Wireless Sensor Network Microcantilever Data Processing using Principal Component and Correlation Analysis

Viktor Zaharov[1], Angel Lambertt[2] and Ali Passian[3]
[1]*Polytechnic University of Puerto Rico, San Juan, PR 00918, U.S.A.*
[2]*Universidad Anahuac Norte, Huixquilucan, Edo. de Mexico, Mexico*
[3]*Oak Ridge National Laboratory, Oak Ridge, TN 37831, U.S.A.*

Keywords: Wireless Sensor Network, Microcantilever, Karhunen-Loève Transform, Correlation Analysis, Data Denoising.

Abstract: One of the main purpose of the wireless sensor network is an identification of unknown physical, chemical and biological agents in monitoring area. It requires the measurement of the microcantilever sensor resonance frequencies with high precision. However, resolving the weak spectral variations in dynamic response of materials that are either dominated or excited by stochastic processes remains a challenge. In this paper we present the analysis and experimental results of the resonant excitation of a microcantilever sensor system (MSS) by the ambient random fluctuations. In our analysis, the dynamic process is decomposed into the bases of orthogonal functions with random coefficients using principal component analysis (PCA) and Karhunen-Loève theorem to obtain pertinent frequency shifts and spectral peaks. We show that using the truncated Karhunen-Loève Transform helps significantly increase the resolution of resonance frequency peaks compared to those obtained with conventional Fourier Transform processing.

1 INTRODUCTION

The fast grow of wireless sensor networks permanently demands to apply some sophisticated techniques to increase the reliability of acquired data for various applications, like physical and environmental conditions monitoring. A wide range of applications for the study of physical, chemical and biological processes require sensitive standoff detection of chemical and biological agents (Parmeter, 2004; Bengtsson et al., 2006; Farahi et al., 2012; Dada and Bialkowski, 2011).

The more advanced sensing device whenever has been implemented, which can be used for detection of unknown agents, is the microcantilever sensor system (MSS) (Buchapudi et al., 2011; Wig et al., 2006). MSS acquires the data by reading a reflected laser spot from the microcantilever sharp nanoscale tip. The reflected beam undergos intensity and shape modifications before reaching the sensitive photodiode, which reads out the microcantilever motion. Further processing of the microcantilever output data allows to find the resonance frequencies helping to identify unknown physical, chemical, or biological agent (Van Neste et al., 2009; Measures, 1984).

Many remote sensor applications operate in atmospheric conditions, dense gases and fluids causing sensitivity difficulties due to gas kinetic and hydrodynamic dissipation and coupling. Depending upon the concentration of the sought analyte and random characteristics of both the environment and the measuring systems, the acquired data can contain a variety of stochastic components that in many applications dominate over the useful signal despite employment of phase sensitive detection. As a result, the measured data are complemented by the high level of random fluctuations that obstruct the systematic pattern (Labuda et al., 2012; Kawakatsu et al., 2002). As for any laser remote sensing measurement the nature of these fluctuations are uncorrelated or weakly correlated random noise, and weakly or strongly correlated instrumental errors (Measures, 1984). Cleaning up the systematic pattern is a big challenge and the solution can be found using quite sophisticated data denoising techniques. In the literature denoising techniques, as a rule, refer to manipulation of some orthogonal decompositions coefficients. The coefficients split out into two orthogonal subsets – the deterministic data pattern and the stochastic data; the first one includes the values which exceed a speci-

Zaharov, V., Lambertt, A. and Passian, A.
Wireless Sensor Network Microcantilever Data Processing using Principal Component and Correlation Analysis.
DOI: 10.5220/0005933200970105
In *Proceedings of the 13th International Joint Conference on e-Business and Telecommunications (ICETE 2016) - Volume 6: WINSYS*, pages 97-105
ISBN: 978-989-758-196-0
Copyright © 2016 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

fied threshold level, and the second one the values which do not exceed a threshold level. Afterward, the denoised data can be reconstructed using an inverse transform over the deterministic data pattern coefficients. To process the data from standoff experiments toward better recognition many well-known denoising techniques are presented in the literature, e.g, Fourier transform (FT), wavelet transform, Haar transform, and so on (Ahmed and Rao, 1975; Wang, 2012).

In this paper we employ more promising type of decomposition that are based on the principal component analysis (PCA) (Wang, 2012) and can be successfully used for the purpose of data denoising – Karhunen-Loève Transform (KLT). We analyze pros and cons of KLT, and discuss its discrete implementation. We show that denoising of data with KLT allows to increase the precision of resonance frequencies measurement because of the highest resolution ability of KLT over any known existing transforms. The simulation result confirms the high performance of KLT.

The paper is organized as follows. Section 2 discusses the microcantilever sensor system analysis and experimental setup, Section 3 introduces the KLT and its discrete implementation. Section 4 presents the application of KLT to sensitive cantilever experimental data processing and the result discussion. In Section 5 we discuss the threshold value determination between deterministic pattern and random fluctuations by involving the correlation analysis. We summarize our work in Section 6.

