A Probabilistic Signal Representation for Detecting Faults in Highly Fragmented and Down-Sampled Vibration Signals

Dean Lee, Vincent Siu, Rick Cruz, Glenn Tolentino, Charles Yetman, and Phu Thoi
SPAWAR Systems Center Pacific
San Diego, CA, USA

***Abstract*—Classical vibration signals analysis techniques are developed assuming idealized conditions from which data is collected. In reality, due to physical constraints imposed by the operating environment, the vibration data is usually filtered by thresholding into segments which are then down-sampled, therefore not readily amenable to the application of classical techniques. In this paper, we introduce *probabilistically reconstructed signals* to represent this particular class of data, and show that the application of classification methods on the reconstructed signal is straightforward with reasonable accuracy.**

***Index Terms*—signal analysis; fault diagnosis; classification algorithms; probabilistic signals representation.**

I. Introduction

In most vibration signals analysis literature, analytic methods are developed based on idealized scenarios where the signals are recorded continuously at sufficient sampling rates. The analysis techniques are then expanded by tweaking the assumptions that underpin the theory, e.g. techniques are developed to account for over- or under-sampling of data. Fourier-transformed-based and Wavelet-based techniques are commonly used in these setup [1].

In reality, vibration signals collection mechanisms vary across different domains, and so the aforementioned analysis methods cannot be readily applied, due to possibly large gaps in data and extreme under-sampling of the data, or both. This type of data is especially prevalent in the avionics domain, where while there are vibration sensors on the aircrafts that collect data continuously, the onboard data bus is limited in bandwidth and the data storage device is limited in storage capacity. As a result the collected vibration data are usually highly fragmented due to thresholding, and the data is further processed by under-sampling to a predeteremined number of data points in order to save storage space.

There has been work done to address irregularly-sampled data [16], and separately, to address under-sampled data [9]. To the best of our knowledge, there has been no serious treatment of data in literature which are both irregularly-sampled and massively under-sampled. The techniques that we have come across from the field analysts in the avionics domain have all assumed the classical signals processing techniques, such as Fourier-Transform based analyses, are sufficient, but often end up with mixed results. We present the *probabilistically reconstructed signal*, or PRS, which is a probabilistic representation of the collected vibration signals, and show that this representation can be used in conjunction with machine learning techniques to detect faults from sub-optimally sampled, highly fragmented data.

II. Background and Related Work

The most basic technique to determine abnormal vibration is to establish a baseline using the Root Mean Square (RMS) or Crest Factor of the vibration levels of the operating environment; the system is considered to have a fault when the vibration exceeds some predetermined threshold.

Fourier-analytic based methods are also tools of choice when it comes to fault diagnosis. For example, time waveform analysis and frequency spectral analysis rely on extensive use of the Fourier transform to shuttle the signals between the time and frequency domains to detect faults; in particular, harmonic analysis is a cornerstone of these type of methods [1].

High frequency detection is yet another method that's used to detect faults in rotating machinery. The idea is that cracking or abrasive wear of the rotating element generates high stress waves, which can then be used as signatures to detect faults [2].

On the other hand, enveloping [3] is a sophisticated technique that was developed to uncover low frequency fault signals, which relies on determining the unique pass frequencies at which various faults can happen. Most enveloping methods seek to determine the optimal window to discern particular faults.

Wavelet-based methods are also popular. While they address many of the short-comings of Fourier-analytic methods, wavelet methods rely on finding the proper bases to be effective [4]. Still, there were some successes using wavelet-based methods combined with machine learning techniques [5].

The above techniques presume the signals are sampled sufficiently at evenly spaced intervals. With irregular sampling, the classical FFT techniques are modified into the Non-Uniform Discrete Fourier Transform (NDFT) [16] to handle non-uniformly sampled data. Additionally, there are a slew of other Non-NDFT methods developed for irregularly sampled data [6]–[8], which arises naturally in many situations, such as astronomy and geophysics. Moreover, the treatment of reasonably under-sampled data is well-understood [9].

