \( \mu \text{m} \) along the horizontal direction, and 507.6 \( \mu \text{m} \) along the vertical direction; the estimated eccentricity is equal to 0.049.

2.2. Experimental Setup

The experimental setup used in the whole process of interrogation of ball resonators is shown in Figure 2.

The interrogator is an optical backscatter reflectometer (OBR; OBR4600, Luna Inc., Roanoke, VA, USA), with the following parameters: scan range 1525–1610.6 nm; gain 0 dB; resolution bandwidth 0.257 GHz.

Spectra were filtered with a low-pass filter: Chebychev type-I filter, 7th order, 0.0084 normalized cut-off frequency. After filtering, and removing the transient of the filter, the analysis was carried out on 40,000 samples, wavelength range 1528.7–1580.1 nm.

All evaluations of KLT performance figures and computation times were implemented in MatLab (Mathworks, Natick, MA, USA), running on a MacBook Pro 2.6 GHz, 16 GB RAM.
In this work, we introduce a new demodulation method for ball resonator sensors that is based on the KLT (Karhunen−Loeve transform) 
 of the undersampled ball resonator spectrum. We show that this simple yet effective method can return an output that depends on the RI, and we show that we can use this method for RI sensing; subsequently, we extend this method to a typical case scenario of a fiber-optic biosensor, by detecting various concentrations of CD44 proteins. The increased CD44 protein level in serum can be used as a biomarker for cancer diagnosis [25,26]. We show that the KLT can work regardless of the shape of the spectral fingerprint.

2.2. Experimental Setup

The experimental setup used in the whole process of interrogation of ball resonators is shown in Figure 2. Figure 1. Photograph of the fabricated ball resonator under an optical microscope; (a) cross-section along the horizontal plane; (b) vertical plane.

![Figure 1. Photograph of the fabricated ball resonator under an optical microscope; (a) cross-section along the horizontal plane; (b) vertical plane.](image)

![Figure 2. Experimental set-up used in this study; (A) fabrication of ball resonator sensor on a CO₂ splicing machine; (B) interrogation of the sensor with optical backscatter reflectometer.](image)

2.3. Refractive Index Calibration

Refractive index (RI) calibration was done by measuring the change in spectra to RI change using solutions with different RI values. Several concentrations of sucrose in water (w/v %) starting from 10% to 13.53% in 9 increments of ~0.49% corresponding to RI values from 1.3478 to 1.3539 for a total change of $6.1 \times 10^{-3}$ RIU in steps of $5.86 \times 10^{-4}$.

2.4. CD44 Functionalization, Measurement, and Specificity

The surface of the ball resonator underwent cleaning with Piranha solution (4:1 = H₂SO₄:H₂O₂) followed by silanization with 1% aminopropyltrimethoxysilane (APTMS) utilized to obtain amine groups on its surface for the further attachment with the gold. The gold sputtering machine (Q150T Plus, Quorum Technologies Ltd., Lewes, UK) allowed coating the surface of the ball resonator with the gold at a thickness of 30 nm followed by annealing at high temperature (200 °C) for 2 h. Furthermore, the gold-coated...
sensor was pretreated with 95% 11-Mercaptoundecanoic acid (MUA) for 16 h at 2–4 °C and activated with EDC/NHS (1-ethyl-3-carbodiimide hydrochloride/N-Hydroxysuccinimide) for 15 min. Finally, the ball resonator was incubated on a shaker for 30 min with 8 mg/mL anti-CD44 antibody and blocked with 1% BSA.

The sensor functionalized with CD44 antibody was employed further for the CD44 protein detection. Optical backscattering reflectometer (OBR) was utilized for the signal detection over the CD44 protein measurement analysis at the concentration range from 0.006 nM to 100 nM. For detection purposes, the functionalized ball resonator was immersed in the vial containing the 200 µL CD44 protein solution.

To conduct the specificity analysis the IL-4 protein measurement at the same concentration range was utilized as a control. To validate the performance of the obtained ball resonators, the experiments were conducted at the same conditions and parameters.

3. Interrogation of Ball Resonators with KLT

The reflection spectrum $R_{Ref}(\lambda)$ of a ball resonator, where $\lambda$ is the wavelength in the infrared range, can be approximated in reference conditions as [20]:

$$R_{Ref}(\lambda) = c \cdot s(\lambda),$$

(1)

where $c$ is a constant, that defines the baseline level of the reflectivity, and $s(\lambda)$ is the spectral signature of the ball resonator, which appears as a fluctuating spectral envelope having a shape different for each sensor. When the RI changes, the reflection spectrum changes in amplitude, by the effect of the variation of reflectivity at the interface between the glass fiber sphere and the outer medium, and experiences a slight wavelength shift as a consequence of the phase change of the reflection coefficient at each reflection. The measured spectrum $R_{Meas}(\lambda)$, therefore, can be approximated as:

$$R_{Meas}(\lambda) = (c + \Delta c) \cdot s(\lambda - \Delta \lambda),$$

(2)

where $\Delta c$ is a term that shows the change of spectral amplitude, and $\Delta \lambda$ is a wavelength shift. In the case of a small change of RI, in the order of $10^{-3}$ or below, as shown in [20], we can approximate the amplitude and wavelength shifts as linear terms:

$$R_{Meas}(\lambda) \approx (c + k_a \Delta n) \cdot s(\lambda - k_w \Delta n),$$