2 MICROCANTILEVER SENSOR SYSTEM ANALYSIS AND SETUP

Resonant microcantilever is a device that absorbs the particles and actuates them into vibration of amplitudes. The resonance cantilever frequencies are identified as peaks of maximal oscillation amplitudes in the frequency domain, and the resonance frequencies strongly depend on the nature of the particles. By measuring a shift in the resonance frequencies the unknown material can be detected and classified. The sensitivity of a cantilever is defined by the quality factor (Q-factor) that determines the resolution, and, as a result, the precision of resonance frequency shift measurement. The Q-factor of a cantilever is a specified value that depends on the cantilever geometry, material elasticity and mass. A change in mass due to interaction with the surrounding gases causes a shift in the resonance frequency of vibrating cantilever.

The higher the Q-factor, the higher the sensitivity of sensor and, as a result, the narrower the resonance peak bandwidth; hence, a shift in resonance frequency can be detected and estimated with high precision. However, despite the high Q-factor provides high sensitivity, the response of the sensor is rather slow. As shown in (Albrecht et al., 1991) for a cantilever with $Q = 50,000$ and a resonance frequency $f_r = 50$ kHz, the maximum available bandwidth is only 0.5 Hz, corresponding to the respond time $\tau = 2Q/2\pi f_r = 0.32$ s, which is too long for many applications. The dynamic range of high sensitivity sensor is also restricted due to high amplitude response on the resonance frequency. Because of mentioned constraints, using the cantilever with very high Q-factor in majority applications is undesirable.

Low Q-factor cantilevers operate with faster response, but because of their low peaks resolution the shifting in the resonance frequency can not be estimated precisely, especially when the shift is rather small. Hence, we have contradictory cantilever implementation requirements: it should operate with a fast response (requires low Q-factor), and in the same time it should be highly sensitive providing high resolution (requires high Q-factor) . Satisfaction to both conditions is a big challenge and an acceptable solution sometimes does not exist. Therefore, the goal of this paper is to achieve the high peak resolution spectra of low Q-factor cantilevers by using KLT.

In our test the microcantilever dynamics is monitored via optical beam deflection in atomic force microscopy (AFM) head. The signal of AFM, $S(t)$, is split and sent to four channels of a digitizing oscilloscope, where the four channels are captured in rapid succession, each channel measurement containing 10,000 points sampled at 200 ns intervals. Farther, we analyze the $4th$ channel data.

Let us consider $S(t)$ as the signal representing the relevant observable in the cantilever dynamics, that is, the deformation $u(x,t)$ at a given x. The vibrations of AFM cantilever, $u(x,t)$, can be described by a partial differential equation using the Euler Bernoulli beam theory

$$EI\frac{\partial^4 u(x,t)}{\partial x^4} + \rho A\frac{\partial^2 u(x,t)}{\partial^2} = 0. \qquad (1)$$

where E is the Youngs modulus, I is the second moment of inertia of the cross section, ρ is the mass density, and A is the cross sectional area (Measures, 1984).

In the absence of any external driving forces, $S(t)$ represents the equilibrium state of $u(x,t)$ and the accumulative random fluctuations in the entire system,

that is, the electronics noise, the Brownian oscillations of the cantilever, thermal, acoustic, mechanical noises, etc.

The solution of (1) has been found in (Passian et al., 2007; Lozano and Garcia, 2009) by the method of separation of variables, resulting

$$u(x,t) = \sum_{k=1}^{\infty} \phi_k(x) e^{jk\omega_k t}, \quad (2)$$

where $\phi_k(x)$ is a set of normalized orthogonal eigenfunctions.

Each term in (2) represents a kth vibration mode, whose dynamic equation is described by a second order differential equation with an effective spring constant μ_k, an effective mass m_k, a frequency ω_k, and a quality factor Q_k (Mokrane and et al, 2012)

$$m_k \frac{\partial^2 u_k(t)}{\partial t} + \frac{m_k \omega_k}{Q_k} \frac{\partial u_k(t)}{\partial t} + \mu_k u_k(t) = F_{noise}. \quad (3)$$

where F_{noise} the variable that represents all external and internal noises.

The fluctuation-dissipation theorem states that the power spectral density (PSD) of the thermal noise at the free end of the cantilever is expressed at a kth vibration mode as

$$\mathbb{S}(\omega) = \frac{2K_bT}{\mu_k \pi \omega/2} \frac{Q_k}{(1-(\omega/\omega_k)^2)^2 Q_k^2 + (\omega/\omega_k)^2}. \quad (4)$$

where K_b is the Boltzman constant and T is the temperature in degrees of Kelvin.

The resonant oscillations of the cantilever, $S(t)$, by stochastic excitation are shown in Fig. 1. The corresponding PSD obtained as $\mathbb{S}(\omega) = S(\omega)S(\omega)^*$ is depicted in Fig. 2, where $S(\omega)$ is FT of $S(t)$. $S(t)$ represents a signal of a free cantilever, that is, when the tip of the micro cantilever probe is outside the range of interfacial force fields, typically a few nanometers away from the surface.