ISBN: 1-60132-453-7, CSREA Press ©

Underlying all of the described techniques is the assumption that there are enough data points so as to recover some properties of the original signal. It is unclear how one should proceed if the collected data is both irregularly sampled (due to thresholding) followed by a down-sampling procedure from which the vast majority of the original signal is lost. In this paper we treat the data from a machine learning perspective. That is, every piece of collected signal fragments is a *sample* from the original signal. Ideally, with enough fragments, some properties of the original signal can be recovered via bootstrapping techniques.

III. Methodology

A. Suboptimally Collected Vibration Signals

Typically, vibration signals are modeled by

$$\eta(\tau) = \sigma(\tau) + \nu(\tau),$$

where σ is the unadulterated signal emitted by the sensors, and ν models noise. Given a measurement period m, the full vibration signals are represented by the sequence

$$(s_0, s_1, ..., s_i, ..., s_m),$$

where $s_i = \eta(\tau_i), 0 \leq i \leq m$.

In many real world scenarios, the limitations imposed by the operating environment, such as data storage capacity, prevent the full, continuous signals from being collected. Often times, the data is collected only when the signal value is above certain predefined threshold t, e.g. when a fault is thrown. Since the original signal η is recorded only when a threshold is crossed, the observed signal is a collection of fragmented sequences $\{f_0, f_1, ..., f_i, ...\}$, where f_i's are subsequences of the original signal. See Figure 1 for an example of how signal fragments are formed.

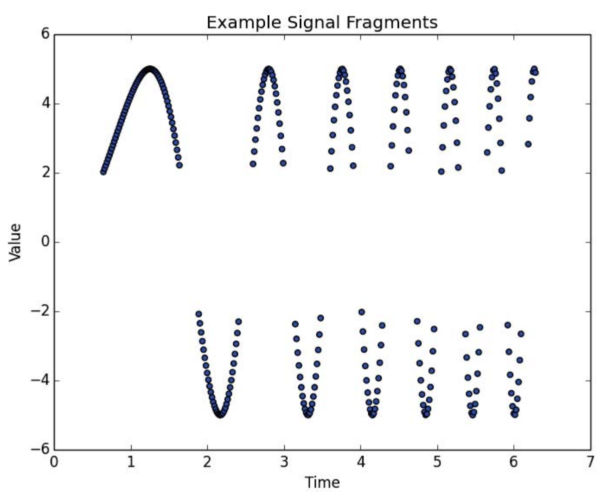

Fig. 1. Signal fragments as determined by threshold values

In austere data collection environments, such as helicopters, the data storage is so minimal that only certain number of data points d are stored after the threshold is crossed. In this case, the fragmented sequences f_i's are artificially down-sampled to exactly d evenly spaced data points. The recorded signals then become:

$$signal_{down-sample} = \{f_0, f_1, ..., f_i, ...\}, \qquad (1)$$

where length$(f_i) = d$ and for all element values s_i in f_i, $|s_i| \geq t$. In this paper, we consider only the type of signals described by Equation (1).

B. Probabilistically Reconstructed Signal

Clearly, the classical signals processing techniques cannot be applied as is to the thresholded and down-sampled data as described. However, with enough samples of the data, we may still recover enough information so that classifications can still be performed with reasonable accuracy.

To understand the problem, we examine (1) more closely. First, note that the length of the actual sampling period in (1) is different for every f_i due to the thresholding procedure. So while there are exactly d data points in every f_i, those data points may not be sampled at the same rate; more precisely, if the sampling period for f_i is w, then the sampling rate for f_i is exactly d/w. Secondly, the occurrence of f_i's are aperiodic; that is, the signals cross threshold t irregularly, possibly due to external forces, and therefore the time between successive f_i's is aperiodic. Without any prior information regarding the underlying signal generation processes, there may be too much information loss from the observed values for signals analysis using classical techniques.