(3)

where the terms $k_a$ and $k_w$ represent the amplitude sensitivity, expressed usually in dB/RIU, and the wavelength sensitivity, typically in nm/RIU, respectively. As shown in [20,23], amplitude and wavelength sensitivity vary in each sensor and depend on the shape of the resonator.

The Karhunen–Loeve transform (KLT), implemented as in Maccone’s algorithm [24,27,28] is then applied on each measured spectrum, treated as a noisy signal. The approach described by Maccone is to decompose the measured spectrum $R_{Meas}(\lambda)$ into a set of orthonormal bases:

$$R_{Meas}(\lambda) = \sum_{n=1}^{\infty} Z_n \cdot \phi_n(\lambda),$$

(4)

where the coefficients $Z_n$ represent the “magnitude” of each $n$-th basis, and the terms $\phi_n(\lambda)$ are orthonormal bases. The orthonormality condition, computed between the minimum ($\lambda_{min}$) and maximum ($\lambda_{max}$) wavelength values of the optical interrogator, reads as:

$$\int_{\lambda_{min}}^{\lambda_{max}} \phi_n(\lambda) \phi_m(\lambda) d\lambda = \delta_{mn},$$

(5)

with $\delta_{mn}$ being the Kronecker delta.

The difference between the KLT and other transform methods is that while in classical algorithms for signal processing the set of basis $\phi_n$ is predefined, in the KLT they are entirely dependent on the signal under analysis. For example, in the Fourier transform
the set of orthonormal basis is made of sine functions of different frequency; in wavelet transforms the set of bases are wavelet functions having different scale and pitch.

However, since the spectra of the ball resonators are not inferred prior to their fabrications, since there is a very poor correlation between the fabrication parameters and the reflection spectrum as clearly observed in previous works [20], the KLT offers a superior method for spectral processing since it can form a set of orthonormal basis that is entirely data-dependent. As a consequence, the coefficients \( Z \) are random variables that measure the magnitude of each basis [29].

To resolve the KLT, we need to find the statistical bases \( \phi_n(\lambda) \); Maccone demonstrated in his work [28] that the set of bases \( \phi_n(\lambda) \) is the eigenfunction of the autocorrelation of the signal \( R_{\text{Meas}}(\lambda) \). The problem of finding the eigenfunctions is integral, and since we can treat the spectral signatures as quasi-random signals, rather than analytical functions, we cannot have a closed-form solution. Maccone demonstrated this method to estimate the frequency of a sine function buried into noise [24], but this problem is different from the one under analysis as, in the case of ball resonator spectra, we need to measure the amplitude and wavelength shifts of an almost flat spectrum with extremely low fringes.

However, from an experimental point of view, we do not work with continuous signals, but rather the OBR interrogator provides a digitally sampled signal. In this case, we can translate the continuous-time processing into a discrete version of the KLT, which is much faster to implement.

We consider the measured spectrum as \( S[m] \) where \( m = 1, 2, \ldots, M \) is the sample number of a digital signal. Whereas in the continuous KLT we need to identify the eigenfunctions of the autocorrelation, in the digital domain the KLT aims at finding the eigenvalues of the autocorrelation matrix. Hence, in other words, the bases \( \phi_n \) become the eigenvectors, and the coefficients \( Z_n \) become the eigenvalues of the autocorrelation matrix.

Computationally, we first apply to compute the autocorrelation matrix \( R_{SS} \) of the spectrum:

\[
R_{SS} = S[m]S^T[m] = \begin{bmatrix}
S[1]S[1] & S[1]S[2] & \cdots & S[1]S[m] \\
S[2]S[1] & S[2]S[2] & \cdots & S[1]S[m-1] \\
\vdots & \vdots & \ddots & \vdots \\
S[m]S[1] & S[m-1]S[2] & \cdots & S[m]S[m]
\end{bmatrix}, \tag{6}
\]

where the operator \( T \) stands for transpose. We note that, from the notation of Equation (6), we omitted the expected value operator, since in our experimental arrangement we do not have a possibility to estimate the expected values of \( S[m] \) but rather we have a single realization of the spectral measurement, and we generate the autocorrelation matrix from this signal alone.