Practically only few resonances can be directly observed during the random fluctuations; therefore, we consider in details only the first and the second resonance frequencies of the cantilever, i.e., $f_r^{(1)} = 10$ kHz, $f_r^{(2)} = 66.7$ kHz. Nevertheless, the data depicted on Fig. 1 allow to analyze much higher resonance frequencies because of the total observation time is $t_{tot} = 2$ ms and a sampling rate is 5 MHz ($\Delta t = 0.2$ μs). Obviously, the sampling time $\Delta t = 0.2$ μs yields an oversampling for the first and for the second resonances, but it is more suitable for the higher order resonances.

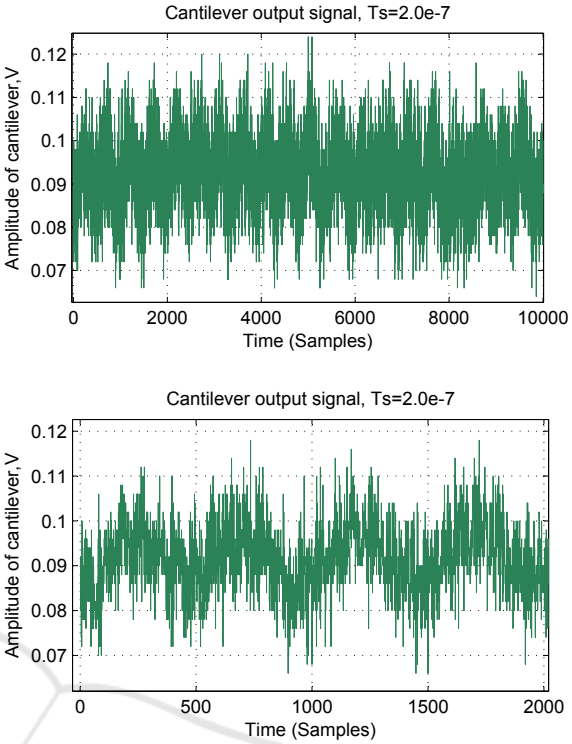

Figure 1: Equilibrium state microcantilever output data, sample time $2 \times 10^{-7} s$ a) 10,000 samples, b) 2,000 samples.

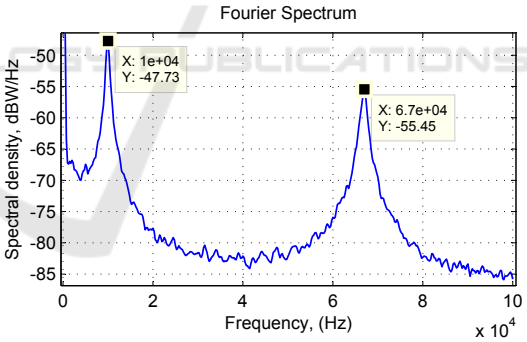

Figure 2: Fourier spectrum of the microcantilever dynamic response. The peak $f_r^{(1)} = 10$ KHz represents the first resonance frequency and at $f_r^{(2)} = 66.7$ KHz represents the second one.

3 KLT AND ITS DISCRETE IMPLEMENTATION

3.1 Karhunen-Loève Decomposition of Stochastic Process

One of the simplest and non-expensive orthogonal decomposition of the stochastic processes from the point

of view of computational complexity is the Fourier transform (FT) with the complex exponents orthogonal bases functions. Unfortunately, some series drawbacks of FT, like low resolution of the spectral components, slow convergence of the truncated decomposition to the origin data (especially when the data vector is short) encourage looking for some other orthogonal decompositions techniques with the lack of mentioned serious weaknesses.

One promising technique that helps to overcome many drawbacks of Fourier transform is the Karhunen-Loève transform (KLT). KLT is an orthogonal transform and it is optimum under the mean squared error (MSE) between truncated and the actual data providing the highest convergence of the data vector into smaller dimension subspace (Loève, 1978; Wang, 2012; Marple, 1987). In the opposite of Fourier transform, where the basis vectors are the complex exponent functions, the orthogonal basis of KLT are the eigenvectors of the data covariance matrix. It succeeds very attractive properties, because of complete decorrelation of the measured data helps to squeeze the data information into minimum number of parameters, and, ultimately, among other orthogonal transforms to reach the highest resolution of the data spectral components (Loève, 1978; Wang, 2012; Marple, 1987).

We propose exploiting the highest resolution property of KLT applying it for denoising of low Q-factor cantilever data, thereby significantly increasing the precision of the resonance frequency estimation. To introduce KLT we consider a decomposition of the stochastic data $S(t), t \in [t_1, t_2]$ as an infinite linear combination of orthogonal functions in $L^2[t_1, t_2]$

$$S(t) = \sum_{k=1}^{\infty} Z_k e_k(t), \tag{5}$$

where

$$Z_k = \int_{t_1}^{t_2} S(t) e_k(t) \tag{6}$$

are uncorrelated random variables and $e_k(t)$ is the set of continuous real-valued functions on $[t_1, t_2]$, which are pairwise orthogonal in $L^2[t_1, t_2]$.