Instead, we treat (1) as values sampled from some distributions $D_{\cdot,t}(\cdot)$ such that for all element values s in f_i, $s \sim D_{\cdot,t}(\cdot), |s| \geq t$. To achieve this, we first assign the time at which the values are sampled some ordinal values which are consistent across all f_i's in (1). We then take all f_i's and create a *probabilistically reconstructed signal* (PRS), such that for every ordinal time value o, there is a distribution $D_{o,t}(\cdot)$, so that for all elements s in f_o, $s \sim D_{o,t}$. More precisely, a PRS is

$$\text{PRS} = (s_{0,t}, s_{1,t}, ..., s_{i,t}, ..., s_{n,t}), \qquad (2)$$

where $|s_{i,t}| \geq t, s_{i,t} \sim D_{i,t}(\cdot)$, and n is the largest ordinal value.

C. Sampling Procedure

Informally, the PRS defines a distribution for every time point o in the aggregated signals space. The PRS then defines a probabilistic model, in which a signal can be sampled by bootstrapping at every time value. In this sense, the PRS is artificially augmenting the signals space in order to generate more samples to train the classifier.

The sampling procedure from the PRS is straight forward. For every ordinal value o on the time axis, a value is drawn from the associated distribution $D_{o,t}$. Since the signal fragments may be very sparse, the resulting PRS can be extremely sparse as well. For example, note that in Figure 2

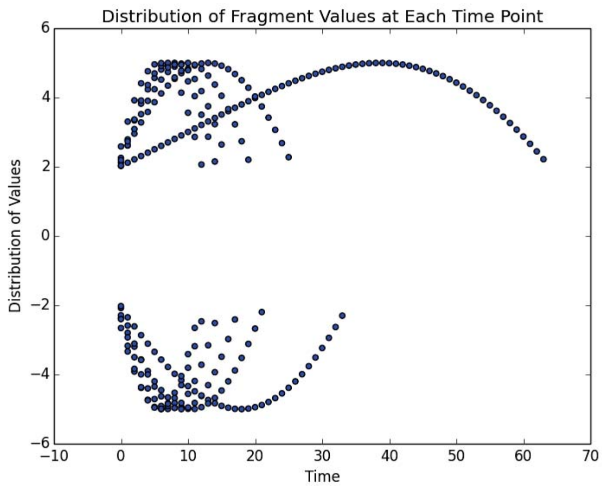

Fig. 2. The distribution of values at every time point, for every fragment in Fig. 1. Conceptually, the fragments are superimposed on top of each other, and this figure is the result. Now every time point becomes a discrete distribution from which to bootstrap.

the distribution of values becomes thinner as the ordinal time value increases. Realistically there may be many gaps where there simply arent any values for those times; in those cases the distributions is replaced with a constant value consistent with the data, such as the mean or median values.

IV. Experiments

We take 3 publicly available vibration data sets with fault diagnosis information, and simulate the conditions in which some proprietary data sets are collected and processed. First, we remove all data points with absolute values below some threshold t, so that the original signal is turned into a collection fragmented signals f_i, *à la* Equation (1). Second, for every f_i in the collection, the signals are further down-sampled so that exactly d evenly spaced data points remain in every fragment. Finally, if a segment does not contain d points after thresholding, it is dropped from the training data set.

In the following experiments, we vary threshold values t and length of signal fragments d to study the performance of PRS. For each data set, we take the down-sampled data and determine the classification accuracies with cross-validation as well as a held-out test data set. As a basis of comparison, we also determine the baseline classification performance without PRS. Random Forest classification is used due to its scale-invariant properties. In each of the experiments, we also note the percentage of the original signal that remains after the thresholding and down-sampling procedures, in order to show the degradation of the original signal.

A. Data Sets

The MFPT data set [10] is made up of three sets of bearing vibration data: 1) a baseline set, sampled at 97656 Hz for 6 seconds in each file; 2) an outer race faults set, sampled at 48828 Hz for 3 seconds in each file; and 3) an inner race faults set, sampled at 48828 Hz for 3 seconds in each file. The data points come from a single-channel radial accelerometer. There are additional data files included in the MFPT data set, but they are not used in the experiments. The data sets used in the experiment contains a total of 4541004 data points, before thresholding and down-sampling procedures.

The Bearing Fault data set [11] is made up of radial vibration measurements on a bearing housing of a test rig, sampled at 51200 Hz, and contains 2 sets of fault data, inner and outer race bearing faults respectively, each with 10 seconds worth of data. The data set contains a total of 1024000 data points.