To estimate the eigenvalues and eigenvectors, we can apply the singular value decomposition (SVD) [30]:

\[
R_{SS} = VDV^{-1}, \tag{7}
\]

in which: \( D \) is a diagonal matrix that has all eigenvalues on the main diagonal; \( V \) is the eigenvectors matrix, that has on each \( i \)-th column the eigenvector corresponding to the \( i \)-th eigenvalue, i.e., the \( D_{ii} \) element of \( D \). Since \( R_{ss} \) is a symmetric matrix, the eigenvalues are all real. We label the eigenvalue string as \( \xi = [\xi_1, \xi_2, \ldots, \xi_M] \), such that \( \xi_i \) is the \( i \)-th element of the diagonal of \( D \).

Through implementing the SVD, we can sort the matrices \( D \) and \( V \) such that the eigenvalues are sorted in ascending order:

\[
\xi_M \geq \xi_{M-1} \geq \ldots \geq \xi_1. \tag{8}
\]

The SVD is the key step of the KLT because the eigendecomposition allows the identification of one eigenvalue having a magnitude much larger than all the other ones. In this sense, the largest eigenvalue \( \xi_M \) represents the main outcome of the KLT.
The main problem in solving Equation (7) is computational, as $R_{SS}$ is an $M \times M$ matrix, and spectral acquisition on the OBR is performed with $M = 65,536$ samples. The eigen-decomposition process implies solving an $M$-size linear system of equations, which is a $\sim O(M^3)$ process. Hence, we adjust the original KLT algorithm by undersampling the acquired reflection spectra, reducing the complexity; this can keep the computation time below 1 s while guaranteeing a reliable output.

In formulas, we downsample the spectrum by a rate $K$, obtaining a new signal $T[p]$ with length $P = M/K$:

$$T[p] = T[1, 2, \ldots, P] = [S[1], S[K + 1], S[2K + 1], \ldots, S[(P - 1)K + 1]$$

and then repeat Equations (6) and (7) to the signal $T$, rather than on the original spectrum $S$, by first building the $P \times P$ autocorrelation matrix and then determining the $P$-size eigenvector string. In this case, the main eigenvalue is $\xi_P$.

The undersampled KLT accomplishes two tasks: it allows the analysis of the whole spectrum of the ball resonators, hence having a much better chance to build the eigenbasis that show the maximum sensitivity to the RI change; at the same time, it keeps the computational complexity low, allowing fast estimation of the KLT output.

As shown by Maccone [28], for an $M$-size KLT calculation, the value of the main eigenvalue $\xi_M$ grows linearly:

$$\frac{\partial \xi_M}{\partial M} = \text{const.} \quad (10)$$

Therefore, we can normalize the main eigenvalue by the length of the analyzed signal, to obtain an output that is independent of the sampling rate (which can therefore allow comparison of the spectra acquired with a different number of samples). Here, we label as $\omega$ the output of the KLT, defined as:

$$\omega = \frac{\xi_P}{P} = \frac{K}{M} \xi_P. \quad (11)$$

To summarize, the undersampled KLT method works as follows (and also as shown in Figure 3):

- Acquire the reflection spectrum of the ball resonator;
- Low-pass filter to clean the higher frequency noise, and obtain the signal $S[m]$ with $M = 40,000$ samples.
- Downsample $S[m]$ by a rate $K$, obtaining the new signal $T[p]$ with size $P = M/K$ samples;
- Build the $P \times P$ autocorrelation matrix of $T[p]$;
- Apply the SVD to the autocorrelation matrix, and identify the eigenvalue string;
- Compute the KLT output $\omega$ as the highest rank eigenvalue normalized by $P$. 
\[ S[1], S[K+1], S[2K+1], \ldots, S[(P-1)K+1] \]

and then repeat Equations (6) and (7) to the signal \( T \), rather than on the original spectrum \( S \), by first building the \( P \times P \) autocorrelation matrix and then determining the \( P \)-size eigenvector string. In this case, the main eigenvalue is \( \xi_P \).

The undersampled KLT accomplishes two tasks: it allows the analysis of the whole spectrum of the ball resonators, hence having a much better chance to build the eigenbasis that show the maximum sensitivity to the RI change; at the same time, it keeps the computational complexity low, allowing fast estimation of the KLT output.

As shown by Maccone [28], for an \( M \)-size KLT calculation, the value of the main eigenvalue \( \xi_M \) grows linearly:

\[
\frac{\partial \xi}{\partial M} = c \text{ const.}
\]

(10)

Therefore, we can normalize the main eigenvalue by the length of the analyzed signal, to obtain an output that is independent of the sampling rate (which can therefore allow comparison of the spectra acquired with a different number of samples). Here, we label as \( \omega \) the output of the KLT, defined as:

\[
\omega = \frac{1}{P} \xi_\omega
\]

(11)

To summarize, the undersampled KLT method works as follows (and also as shown in Figure 3):

- Acquire the reflection spectrum of the ball resonator;
- Low-pass filter to clean the higher frequency noise, and obtain the signal \( S[m] \) with \( M = 40,000 \) samples.
- Downsample \( S[m] \) by a rate \( K \), obtaining the new signal \( T[p] \) with size \( P = M/K \) samples;
- Build the \( P \times P \) autocorrelation matrix of \( T[p] \);
- Apply the SVD to the autocorrelation matrix, and identify the eigenvalue string;
- Compute the KLT output \( \omega \) as the highest rank eigenvalue normalized by \( P \).