Varying the set of orthogonal functions (6) bearing the transformations like Fourier transform, wavelet transform, Haar transform, and so on (Ahmed and Rao, 1975; Wang, 2012). Kari Karhunen and Michel Loève (Karhunen, 1947; Loève, 1978) had found the set of orthogonal functions that provide the optimum approximation of the original stochastic process in the sense of the minimum total mean-square error $[S(t) - \tilde{S}(t)]^2$, where $\tilde{S}(t) = \sum_{k=1}^{L} Z_k e_k(t)$ is a truncated to L terms decomposition of $S(t)$. The solution has been found by solving the homogeneous Fredholm integral equation of the second kind

$$\int_{t_1}^{t_2} K_S(\tau, t) e_k(\tau) d\tau = \lambda_k e_k(t). \tag{7}$$

where $K_S(\tau, t)$ the covariance function (kernel) that satisfies the definition of a Mercer theorem (Wang, 2012) postulating that such a set of eigenvalues λ_k and eigenfunctions $e_k(t)$ exists, which forms an orthonormal basis in $L^2[t_1, t_2]$, and $K_S(\tau, t)$ can be expressed as

$$K_S(\tau, t) = \sum_{k=1}^{\infty} \lambda_k e_k(\tau) e_k(t). \tag{8}$$

As a result, (6) becomes the Karhunen Loève Transform (KLT), and its basis provide the fastest convergence of $\tilde{S}(t)$ toward $S(t)$. However, the serious drawback of the KLT is the high numerical cost of determining the eigenvalues and eigenfunctions of its covariance operator, because it requires the solution of the integral equation (7). However, the implementation of KLT can be significantly simplified when input data is a discrete random process, because the linear algebra tools like the eigenvalues decomposition of discrete covariance matrix can be used. Furthermore, the fast KLT algorithms can help in some particular cases (Wang, 2012; Reed and Lan, 1994). Below we discuss the discrete KLT implementation in details.

3.2 KLT Discrete Implementation

Implementation of discrete KLT is interconnected with several inconveniences. Firstly, before KLT can be used for data processing, some preliminary data preprocessing should be done. It includes, a) storing in a memory the "training" set of representative data samples, b) forming the data covariance matrix, and c) finding the proper basis functions by computing the data covariance matrix eigenvalue decomposition. This preprocessing makes KLT a data dependent linear transform, and the proper basis functions are never known *a priori* excepting when the model of the data is known. Secondly, KLT has much higher computational cost and memory requirements that any other known orthogonal transform restricting both its software and hardware implementations. Nevertheless, KLT is a very attractive tool for data denoising because of data can be split into two orthogonal subspaces, the useful data and the noisy data, and afterward the noisy data can be just ignored. The data denoising algorithm that involves KLT includes the following sequence of operations (Wang, 2012).

1. Find the data covariance matrix $R = \frac{1}{M}\sum_{i=1}^{M} S_i S_i^T$, $M \geq N$, where S_i is a data $N \times 1$ vector, and M is a number of data vector running.

2. Find the eigenvalue decomposition $R = \Phi\Lambda\Phi^T$, where $\Lambda = [\lambda_1, \lambda_2, \ldots, \lambda_N]$ is a $N \times N$ diagonal matrix with elements sorted in descending order, and matrix $\Phi = [\phi_1, \phi_2, \ldots, \phi_N]$ is an eigenvectors matrix, where ϕ_i, $i = 1, \ldots, N$, is a $N \times 1$ vectors.
3. Determine the threshold that split the data space into two subspaces, the subspace of the useful eigenvalues, and subspace of the unwanted eigenvalues, and determine m significant eigenvalues which succeed the threshold ($m \leq N$).
4. Form $N \times m$ truncated KLT transform matrix $\Phi_m = [\phi_1, \phi_2, \ldots, \phi_m]$ that includes only eigenvectors belonging to larger m eigenvalues of R.
5. Compute direct KLT as $y = \Phi_m^T x$.
6. Compute inverse KLT as $x = \Phi_m y$ for reconstruction and data denoising .

Unfortunately, we are not able to practically implement step 1 because of lack of sufficient statistics – only one realization of the data vector is available. To form the full rank matrix R the number of available realizations must be at least equal to the number of data points, otherwise the matrix R is a deficient one. Furthermore, forming the full rank matrix R requires enormous numbers of measurements (at least 10,000 in our case), and posterior tough computations. However, the computational work can be shortened if the true covariance matrix R (it is asymptotical covariance matrix when $M \to \infty$) can be successfully replaced by the analytical matrix. Various types of random processes and their analytical covariance matrices have been derived and presented in (Maccone, 2009). It is well known that in the absence of any external driving forces, the microcantilever accumulative random fluctuations is the Brownian oscillations process (Mokrane and et al, 2012; Measures, 1984). Therefore, the analytical Brownian covariance matrix is formed according to (Maccone, 2009).