The High Speed Gear Fault data set [12] is made up of radial vibration measurements taken from a 3MW wind turbine pinion gear, and contains a faulty data set and a baseline data set. The data files contain 6 seconds of measurements sampled at 97656 Hz. The data set contains a total of 14062464 data points.

B. Results

1) MFPT: See Figures 3 and 4 for classification results of the MFPT data set. See Table I for a perspective on how the original signal has degraded under thresholding and down-sampling. It is clear that the PRS method is able to capture the intrinsic properties of the original signal, since the classification accuracy is extremely high even when only 0.18% of the original data is present. Comparatively, the baseline classification performs quite poorly.

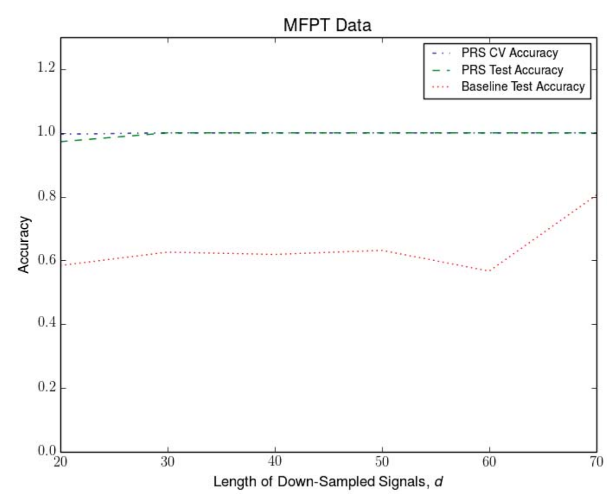

Fig. 3. Classification accuracy of MFPT Data by varying length of signal fragments d

2) Bearing Fault: See Figures 5 and 6 and Table II for results of the Bearing Fault data set. Overall, the PRS performs better than the baseline classifier when d is varied, although when d is small, the performance of PRS tends to be uneven. In particular, the classification of the test set experiences wild swings at smaller values of d. We also note that when at $d =$ 100 the dramatic drop of the baseline classifier, and the uptick

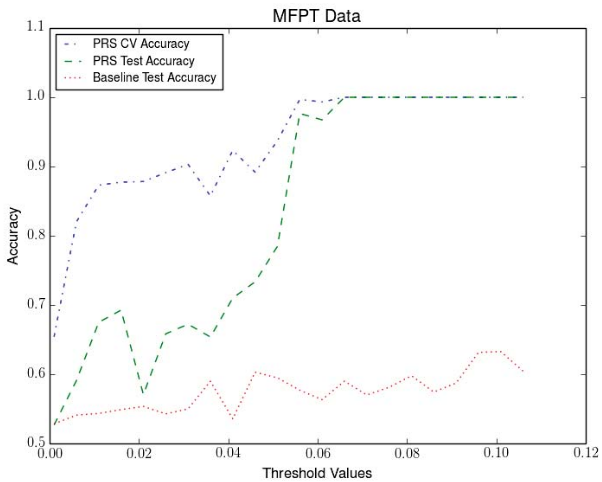

Fig. 4. Classification accuracy of MFPT Data by varying threshold values

Length of Down-Sampled Signals, d	Pct. of Original Data
20	3.45%
30	1.61%
40	0.82%
50	0.40%
60	0.24%
70	0.18%

TABLE I
PERCENTAGE OF RETAINED MFPT DATA, AS A RESULT OF THRESHOLDING AND DOWN-SAMPLING

in the classification accuracy of the PRS. On the other hand, as evidenced in Figure 6, the performance of both the PRS and baseline classifiers seem to have big variances as the threshold values are varied. We suspect this is due to the scarcity of the original data to begin with, at 1024000 data points this data set is the smallest of the three data sets under consideration. Additionally, with 0.83% of the original data (see Table II), the classifiers are working with just 8700 points. After some point, there just isn't enough data to work with, hence the poor performance of the classifiers.