**Figure 3.** A flow chart of the undersampled KLT is used to demodulate the spectrum of the ball resonators.

### 4. Results

#### 4.1. Operation and Performances of the KLT

In this section, we analyze the typical output of the KLT, and how the performances are affected by the main parameters. We analyze the sensor described in Section 2, exposed to a reference RI value, as a benchmark.

Figure 4 shows the operation of the KLT, with \( K = 200 \) downsampling rate (201 samples under analysis). The first chart shows the reflection spectrum of the ball resonator, which appears in analogy with previous works [20] and shows an almost random pattern of shallow peaks, with amplitude <1 dB.

By applying the KLT, we can obtain the eigenvalues and eigenvectors of the autocorrelation matrix of the downsampled signal. The eigenvector corresponding to the maximum eigenvalue is shown in the second chart. We see that the envelope resembles the flipped spectrum of the ball resonator, which shows that the SVD operation can single out the main eigenbasis of the ball resonator spectrum.

In Figure 4c, we show the eigenvalue string, with all values ordered in ascending order; for better visibility, we display the data in logarithmic units. We can see that the main eigenvalue stands out from the other values: the highest rank eigenvalue is equal to \( 4.8 \times 10^5 \), while the second eigenvalue \( 7.7 \times 10^{-11} \). This means that the KLT entirely encodes the ball resonator spectrum in the highest rank eigenvalue, whereas all the other eigenvalues are 15 orders of magnitude below this value.

This observation was reported by Maccone in the problem of extracting a signal buried in noise [24]; in this work, as noise is not very dominant in the optical spectrum, we have a wider divergence between the highest eigenvalue and the other ones. In the case of a spectrum with more noise (for example, by omitting the low-pass filter), we would also observe a rise of the lower-rank eigenvalue, while the main one would maintain a similar value and shape of its eigenvector.
4. Results

4.1. Operation and Performances of the KLT

In this section, we analyze the typical output of the KLT, and how the performances of the KLT depend on the choice of undersampling rate $K$: by choosing a higher value of $K$, the KLT is faster, as the SVD operates on a matrix with a size proportional to $\sim 1/K$. On the other hand, by working with severely undersampled spectral signals, we lose spectral features that can become part of the KLT eigenbasis.

We can evaluate experimentally the range of $K$ values that optimizes the KLT; this process is shown in Figure 5. The left chart shows the computation time, averaged over 100 different KLT calculations, for $K$ values ranging from 45 ($P = 889$) to 2000 ($P = 21$). We can see that the computation time increased significantly for lower values of $K$, following an almost linear pattern when data are displayed in logarithmic units. A good selection was obtained by imposing a fast computation, which would not slow down the acquisition rate of the OBR. Considering that the OBR works at $\sim 1$ Hz on the continuous acquisition on the full bandwidth, by imposing a computation time $\leq 10$ ms we guaranteed that the data processing time was $<1\%$ of the acquisition sampling time; this value corresponds to $K = 172$, and therefore a choice of $K > 172$ satisfies the fast computation criteria. Considering that the data processing was implemented on a capable processor, this criterion would also guarantee a rapid computation in most experimental conditions. By imposing, instead, a looser constraint on the acquisition time as $\leq 100$ ms, we obtained $K > 66$.

The second chart shows how the KLT output $\omega$, computed on the same dataset, changed as a function of $K$. We have seen in Equation (10) that the output of the KLT should be constant with respect to the size of the autocorrelation matrix; a good evaluation of the quality of the KLT is to evaluate the accuracy of this statement, and therefore evaluate the change of $\omega$ from the reference level (hereby obtained with the lowest value $K = 45$) as $K$ increases. At first, we notice that in all conditions the variation of $\omega$ was limited to 0.46% change from 2393.3 to 2404.1, and therefore in all conditions, even working on very small matrices, we can guarantee that the KLT has good effectiveness. Empirically, we can define it as a good choice of $K$, a value that limits the deviation of $\omega$ to $\leq 0.1\%$ from the reference value; this corresponds to $K < 393$.

Hence, a good choice according to this experimental analysis is $172 < K < 393$; in the following analysis, we always set $K = 200$. 

Figure 4. Overview of the operation of the undersampled KLT applied to the spectrum of a ball resonator. (a) The reflection spectrum of the ball resonator, acquired by the OBR and processed. (b) The eigenvector corresponding to the highest rank eigenvalue, acquired over 201 samples. (c) Eigenvalue string, reporting the magnitude (in logarithmic units) of all 201 eigenvalues sorted in ascending order.