4 MICROCANTILEVER DATA PROCESSING WITH KLT

We find the KLT of the data depicted in Fig. 1 using the algorithm described in Section 3, where all 10,000 eigenvalues are involved. Fig. 3a shows resulting KLT spectrum for the frequency range 0 - 1 MHz (despite the whole available range of frequencies for the sample rate $T_s = 2 \times 10^{-7}$ s is 0 - 2.5 MHz), and Fig. 3b shows KTL spectrum in details in the frequency range 0 - 100 KHz just helping to compare the KLT spectrum with the reference FT spectrum.

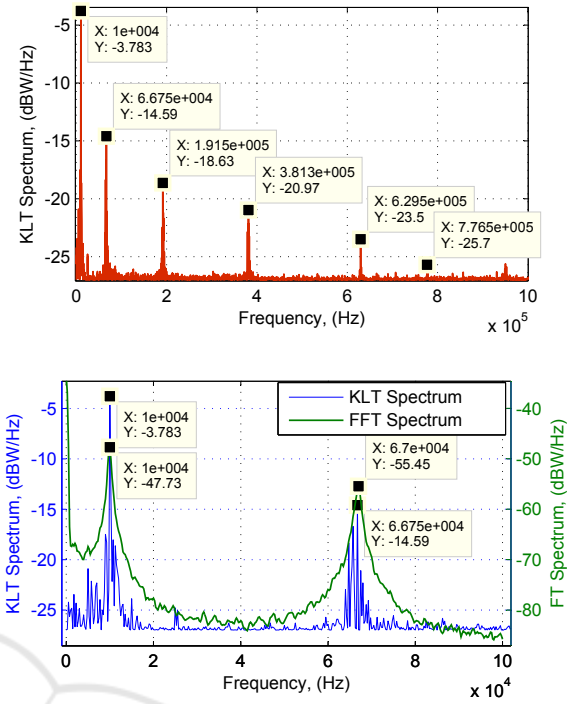

Figure 3: Cantilever output data Fourier spectrum vs. KLT spectrum a) $f_{max} = 1\ MHz$, b) $f_{max} = 100\ kHz$.

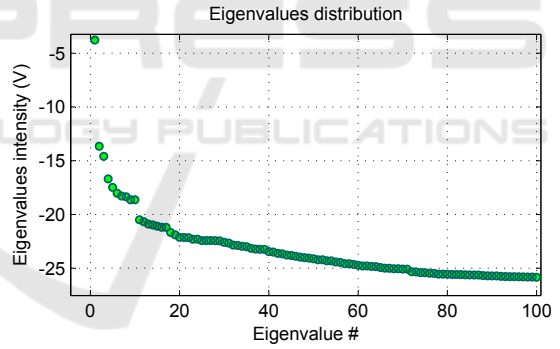

Figure 4: Distribution of the first 100 eigenvalues of data.

As follows from Fig. 3 the KLT spectrum has much higher resolution than the reference FT spectrum, even in the presence of complementary noise that can both distort and displace the result of resonance peak frequencies measurements. However, denoising of the original data helps to identify the resonance frequencies with more higher precision. We start the data denoising with finding of the data covariance matrix eigenvalues distribution, and the result is presented in Fig. 4. Applying the KLT to the output data and afterward reconstructing the data involving the very first largest eigenvalue resulting the perfect emerging of high resolved peak on the first resonance frequency 10 kHz. Then involving three largest eigenvalues results the second peak at the fre-

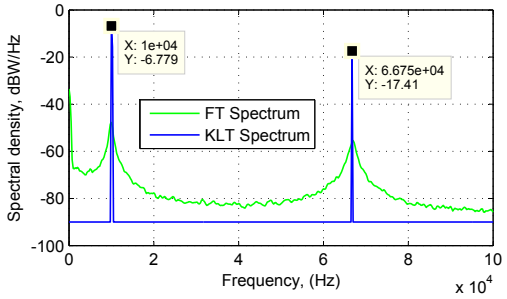

Figure 5: Estimation of the first and the second resonance frequency peaks involving three largest eigenvalues.

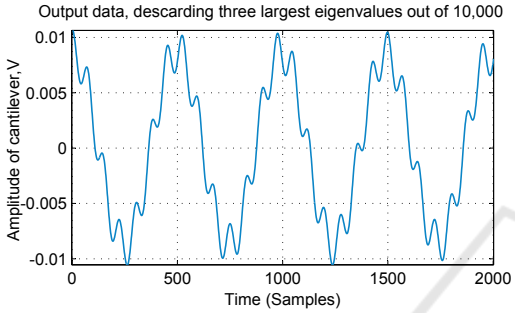

Figure 6: Data reconstructed using three largest eigenvalues.

quency 66.7 kHz as Fig. 5 shows. Emerging strong and narrow peaks helps to estimate resonance frequencies with much higher precision than those that can be provided by the reference spectrum obtained by FT. Fig. 6 depicts the time domain data when the first three eigenvalues are used in computing of the inverse KLT , where the first and the second resonance frequencies can be recognized as the lower and the higher frequencies oscillated waves, respectively.