3) High Speed Gear: See Figures 7 and 8 and Table III for results of the High Speed Gear data set. Overall, the PRS method outperforms the baseline data set, although the PRS test accuracy hovers slightly above the baseline accuracy for most parts of the experiment.

V. DISCUSSION AND FUTURE WORK

As seen in the Section IV-B, the performance of PRS can be uneven at times, especially when the length d of the signal fragment is relatively small. In most cases, however, the performance of the PRS is at least as good as the baseline tests. In general, we find that when d is bigger, the classifier performance becomes more stable. However, note that due to the way the data was prepared, a bigger value of d correlates with dropping more data (see Tables I, II, and III), so in essence, the classifiers are dealing with less data, but more closely related data points. We conjecture this is why the PRS

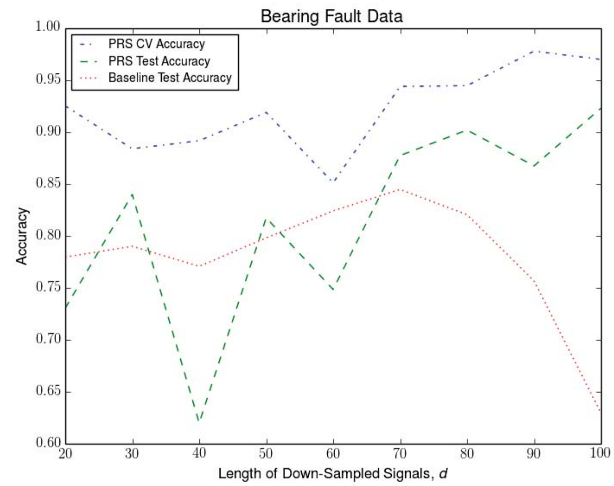

Fig. 5. Classification accuracy of Bearing Fault Data by varying the length of signal fragments d

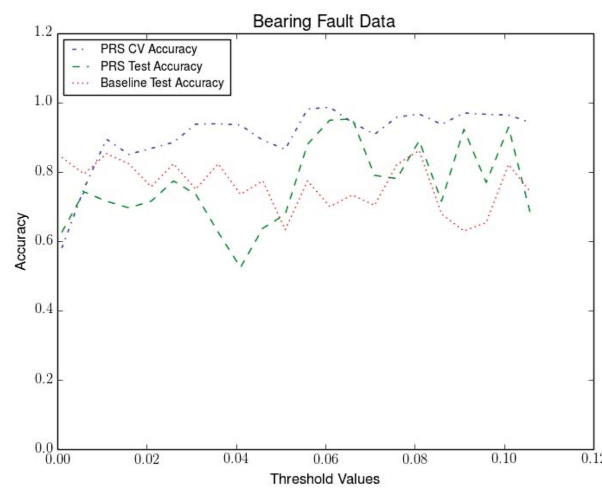

Fig. 6. Classification accuracy of Bearing Fault Data by varying threshold values

is able to extract more intrinsic properties from the degraded signals. It may be interesting to study the theoretical limit at which the performance of PRS starts to degrade. As seen in the experiments conducted on the Bearing Fault Data, since the size of the data set is relatively small to begin with, the PRS classifier performance tends to be uneven.

It is also common practice to standardize data sets before machine learning algorithms are applied. However, in the case of degraded signals, such as the ones we consider in this paper, we believe that standardizing the raw signal values (after the thresholding and down-sampling procedures) can actually be detrimental to the classifier performance. Some preliminary work (not shown in this paper) was done and showed that indeed the PRS does not work well when the data is standardized. Perhaps it is because the standardization

Length of Down-Sampled Signals, d	Pct. of Original Data
20	4.87%
30	3.34%
40	2.49%
50	1.76%
60	1.77%
70	1.31%
80	0.99%
90	1.08%
100	0.85%