The performance rates of the KLT depend on the choice of undersampling rate $K$.
The sensitivity can be estimated by applying a linear fit to the $\omega$ vs. RI calibration function; the resulting sensitivity was $-1593.6 \text{ RIU}^{-1}$, obtained with the coefficient of determination $R^2 = 0.9857$ which showed an excellent agreement between the $\omega$ decrease and the RI increase, following a path similar to other fiber-optic refractometers interrogated on a narrow RI range [7].

The KLT, as shown on the right chart, was, however, effective in returning an output that had a clear RI dependence, despite the lack of clear spectral features. The KLT output $\omega$ changed linearly as the RI increased, following a path similar to other fiber-optic refractometers interrogated on a narrow RI range [7].

The KLT, as shown on the right chart, was, however, effective in returning an output that had a clear RI dependence, despite the lack of clear spectral features. The KLT output $\omega$ changed linearly as the RI increased, following a path similar to other fiber-optic refractometers interrogated on a narrow RI range [7].
and the increase of RI. This corresponds to a decrease of $\omega$ equal to 0.066% per each RI increase of $10^{-3}$ RIU.

4.3. Biosensing Capability: Detection of CD44

Figure 7 shows the response of the functionalized ball resonator to the CD44 protein, with concentrations ranging from 6 pM to 100 nM. The first chart shows the spectral response of the sensor. In comparison to Figure 6, we observe that the spectrum maintained a similar envelope with a smoother pattern near the peaks, but with a much higher reflectivity value due to the gold-coating process, which increases the wall reflectivity at the interface between the spherical tip and the outer medium. Once again, we observe that the detection was hampered by the uneven shape of the spectral fingerprint, which prevented a simple feature extraction algorithm to work.

![Figure 7](image)

**Figure 7.** Detection of CD44 protein with a biofunctionalized gold-coated ball resonator. (a) The reflection spectrum of the ball resonator for different CD44 concentrations, ranging from 6 pM to 100 nM; (b) KLT response for each concentration, reporting the standard deviation of data acquired over 10 min, the log–linear fit, and the LoD determination.

By processing the spectral envelope with the undersampled KLT, on the other hand, we obtained a much clearer response; this is shown in Figure 7b, where we report the change of $\omega$ as related to the CD44 concentration change. Hereby, the datapoint represents the average of the sensor response over 10 min of acquisition, while the error bars report $\pm 3$ times the standard deviation, as commonly performed for fiber optic biosensors [1].

We observed that the baseline value of $\omega$ changed from the previous RI calibration; this is due to the different spectral shape, and particularly the different change of spectrum observed over the acquisition window; for the RI calibration this change was 5.2 dB, while after biofunctionalization (and particularly, after gold coating) the change was limited to 4.0 dB. However, once again we have a clear pattern that increases with the concentration of CD44. The standard deviation appears to be lower at the lowest concentrations, while it tends to increase at higher concentrations reaching its maximum value at 25 nM.

The quality of the detection is shown by a log–linear fit, that relates the increase of $\omega$ to a $10 \times$ increase of CD44 concentration; the fit shows that $\omega$ increases by 4.844 for each $10 \times$ concentration increase, with an $R^2$ parameter equal to 0.9556 which shows a good match between experimental data and the fit.

We can also infer the limit of detection (LoD) of this sensor, by using the method reported by Chiavaioli et al. [7]: considering as blank level the lowest concentration (6 pM) and adding three times the worst-case standard deviation, we can obtain the level of $\omega$ that defines the LoD (indicated in red in Figure 7b): this level corresponds to 19.7 pM, which shows that the proposed method based on KLT can successfully demodulate the spectral
fingerprint with shallow and uneven patterns into a biosensor with a clear response, and LoD at picomolar-level concentration.

The specificity of the sensor has been evaluated by comparing the response to CD44 to the response to a control (IL4). Specificity tests have been performed by analyzing the response to CD44 at four different concentrations (390 pM, 1.56 nM, 25 nM, 100 nM), compared to the control analyte IL4 detected at the same concentration levels.

The results of the specificity analysis are shown in Figure 8, where we compare the change of KLT output $\omega$, calculated from the reference condition (6 pM) observed for both CD44 and control IL4. We observe that the response to control is approximately 2.5 times inferior to the response to CD44, which validates the concept of the sensor. The specificity, observed between 390 pM and 100 nM, ranges from 30% to 41%, computed as the response to control normalized by the response to the CD44 analyte. This value is in analogy with other CD44 biosensors with lower signal (impedance decrease rate and current change, respectively) for control proteins [31,32], but these sensors were capable of working at lower concentrations.