5 THRESHOLD DETERMINATION BETWEEN USEFUL PATTERN AND NOISE USING CORRELATION ANALYSIS

In previous section to detect and identify the resonance frequencies we involve the fist three larger eigenvalues that correspond to the first and to the second resonances, e.i., 10.0 kHz and 66.7 kHz, and the rest of the eigenvalues were discarded as unwanted. However, as Fig. 3a shows, KLT spectrum allows to estimate the higher order resonance frequencies as well. Particularly, 191.5 kHz, 381.3 kHz, 629.5 kHz and 776.5 kHz, which can provide additional meaningful information about the measured data. It means that all eigenvalues that represent the higher order resonance frequencies belong to the set of eigenvalues that bearing additional useful information. Hence, discrimination of the desire eigenvalues, which represent the useful data and unwanted eigenvalues which represent the purely stochastic components, refers to the threshold determination problem. Finding appropriate threshold helps to solve the data denoising problem, i.e., to suppress the interfering random fluctuations that distort the resonance frequencies measurement result. The threshold determination problem is formulated as following (Marple, 1987). In the whole set of N eigenvalues $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m, \lambda_{m+1}, \ldots, \lambda_N\}$ obtained by eigenvalues decomposition of the data covariance matrix determine the value m, $1 < m < N$, that splits the set Λ into two orthogonal subspaces, one of them is a signal subspace, Λ_S, that contains the useful deterministic components (excited the resonances) and other one, Λ_n, is a subspace of the random noise that contains uncorrelated residual components (interferer), i.e.

$$R = \sum_{k=1}^{m} \lambda_k \phi_k \phi_k^H + \sum_{k=m+1}^{N} \lambda_k \phi_k \phi_k^H, \qquad (9)$$

where unknown value m must be determined. In the literature numerous attempts are done to determine the value m. For example, Konstantinides (Konstantinides and Yao, 1988) propose firstly to find the ratio $(\sum_{k=1}^{m} \lambda_k / \sum_{k=1}^{N} \lambda_k)^{1/2}$, and then compare it with some predetermined value ε, which depends on the variance of the random noise; however it require *a priory* information about the variance of the noise, hence, some additional measurements or analysis of data is needed. Marple (Marple, 1987) propose to use the rule based on the Akaike information criterion (AIC) (Sakamoto et al., 1986), that is

$$AIC\{m\} = (N-m)\ln\left(\frac{\frac{1}{N-m}\sum_{k=m+1}^{m}\lambda_k}{\prod_{k=m+1}^{m}\lambda_k^{N-m}}\right) + m(2N-m). \qquad (10)$$

Afterward, the minimum AIC over all m should be chosen. Despite the Akaike criteria estimates a relative information lost for a given data model, it does not provide the testing of a null hypothesis, meaning that some weak amplitude components of a useful signal can be lost.

Below we propose the solution of the threshold estimation problem by using the correlation analysis of the stochastic data subspace in order to find the subset that contains only the random uncorrelated components. Other words, we find the value m using the trial and errors approach ensuring that $\sum_{k=m+1}^{N} \lambda_k \phi_k \phi_k^H$ is a random uncorrelated data. For this, we successively

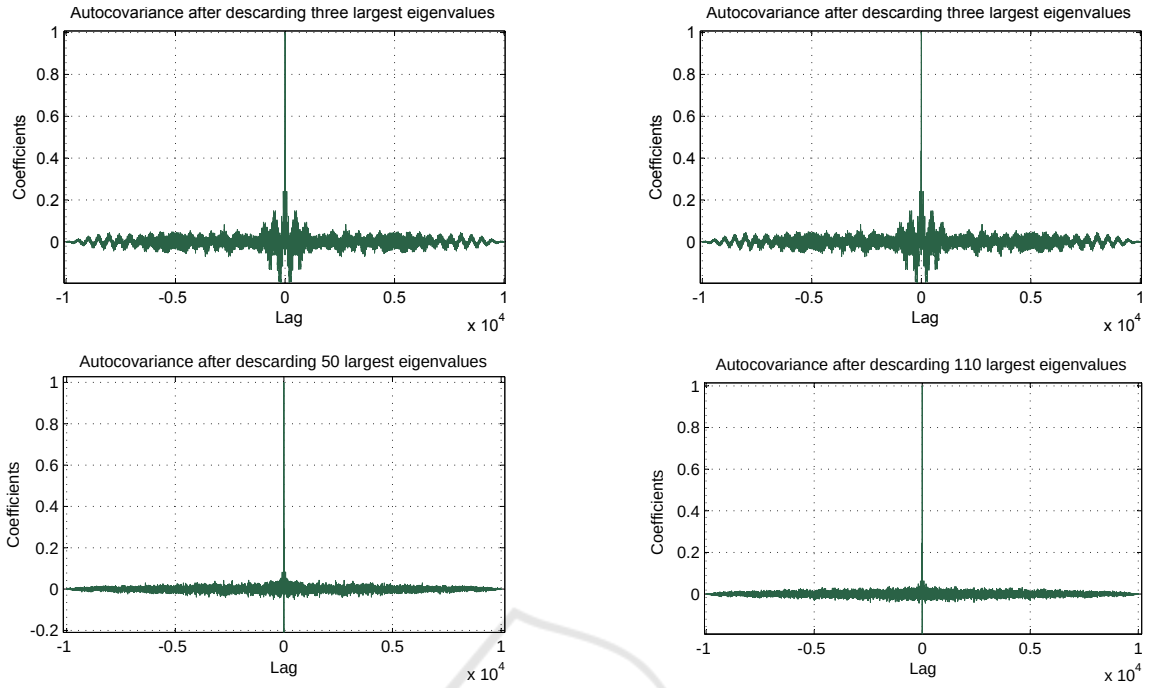

Figure 7: Data autocovariance after a) discarding three larger eigenvalues, b) discarding 20 larger eigenvalues, c) discarding 50 larger eigenvalues, d) discarding 110 larger eigenvalues.