TABLE II
PERCENTAGE OF RETAINED BEARING FAULT DATA, AS A RESULT OF THRESHOLDING AND DOWN-SAMPLING

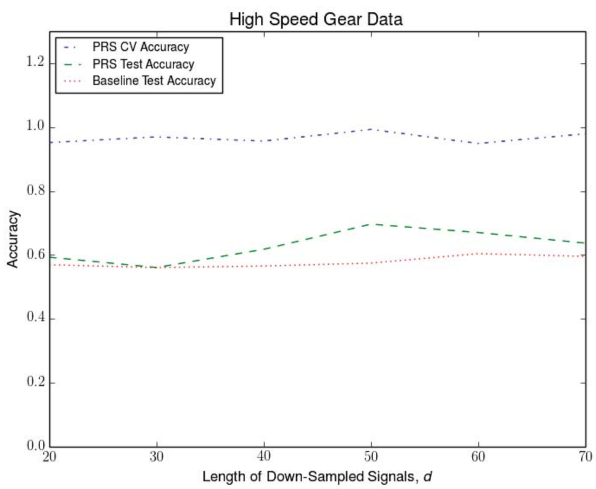

Fig. 7. Classification accuracy of High Speed Gear Data by varying length of signal fragments d

procedure introduces additional information loss on an already severely degraded signal. Thus, every bit of the degraded signal must be leveraged to its fullest in order to extract maximal information out of the signal. Hence we chose the Random Forest classifier as it is invariant to the scale of the data values.

The PRS method relies on bootstrapping the degraded signals in order to extricate information about the original signal. However, bootstrapping assumes the data values are inherently discrete; clearly vibration signals are continuous. In recent years, there has been advances in the deep learning community that leverages ideas from variational inference [13,14] to learn continuous distributions with much success. Briefly, the variational inference technique, together with the reparameterization trick [15], turns the computation of an intractable prior in an Bayesian framework into an optimization problem, therefore making the finding of the distribution of priors tractable and possible. We believe that variational inference can be readily applied to the PRS, whether to determine individual $D_{\cdot,t}$, or to determine the distribution of the signal as a whole so that dependencies can be explicitly modeled, by sampling from these continuous distributions it will further enhance predictive power of the PRS. Moreover,

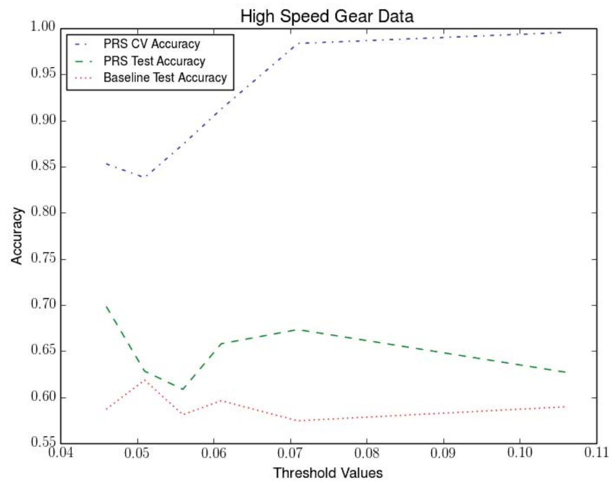

Fig. 8. Classification accuracy of High Speed Gear Data by varying threshold values

Length of Down-Sampled Signals, d	Pct. of Original Data
20	4.39%
30	2.49%
40	1.68%
50	1.07%
60	0.66%
70	0.43%

TABLE III
PERCENTAGE OF RETAINED HIGH SPEED GEAR DATA, AS A RESULT OF THRESHOLDING AND DOWN-SAMPLING

inherent in the sampling procedure stated in Section III-C is the simplifying assumption that every distribution $D_{\cdot,t}$ is independent of each other. With variational inference, we may be able to model much more complex and intricate interactions among all the distributions.

VI. CONCLUSION

We presented the *probabilistically reconstructed signal*, a technique which reconstructs a highly fragmented and down-sampled signal so that intrinsic properties of the original signal can be still be extracted. We took three publicly available vibration data sets, and curated the data by thresholding and down-sampling to match the conditions of the data collected in austere operating environments. We then showed the classification of faults using the PRS on severely degraded signals is still possible, and sometimes beats the baseline classifiers by big margins. While there are many possible improvements to be made to the PRS, such as sampling from continuous distributions instead of bootstrapping on discrete values to reconstruct the original signal, we believe what was demonstrated in this paper shows PRS is a viable approach to dealing with extremely degraded signals.