![Figure 8](attachment:image.png)

**Figure 8.** Evaluation of the specificity of the CD44 biosensor, using the KLT interrogation method. The chart shows the percentage change of the KLT output $\omega$, calculated from the reference condition (6 pM, lowest concentration), for 4 different concentrations (390 pM, 1.56 nM, 25 nM, 100 nM) of the CD44 protein analyte compared to a control (IL4).

### 4.4. Comparison between KLT and Feature Extraction

As a key validation of the KLT capabilities, we compared its performance to the spectral feature extraction and analysis method that has been proposed in previous works [20] for ball resonator demodulation. This benchmark assessed the quality of the spectral detection by comparing methods providing different outputs, by normalizing the responses in calibration with a known response.

For the KLT, as previously reported, we evaluated the change of the main normalized eigenvalue $\omega$, computed by processing the reflection spectrum acquired through the OBR. For the feature extraction (FE) method, similarly to [20] we scanned the same reflection spectrum (filtered with the same low-pass digital filter) and measured the reflectivity change of the most significant feature. As observed in [20], the analysis of the intensity change tended to return a better accuracy (i.e., better $R^2$) than the measurement of wavelength shift, when the detector was based on an accurate power meter as the OBR.

Since both responses are based on distinctive methods and different measurement units, we normalized their responses between 0 and 1, using a simple normalization as in the formula below:

$$Y = \frac{y - y_{\text{min}}}{y_{\text{max}} - y_{\text{min}}}$$  \tag{12}

where $Y$ is the normalized response, $y$ is the output of each demodulation method (KLT and feature extraction), and $y_{\text{max}}, y_{\text{min}}$ are the maximum and minimum of the responses for
each method ($\omega$ for then KLT, the spectral intensity for the FE). This normalization allows us to compare the different methods in the same framework.

The calibration scenario was the same as presented for RI calibration; this was a good choice because, as shown previously for several fiber-optic biosensors [2], for a small RI change (below $10^{-2}$ RIU) we can perform a small-signal analysis in which we can approximate that the sensor response changes linearly with the RI. As such, our analysis consisted of 10 data points of RI between 1.3478 to 1.3539.

In order to evaluate the quality of the detection, we performed a linear fit between the RI change and the normalized response $Y$. As a figure of merit, we consider the coefficient of determination $R^2$ between the datapoints and the fitted data; a high $R^2$ ($>0.95$) ensures that the method correctly demodulates the sensor, while a low value of $R^2$ shows that datapoints are too scattered and therefore the demodulation method fails to correctly identify spectral changes.

The comparison between KLT and FE, presented for the previously described resonator, is shown in Figure 9. We see that while the KLT correctly identified the RI change with a reliable method, returning an $R^2$ coefficient of 0.9858, the standard FE method achieved only an $R^2$ value of 0.6257, hence failing to correctly demodulate the signal. The main reason is that while the KLT isolates the main eigenvector of the spectrum, which encodes the spectral change, the shallow spectral peaks overlapped with noise prevent the FE from achieving a clean detection, which results in poor calibration quality.

![Figure 9](image-url)  
**Figure 9.** Comparison between KLT and feature extraction (FE) methods for demodulation of the ball resonator. The chart shows the normalized output ($Y$) as a function of the RI, obtained for each method. Datapoints have been interpolated with a linear fit; the $R^2$ coefficient of determination shows the quality of the agreement between measured data and the fit.

We can further extend the comparison between KLT and FE methods, in order to provide a more encompassing view of the capabilities of the KLT compared with the standard feature extraction and analysis. For this task, we fabricated five additional ball resonators, using the same method described in Section 2.1, and having a diameter ranging from 290 $\mu$m to 532 $\mu$m (respectively: 290, 466, 515, 521, 532 $\mu$m). Reflection spectra were acquired and measured using the same method described in Section 2.2, and further processed using the KLT and the FE methods, and each output was normalized in order to provide a comparison between the methods.

We can observe, from the view in Figure 10, that in all the cases, the KLT was able to process the spectral fingerprints of the ball resonators with success, returning an $R^2$ coefficient ranging between 0.965 and 0.993; conversely, the FE method performed a correct ($R^2 > 0.95$) identification of the spectral feature only in one case ($R^2 = 0.984$), while for all
the other four cases the coefficient of determination ranged between 0.879 and 0.913, below the threshold for correct identification. Of all the five ball resonators used in this study, the KLT method returned a better $R^2$ than the FE method, at no computational expense, since, in both cases, the computation time is instantaneous (<10 ms).

**Figure 10.** Comparison between KLT and feature extraction, for 5 different resonators. The charts report the normalized output $Y$ for each method, as a function of refractive index (10 RI datapoints from 1.3478 to 1.3539). All ball resonators were fabricated using the same method based on CO$_2$ laser splicer, and had different diameters: (a) 290 $\mu$m; (b) 466 $\mu$m; (c) 515 $\mu$m; (d) 521 $\mu$m; (e) 532 $\mu$m.