discard $m = 1, 2, \ldots, N$ larger eigenvalues from the data and for each m we are testing the null hypothesis of the rest of the data. The resulting autocovariance function when m is 3, 20, 50 and 110 are presented in Fig. 7. As follows from Fig. 7, when only 3 or even 20 largest eigenvalues are discarded, the correlation is still high (lagged correlation coefficients oscillat with the high amplitudes), but when the number of largest eigenvalues to be discarded are increasing, then lagged correlation coefficients are getting lower approaching the level $-N \pm 2/\sqrt{N}$, and for $m = 110$ the null hypothesis (the sequence represents the random process with zero correlation) can be accepted inside the 95% confidence interval limits (Kendall et al., 1977). Resulting plots of splitting random noise and deterministic pattern when $m = 110$ in both time and frequency domains are depicted in Fig. 8, and Fig. 9, respectively. Visually Fig. 8 represents the poorly random data showing stochastic fluctuation in the time domain and almost equal amplitudes of all eigenvalues in the frequency domain (consist with PSD of the random noise), and Fig. 9 represents deterministic data, showing predicted behavior in the time domain and expressive peaks on the frequency domain (consist with deterministic pattern behavior). It just confirms that after the null hypothesis testing the determined threshold is found correctly.

6 CONCLUSIONS

We present and analyze microcantilever data denoising technique using one of the more advanced orthogonal transforms – Karhunen-Loève Transform (KLT), which helps to split the eigenvalue decomposition of measured data covariance matrix by two independent subsets – subset of useful signal that contains the larger eigenvalues, and subset of unwanted signal that contains the rest of eigenvalues. Processing of data with the first three larger eigenvalues and discarding the rest of them we perfectly determined the first and second resonance frequencies of the microcantilever with high precision.

We proposed to use the correlation analysis and the null hypothesis testing to determine the optimum threshold between subsets of deterministic eigenvalues and eigenvalues that belong to the random fluctuations. The simulation result confirm the truthful of proposed approach.

The KLT has been demonstrated the ability of effective improvement of the spectral identification of the micro cantilever resonances that were driven by the Brownian noise (thermal, mechanical, electronic and other type of random fluctuations). The simulation results and analytical analysis illustrate that KLT can be adapted as a powerful data denoising tool for the cantilever based sensing applications.

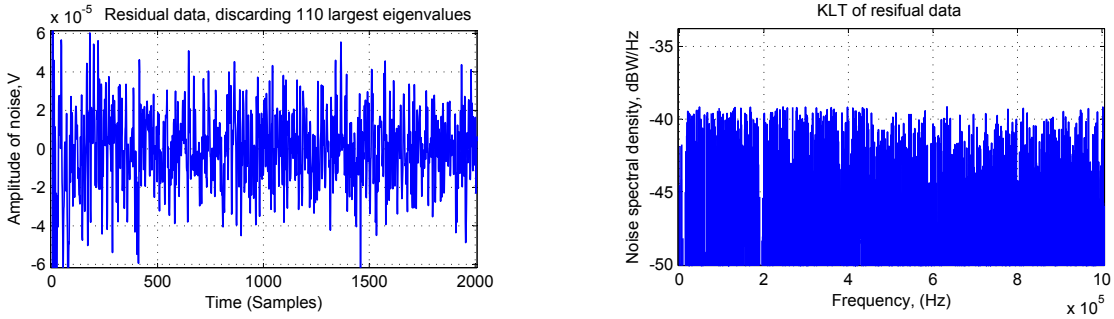

Figure 8: Residual noise after discarding of 110 larger eigenvalues a) time domain, b) frequency domain.

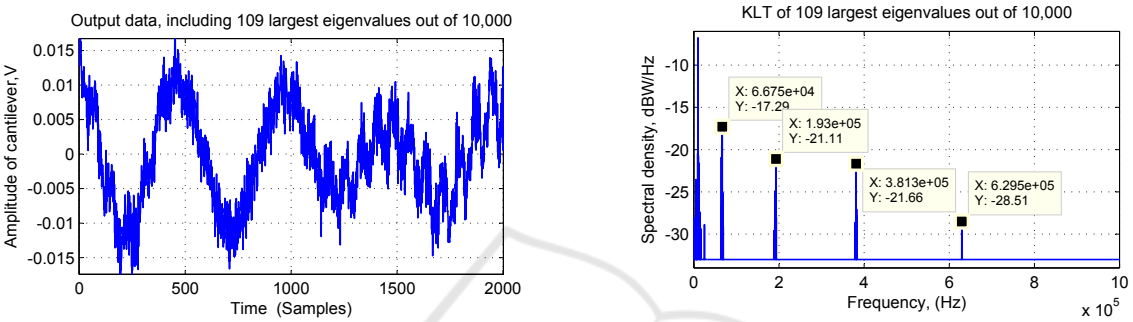

Figure 9: Deterministic pattern involving 109 larger eigenvalues a) time domain, b) frequency domain.