REFERENCES

[1] J. I. Taylor, *The Vibration Analysis Handbook*, 2nd ed. Vibration Consultants, 2003.
[2] S. J. Lacey, “An overview of bearing vibration analysis,” 2007.

[3] E. Mendel, T. W. Rauber, F. M. Varejão, and R. J. Batista, "Rolling element bearing fault diagnosis in rotating machines of oil extraction rigs," in *17th European Signal Processing Conference, EUSIPCO 2009, Glasgow, Scotland, UK, August 24-28, 2009*, 2009, pp. 1602–1606.

[4] L. Tóth and T. Tóth, "On finding better wavelet basis for bearing fault detection," *Acta Polytechnica Hungarica*, vol. 10, no. 3, pp. 17–35, 2013.

[5] P. K. Kankar, S. Sharma, and S. P. Harsha, "Rolling element bearing fault diagnosis using wavelet transform," *Elsevier - Neurocomputing*, vol. 74, pp. 1638–1645, 2011.

[6] H. G. Feichtinger and K. Gröchenig, "Irregular Sampling Theorems and Series Expansions of Band-Limited Functions," *Journal of Mathematical Analysis and Applications*, vol. 167, pp. 530–556, Jul. 1992.

[7] P. J. S. G. Ferreira, "Incomplete sampling series and the recovery of missing samples from oversampled bandlimited signals," *IEEE Transactions on Signal Processing*, vol. 40, pp. 225–227, Jan. 1992.

[8] P. J. S. G. Ferreira, A. M. P. Tome, A. M. P. Tom'e, and J. M. N. Vieira, "On two recent approaches to the reconstruction of signals from nonuniform samples," in *In Proceedings of SampTA'95, Jurmala*, 1995, pp. 263–267.

[9] Analog Devices, Inc., *Mixed-Signal and DSP Design Techniques*, Walt Kester, Ed. Newnes, 2003.

[10] http://www.mfpt.org/FaultData/FaultData.htm.

[11] http://data-acoustics.com/measurements/bearing-faults/bearing-1/.

[12] http://data-acoustics.com/measurements/gear-faults/gear-1/.

[13] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," *ArXiv e-prints*, Dec. 2013.

[14] L. Mescheder, S. Nowozin, and A. Geiger, "Adversarial Variational Bayes: Unifying Variational Autoencoders and Generative Adversarial Networks," *ArXiv e-prints*, Jan. 2017.

[15] L. Devroye, "Random variate generation in one line of code," in *Proceedings of the 28th Conference on Winter Simulation*, ser. WSC '96. Washington, DC, USA: IEEE Computer Society, 1996, pp. 265–272.

[16] F. Marvasti, *Nonuniform Sampling: Theory and Practice*. Plenum Publishers Co., 2001.

[17] S. Tao, T. Zhang, J. Yang, and X. Wang, "Bearing fault diagnosis method based on stacked autoencoder and softmax regression," in *Control Conference (CCC), Hangzhou, China, July 28-30, 2015*. IEEE, July 2015, pp. 6331–6335.

[18] W. Yan and L. Yu, "On accurate and reliable anomaly detection for gas turbine combustors: A deep learning approach," in *Annual Conference of The Prognostics and Health Management Society 2015*, vol. 6, 2015. [Online]. Available: http://www.phmsociety.org/node/1652/

[19] T. Junbo, L. Weining, A. Juneng, and W. Xueqian, "Fault diagnosis method study in roller bearing based on wavelet transform and stacked auto-encoder," in *Control and Decision Conference (CCDC), 2015 27th Chinese*. IEEE, 2015, pp. 4608–4613.

[20] Z. Chen, C. Li, and R.-V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock and Vibration*, vol. 2015, 2015.

[21] H. Balderston, "Incipient failure detection: Incipient failure detection in ball bearings." DTIC Document, Tech. Rep., 1969.

[22] J. L. Frarey, "The history and application of the envelope detector," DTIC Document, Tech. Rep., 1996.