Overall, this benchmark shows that the KLT adapts the output to each spectral signature, without inputting a priori information into the detection; on the other hand, the FE method works well only in ball resonators that exhibit a clear spectral feature, which leads to a much inferior rate of successful detection.

5. Discussion

Optical fiber ball resonators are a powerful tool for refractive index sensing [20] and biosensing [21], as they combine a single-mode fiber-optic structure, which allows the use of telecom-grade interrogators with high precision and wavelength sensitivity, a high sensitivity rating comparable with plasmonic grating sensors [13] and U-bent fiber biosensors [33], and a splicing recipe that can be fabricated in large volume. In particular, the manufacturing process requires a very short fabrication time, of just a few seconds duration, reusing a method developed for the fabrication of diffractive lenses with fibers having a larger core [34].

In this scenario, the KLT plays a significant role because it deals with the main weakness of ball resonators used for RI sensing and biosensing, which is the quasi-random pattern of the spectral signature of each ball resonator, which also appears to be not significantly related to geometrical parameters, such as size and eccentricity.

In particular, the KLT operates as a "black box", taking as input the reflection spectra and returning an output that is neither an intensity change nor a wavelength shift, but it directly links to the change of intensity and wavelength of the spectrum. While refractometers and biosensors express their performance figures in terms of sensitivity to wavelength shift (usually expressed in nm/RIU) or intensity change (expressed in dB/RIU), the KLT applied to the ball resonators reports a sensitivity rating in terms of variation of $\omega$ as a function of the RI, which can be expressed as RIU$^{-1}$ or in %/RIU. Nevertheless, since the quality of the detection is the most important factor in the demodulation of the ball resonator, we can observe that the KLT achieves a performance rating superior to the
spectral feature extraction and analysis. In particular, when detecting the RI in a small range, the KLT has always succeeded in detecting the spectrum of ball resonators having a reflectivity level sufficient to perform a detection (return loss > −70 dB), while the FE achieves a successful detection in only a few resonators (about one successful detection for every six devices fabricated with sufficient reflectivity).

The KLT method finds important applications in signal processing, as it is a versatile method for analyzing complex signals. This method was originally developed in order to extract very weak signals from noise, using the bordered autocorrelation method [28] that allows the complexity to be reduced into incremental contributions. This method deals with signals having a large sample size; however, the KLT has recently found applications in coarsely sampled signals [35], exploiting its capability to extract signals with lower noise contribution but having low sample sizes. Table 1 reviews some of the applications of KLT, in its original implementation and more recently in fiber optic sensing.

### Table 1. Review of some recent implementations of KLT in signal processing related to sensing data and analysis, and practical applications.

| Reference        | KLT Implementation                              | Application                                  | KLT Output                           |
|------------------|------------------------------------------------|---------------------------------------------|--------------------------------------|
| Maccone 2010 [24] | KLT—Bounded autocorrelation matrix              | Extract a weak narrowband tone buried in white noise | Extraction of the principal eigenfunction |
| Maccone 2006 [29] | KLT of Brownian motion                         | Optimization of special relativity computations | Rescaled-time eigenfunctions          |
| Lee 1999 [36]     | KLT preprocessing of discrete cosine transform  | Multispectral image compression             | Principal component analysis          |
| Sharma 2018 [37]  | KLT fused to discrete wavelet transform         | Cryptography: hiding imperceptible data in images | Eigenfunctions of the SVD             |
| Tosi 2015 [35]    | KLT-SVD of coarsely sampled signals            | Interrogation of fiber Bragg grating sensors | Principal eigenvalue                  |
| This work         | Undersampled KLT                               | Interrogation of optical fiber sensors having quasi-random spectrum | Normalized principal eigenvalue       |

This work proposes to convert the KLT from the analysis of long signals, which is a high-computational complexity problem, into a low complexity problem that can be solved in few milliseconds by using a downsampling routine. We expect that, through analyzing the main eigenvalue that constitutes the most important feature of the spectrum, we can even further increase the performance, particularly when looking at interrogating ball resonators with analyzers with a lower cost and inferior resolution bandwidth. In terms of application, the key benefit of the KLT, besides improving the detection, is that the user should no longer search for the most significant features, which might also change when the medium surrounding the sensor varies, but the KLT automatically identifies the most significant feature; then, by considering only the main eigenvalue, this is singled out from the surrounding noise without having to apply any prior knowledge of the sensor spectra.

The availability of a powerful detection method, such as the KLT, is an important asset for interrogating sensors fabricated with CO₂ laser splicers [38,39], which yield rapid and repeatable fabrication but with spectral characteristics that are often hard to process with an FE method.

### 6. Conclusions

In conclusion, we presented a method based on the Karhunen–Loeve transform for the interrogation of optical fiber ball resonators used in refractive sensing and biosensing. Optical fiber ball resonators have the advantage of a single-step rapid and reliable fabrication process, which can make a significant step up in terms of volume manufacturing.
The main drawback, however, is that the spectrum has a quasi-random pattern, which significantly hampers the detection method.

The undersampled KLT method presented in this paper shows great effectiveness in the demodulation of the reflection spectra of the ball resonator, outperforming the feature extraction methods used so far [20]. In particular, its main strength is that the eigendecomposition method at its core can self-adapt to any spectrum, which fits very well the almost random nature of the spectral fingerprints observed in the ball resonators. By optimizing the downsampling rate, the KLT method is both accurate and computationally fast.

We presented the effectiveness of the KLT and its application to RI sensing, achieving a sensitivity rate of 1593.6 RIU$^{-1}$; subsequently, we validated its application to biosensing by functionalizing one ball resonator to CD44 biomarker detection, achieving a 19.7 pM limit of detection and a ~40% specificity in the worst-case scenario.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/s21206721/s1, Table S1: The fabrication parameters of the ball resonator using CO$_2$ laser splicer (Fujikura LZM-100).

Author Contributions: Conceptualization, D.T.; methodology, all authors; software, D.T.; validation, all authors; formal analysis, D.T.; investigation, Z.A., A.B., Z.M., A.R., T.A., M.S.; data curation, D.T.; writing—original draft preparation, D.T., Z.A., A.B.; writing—review and editing, all authors; funding acquisition, D.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Nazarbayev University, under grants SMARTER (code: 091019CRP2117) and EPICGuide (code: 240919FD3908).

Institutional Review Board Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Socorro-Leránzoz, A.; Santano, D.; Del Villar, I.; Matías, I. Trends in the design of wavelength-based optical fibrebiosensors (2008–2018). Biosens. Bioelectron. X 2019, 1, 1000015. [CrossRef] [PubMed]
2. Guo, T.; González-Vila, A.; Loyez, M.; Caucheteur, C. Plasmonic Optical Fiber-Grating Immunosensing: A Review. Sensors 2017, 17, 2732. [CrossRef] [PubMed]
3. Loyez, M.; Lobry, M.; Hassan, E.M.; DeRosa, M.C.; Caucheteur, C.; Wattiez, R. HER2 breast cancer biomarker detection using a sandwich optical fiber assay. Talanta 2021, 221, 121452. [CrossRef]
4. Ran, Y.; Long, J.; Xu, Z.; Yin, Y.; Hu, D.; Long, X.; Zhang, Y.; Liang, L.; Liang, H.; Guan, B.-O. Harmonic optical microfiber Bragg grating immunosensor for the accelerative test of cardiac biomarker (cTn-I). Biosens. Bioelectron. 2021, 179, 113081. [CrossRef] [PubMed]
5. Guo, T.; Liu, F.; Liang, X.; Qiu, X.; Huang, Y.; Xie, C.; Xu, P.; Mao, W.; Guan, B.-O.; Albert, J. Highly sensitive detection of urinary protein variations using tilted fiber grating sensors with plasmonic nanocoatings. Biosens. Bioelectron. 2016, 78, 221–228. [CrossRef] [PubMed]
6. Gong, Z.; Xiang, Z.; Ouyang, X.; Zhang, J.; Lau, N.; Zhou, J.; Chan, C.C. Wearable Fiber optic Technology Based on Smart Textile: A Review. Materials 2019, 12, 3311. [CrossRef]
7. Chiavaioli, F.; Gouveia, C.A.J.; Jorge, P.A.S.; Baldini, F. Towards a Uniform Metrological Assessment of Grating-Based Optical Fiber Sensors: From Refractometers to Biosensors. Biosensors 2017, 7, 23. [CrossRef] [PubMed]
8. Agrawal, G.P. Fiber-Optic Communication Systems. In Phased Array Antennas; Wiley: Hoboken, NJ, USA, 2010; Volume 222, pp. 1–603.
9. Presti, D.L.; Massaroni, C.; Leitao, C.S.J.; Domingues, M.D.F.; Sypabekova, M.; Barrera, D.; Floris, I.; Massari, L.; Oddo, C.M.; Sales, S.; et al. Fiber Bragg Gratings for Medical Applications and Future Challenges: A Review. IEEE Access 2020, 8, 156863–156888. [CrossRef]
10. Bekmurzayeva, A.; Dukenbayev, K.; Shaimerdenova, M.; Bekniyazov, I.; Ayupova, T.; Sypabekova, M.; Molardi, C.; Tosi, D. Etched Fiber Bragg Grating Biosensor Functionalized with Aptamers for Detection of Thrombin. Sensors 2018, 18, 4298. [CrossRef]
11. Chiavaioli, F.; Baldini, F.; Tombelli, S.; Trono, C.; Giannetti, A. Biosensing with optical fiber gratings. Nanophotonics 2017, 6, 663–679. [CrossRef]