ACKNOWLEDGEMENTS

This work was supported by the laboratory directed research and development (LDRD) fund at Oak Ridge National Laboratory (ORNL). Viktor Zaharov acknowledges the financial support received from the Department of Energy (DOE) Visiting Faculty Program (VFP). ORNL is managed by UT-Battelle, LLC, for the US DOE under contract DE-AC05-00OR22725.

REFERENCES

Ahmed, N. and Rao, K. (1975). *Orthogonal transforms for digital signal processing*. Springer-Verlag.

Albrecht, T. R., Grtitter, P., Horne, D., and Rugar, D. (1991). Frequency modulation detection using highdkantilevers for enhanced force microscope sensitivity. *J. Appl. Phys.*, 69(2).

Bengtsson, M., Grönlund, R., Lundqvist, M., Larsson, A., Kröll, S., and Svanberg, S. (2006). Remote laser-induced breakdown spectroscopy for the detection and removal of salt on metal and polymeric surfaces. *Applied Spectroscopy*, 60:1188–1191.

Buchapudi, K. R., Huang, X., Yang, X., Ji, H. F., and Thundat, T. (2011). Microcantilever biosensors for chemicals and bioorganisms. *Analyst*, (136(8)):1539–1556.

Dada, O. O. and Bialkowski, S. E. (2011). A compact, pulsed infrared laser-excited photothermal deflection spectrometer. *Applied Spectroscopy*, 65(2):201–205.

Farahi, R. H., Passian, A., Jones, Y. K., Tetard, L., Lereu, A. L., and Thundat, T. G. (2012). Pumpprobe photothermal spectroscopy using quantum cascade lasers. *Journal of Physics D*, 45:125101.

Karhunen, K. (1947). Über lineare methoden in der wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae*, Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.(37):1–79.

Kawakatsu, H., Kawai, S., Saya, D., Nagashio, M., Kobayashi, D., Toshiyoshi, H., and Fujita, H. (2002). Towards atomic force microscopy up to 100 MHz. *Review of Scientific Instrument*, 73(2317).

Kendall, M. G., Stuart, A., and Ord, J. K. (1977). *The advanced theory of statistic*. Griffin.

Konstantinides, K. and Yao, K. (1988). Statistical analysis of effective singular values in matrix rank determination. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(5):757 – 763.

Labuda, A., Bates, J. R., and Grütter, P. H. (2012). The noise of coated cantilevers. *Nanotechnology*, 23(025503).

Loève, M. (1978). *Probability theory. Graduate Texts in Mathematics*, volume 2. Springer-Verlag, 4th edition.

Lozano, J. and Garcia, R. (2009). Theory of multifrequency atomic force microscopy. *Phys. Rev. B*, 79(014110).

Maccone, C. (2009). *Deep Space Flight And Communication*. Springer.

Marple, S. L. (1987). *Digital Spectral Analysis with Application*. Prentice-Hal.

Measures, R. M. (1984). *Laser Remote Sensing: Fundamentals and Applications*. Wiley-Interscience.

Mokrane, B. and et al (2012). Study of thermal and acoustic noise interferences in low stiffness atomic force microscope cantilevers and characterization of their dynamic properties. *Review of Scientific Instruments*, 83(1).

Parmeter, J. E., editor (2004). *The challenge of standoff explosives detection*. 0-7803-8506-3/02. IEEE.

Passian, A., Lereu, A. L., Yi, D., Barhen, S., and Thundat, T. (2007). Stochastic excitation and delayed oscillation of a micro-oscillator. *Physical Review B*, 75:233403.

Reed, I. S. and Lan, L.-S. (1994). A fast approximate karhunen-loève transform (aklt) for data compression. *Journal of Visual Communication and Image Representation*, 5:304–316.

Sakamoto, Y., Ishiguro, M., and Kitagawa, G. (1986). *Akaike Information Criterion Statistics*. Springer.

Van Neste, C. W., Senesac, L. R., and Thundat, T. (2009). Standoff spectroscopy of surface adsorbed chemicals. *Analytical Chemistry*, 81(5):1952–1956.

Wang, R. (2012). *Introduction to Orthogonal Transforms: With Applications in Data Processing and Analysis*. Cambridge University Press.

Wig, A., Arakawa, E. T., Passian, A., Ferrell, T. L., and Thundat, T. (2006). Photothermal spectroscopy of bacillus anthracis and bacillus cereus with microcantilevers. *Sensors and Actuators B*, 114:206.

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS