# Compositional modelling of partial discharge pulse spectral characteristics

P. C. Baker, *Member, IEEE,* B. Stephen, *Member, IEEE,* and M. D. Judd, *Senior Member, IEEE*

*Abstract*—**Partial discharge (PD) monitoring is an established method for insulation health monitoring in high voltage plant. A number of different approaches to PD defect diagnosis have been developed to extract defect-specific information from PD pulse data in both the time and frequency domains. Frequency based PD pulse analysis has previously been demonstrated to offer a low-power approach to PD defect identification, where a mixture of passive and active analog electronics can be used to generate diagnostic features in a low-power device suited to wireless sensor network operation.**

**This paper examines approaches to implementing diagnostic methods for frequency-based PD pulse diagnosis targeted at compositional frequency spectrum features in a computationally efficient manner. Dirichlet and Gaussian distributions are used to demonstrate the complex probabilistic form of fault class decision surfaces, which motivates the proposed application of the log ratio transform to frequency composition data.**

**The results demonstrate that PD defects can be differentiated using these frequency-based methods and that employing the log ratio transform to the compositional frequency content data yields increases in classification accuracy without necessarily resorting to more complex classifiers.**

*Index Terms*—**Partial discharges, machine learning, condition monitoring, RF signals, fault diagnosis, wireless sensor networks**

## I. INTRODUCTION

**P**ARTIAL discharge (PD) monitoring is becoming an increasingly valuable tool for determining the state of electrical plant, which, due to regulatory and financial constraints on electricity utilities has become an area receiving significant research in recent years. Successful applications of PD monitoring include rotating machines [1], underground cables [2], gas insulated substations (GIS) [3], air-insulated substations [4] and power transformers [5, 6].

Partial discharges arise within electrical plant due to localised ionisation in regions where the electric field exceeds the dielectric strength of the insulation. PD can indicate the presence of faults long before equipment failure occurs. As an electrical plant asset management tool, PD monitoring provides a means of detecting, tracking and managing incipient insulation defects well in advance of their development into costly or unsafe failures. Partial discharges also contribute to the gradual degradation of insulation, with the end result being total, and potentially catastrophic, dielectric breakdown. To monitor PD in HV plant, a number of established approaches to PD detection have been developed including electrical, radiometric (RF) and acoustic techniques.

P. C. Baker, B. Stephen and M. D. Judd are with the Institute for Energy and Environment, University of Strathclyde, Glasgow, UK (email: m.judd@eee.strath.ac.uk)

CIGRE recommendations for power transformer condition monitoring suggest that the level of condition monitoring applied to plant should be determined by the strategic value of the asset [7]. The relative cost of on-line PD monitoring systems compared to asset value have resulted in their use being reserved for the most critical units. Hence, the development of a low-cost approach would allow PD monitoring to be economically viable on lower-valued assets where PD monitoring would not otherwise be considered. To support this, wireless sensor networks (WSNs) can provide the necessary processing, storage and communication capabilities for this to be realised. An RF PD pulse detector intended for wireless sensor network operation was previously developed to offer a low-cost approach to PD monitoring [8]. The device incorporated analog filters and low-power RF detectors in an integrated form factor, detecting PD pulse spectral energies across three bands between DC and 3.2 GHz. The device was designed to be incorporated into a wireless sensor that could feasibly be deployed onto a wide range of plant items such as oil-filled transformers and GIS.

To complement the low-power frequency-based detection method, a suitable diagnostic method is required that can distil the features generated by a physical detector into useful diagnostic information. The spectral energy content distributed across $D$ frequency bandwidths of a partial discharge pulse represent a composition of the overall spectral content. $D$-dimensional compositional data such as this can be represented as points in a $D$-simplex, where a Barycentric coordinate system allows for points to be plotted on a $D-1$ dimensional plane: a restricted space that accommodates the inherent dependence between compositional variables. In this paper, three frequency-based PD pulse diagnostic methods are presented. In the first instance, a simple discriminative classifier in the form of a binary decision tree is shown to automatically segment the simplex into regions that characterise the type of partial discharges. This concept is then demonstrated as a generative model over the simplex based around the Gaussian and Dirichlet distributions. Dirichlet distributions are useful in modelling the probability of proportional data; however, owing to the multiple causes of partial discharges it is necessary to incorporate more than one distribution to accommodate such complexities. Finally, the log ratio transform is used to transform data in the simplex back into Euclidian space, where linear classification methods can be applied and compared.

## II. MOTIVATION FOR WIRELESS PD SENSORS

For substation applications, wireless sensor networks can be used to instrument existing equipment without the need for

wired communications links, fulfilling the role of a condition monitoring network by combining remote sensing, data management and processing capabilities in a single architecture. There have been two recent notable surveys on the use of wireless technology for the electric power system. The first, published by EPRI [9], investigated wireless connectivity options for electric substations and made technology recommendations for specific application classes. The second, published by the US Department of Energy [10], took a wider view to look at wireless communications standards for use within the electricity industry as a whole. These strategic studies demonstrate that there is a an impetus for the use of wireless sensor networks for power system monitoring brought on by the requirement for increased observability of power system assets for the 'smart grid', the high capital cost of cabling within substations (upwards of $2000 USD per foot [11]), coupled with maturing technology and ratification of industrial wireless sensor network standards such as ISA100.11a and WirelessHART [12]. In practical terms, the largest substation wireless sensor network deployment to date consisted of over 250 wireless sensor nodes, measuring parameters including temperature, vibration and $SF_6$ insulating gas density [13]. Looking to the longer term, wireless condition monitoring sensors such as these are likely to become 'self-powering', drawing power from renewable sources or the ambient electromagnetic field found within substations [14].

## III. PARTIAL DISCHARGE MONITORING AND DIAGNOSTICS

### A. PD measurement

The physical nature of partial discharge has resulted in a number of approaches to PD measurement being developed. The commonly used techniques are:

*1) IEC 60270:* This is an electrical PD measurement method for high voltage equipment [15], where the 'apparent charge' equivalent of PD pulses are measured via capacitively coupled sensors attached to the terminals of a single phase of the plant. The IEC technique is best suited to off-line testing in a laboratory or during factory validation to verify the insulation condition before new plant is dispatched to customers, as the level of electrical noise commonly present within substations means that the technique does not perform well in the field (although this has been addressed in part through noise rejection methods [16]).

*2) Acoustic:* The acoustic PD detection method measures PD pulses in the ultrasonic range, usually between 20 kHz and 1 MHz. This method, first applied in [17] to transformers, is not subject to the electrical noise problems of the IEC method because the measurable ultrasonic wave emission is not affected by strong electric fields, however piezoelectric sensors can also be subject to large amounts of environmental noise within substations [5].

*3) RF:* The RF technique captures emissions from PD at frequencies up to the UHF band, resulting in less susceptibility to noise than the IEC and acoustic methods as electrical interference tends to be concentrated at lower frequencies. However, despite being more immune to noise than the IEC and acoustic methods, RF PD pulse measurements are still

affected by factors including signal reflection, refraction and attenuation, based upon the geometry and distance between PD source and sensor. RF transducers are increasingly being incorporated into new plant [7], although in some circumstances sensors can be retrofitted onto existing equipment. For example in high-voltage transformers either onto inspection hatches or through oil drain valves as shown in Fig. 1.

### B. PD pulse diagnostics

To complement PD measurement techniques, a range of PD defect diagnostic methods have been developed, largely drawing from signal processing and machine learning techniques with the aim of automating the otherwise labor-intensive process of analysing PD pulse measurements. These can be classified by their fundamental approach to extracting defect-specific information from PD pulses.

*1) Phase resolved diagnostics:* As PD activity is under the influence of the applied electric field which varies with the external voltage, PD pulses can be resolved against electrical phase for diagnostic purposes. Typically requiring high-speed sampling to capture PD pulses with sufficient resolution, phase-resolved partial discharge (PRPD) plots can reveal defect-specific information that can be used for classification and defect diagnosis. The RF method has been applied to on-line plant items including transformers [18], GIS [19] and cables [20], and automated phase-resolved diagnostic methods have been developed based on techniques including fuzzy classifiers [21], support vector machines [22] and knowledge-based systems [23].

*2) Pulse-sequence analysis:* The Pulse Sequence Analysis (PSA) method [24] analyses sequences of PD pulses in the time domain, rather than against an absolute phase reference. This incorporates the external voltage waveform magnitude as a relative measure of externally applied electric field strength. Automated analysis of pulse-sequence measurements has been achieved through the use of artificial neural networks (ANNs) [25].

*3) Frequency-domain diagnostics:* Frequency-domain diagnosis of PD pulses analyses the spectral content of PD pulses as a measure to identify and discriminate defect types. The frequency spectrum of different PD defect sources contains varying proportions of RF energy across different frequency bands, correlated to the geometry of the defect [26, 27], illustrated in Fig. 2. The frequency spectrum of an observed PD pulse is also dependent on the propagation path between



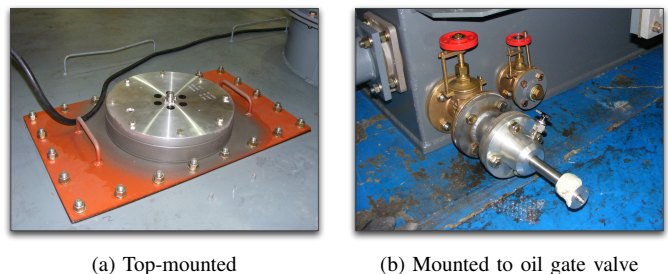(a) Top-mounted      (b) Mounted to oil gate valve

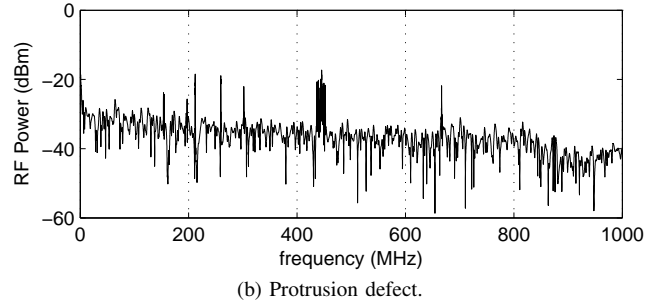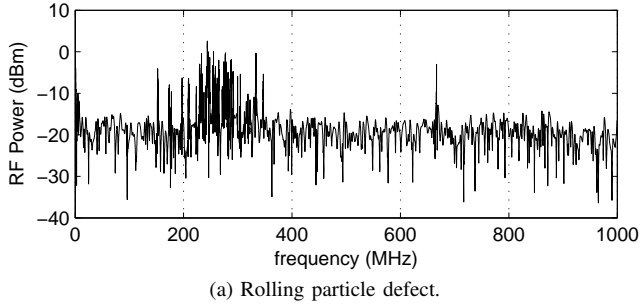Fig. 1. Examples of RF sensors mounted on power transformers

Fig. 2. PD pulse frequency spectra of two types of PD defect pulses generated under laboratory conditions. $f_s = 10 GSs^{-1}$, data length = 20002.

the defect source and sensor, incorporating factors including resonance and frequency-selective attenuation. The effect on observed PD pulse spectra will therefore differ between plant types as the complexity of the plant geometry varies, however defect-specific information is still present is the measured signal. Frequency-based analysis has previously been applied to a range of HV plant, including Gas Insulated Substations (GIS) [8, 27], XLPE cables [28], oil-filled transformers [29, 30] and induction motors [31]. In related work, a data-driven approach to PD frequency spectra feature reduction was achieved in [30]. This approach employed a wideband sampling method where principle component analysis and cluster analysis were applied for feature reduction. Euclidean and Mahalanobis distance measures were employed, both of which cannot accommodate for non-linearities in the data; an issue which is addressed in this paper.

## IV. FREQUENCY-BASED PD DETECTOR

The realisation of a diagnostic sensor based on the RF sensing of PD pulses was achieved in [8, 32] to detect PD pulses and perform rudimentary diagnostics based upon frequency content and pulse magnitude. The motivation for developing this device was to implement a low-cost, low-power package which could be incorporated into an industrial wireless sensor network and feasibly be powered by an energy harvesting device within a substation [14]. This detection method does not require an external voltage or timing reference and statistical feature generation is carried out by multiple analog detector channels.

The detector applies band filtering to incoming RF signals to evaluate the PD spectral energy over the three bands shown in Table I. Frequency bands were selected based upon previous PD pulse frequency content analysis carried out by Meijer [33, 34]. The PD detector was used in the laboratory to capture PD pulse data from a set of $SF_6$ gas-filled test cells, data which has also been used for the study presented in this paper. A full description of the detector and the experimental study can be found in [8, 35].

## V. COMPOSITIONAL DATA

One key aspect of the spectral information captured in the frequency-based method is that the data is compositional. In some multivariate data sets, including the PD pulse dataset used here, the information conveyed is from the values relative

TABLE I
FILTER BANDS EMPLOYED BY THE PD DETECTOR

| Band | Frequency (MHz) |
|------|-----------------|
| 1 | 0 - 450 |
| 2 | 400 - 750 |
| 3 | 700 - 3200 |

to each other rather than their explicit values or magnitudes. Compositional data is multivariate and independent save for a constant sum constraint (although variables in compositional data are deemed independent, there is an element of spurious correlation introduced through the common denominator [36]). Individual variates contain information that the whole does not convey. In reality only relative information is conveyed by each variable with the remainder being from consideration of multiple variables. The natural sample space for compositional data $x$ is the D-simplex $S^D$ which is a restricted part of Real space:

$$S^D = \left\{ x = [x_1, \ldots x_D] \in \Re^D \middle| x_i \geq 0, \sum_{i=1}^{D} x_i = \lambda \right\} \quad (1)$$

This constrains the variables to sum to a constant $\lambda$ (which is often unity) and enforces a constraint of strict positivity which reflects the two characterising attributes of compositional data. A 3-simplex can be represented in a ternary diagram which displays the points using a Barycentric co-ordinate system. In a 3-simplex, this is an equilateral triangle, in a 4-simplex it is a generalised tetrahedron. Although $D$-dimensional, the composition is noted by [37] to be completely specified by $D - 1$ variables with the absent component being inferred from the sum to unity.

### A. Example of compositional PD data

Using a data set of 1673 PD pulses generated under laboratory conditions from $SF_6$ test cells of 3 different defect PD [8], a composition over the frequency bands 0-450 MHz, 400-750 MHz and 700-3200 MHz was generated.

Using this set PD signals, labelled according to type, allows supervised learning techniques to be used to identify class-specific decision boundaries. For future observations, PD type can be inferred from the range of values each variable takes.
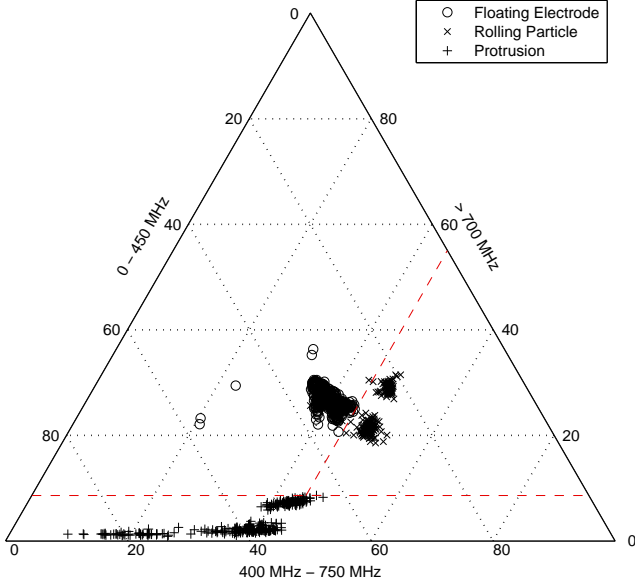
Fig. 3. 3-simplex ternary plot of PD pulse frequency content, split by regions generated from C4.5 indicated by dashed lines

One such technique is C4.5 [38] which uses the information content of each observation instance to partition the observations hierarchically, producing a decision tree which branches at particular values and terminates in leaf nodes when no ambiguity remains regarding the class to which the observation belongs. In [32], a decision tree was generated using C4.5 from the 1673 samples. This partitioned the simplex into regions that broadly represented the defect specific sub-compositions, as shown in Fig. 3.

The tree classifier has several limitations as the representation of the tree in Barycentric co-ordinates shows. The classifier decision boundaries are linear and restricted to being aligned with the axes. The Floating Electrode class contains almost equal energies over all three bands with relatively little variance in the composition. Despite this, the classifier extends the plausible region to include even zero energy in the 400 - 750 MHz band. Similarly, the Protrusion class is permitted to have a sub-450MHz band content ranging from 0 - 100 % even though it is never observed to have less than 50% of its spectral energy in this band. Also, in [32] it is noted that narrowband noise sources such as terrestrial TV transmissions, cellphone and wireless communications will typically exhibit a composition of 100% in the corresponding band. The presence of any of these interference types would produce spurious diagnostics from the decision tree classifier. What is lacking is a measure of plausibility for the regions where compositions have not been observed.

## VI. PROBABILISTIC CHARACTERISATION OF SIMPLICIAL PD DATA

The Dirichlet distribution for a $D$-dimensional composition is given by [37]:

$$P(X; \alpha) = \frac{\Gamma\left(\sum_{i=1}^{D} \alpha_i\right)}{\prod_{l=1}^{D} \Gamma(\alpha_l)} \prod_{j=1}^{D} x_j^{a_j} \qquad (2)$$

Where $\alpha$ is the parameter dictating the shape of the distribution and $\Gamma$ is the gamma function:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} \, dt \qquad (3)$$

The Gamma function effectively serves the same purpose for continuous data as the factorial does for integer data. The Dirichlet has a single hyperparameter $\alpha$, which is normally a $D$-dimensional vector although it can be a scalar in the case of a symmetric Dirichlet (where all $\alpha$ are equal). The Dirichlet can be reparameterised to provide a more intuitive view of how $\alpha$ values reflect the shape of the resulting distribution. By summing over all $\alpha$, the precision (the inverse of the variance) of the distribution can be obtained:

$$s = \sum_{d=1}^{D} \alpha_d \qquad (4)$$

In conjunction with the precision, the mean can be found as follows:

$$m = \left\{ \frac{\alpha_1}{s}, \frac{\alpha_2}{s}, \ldots, \frac{\alpha_D}{s} \right\} \qquad (5)$$

As a unimodal distribution, the mean defines the position where the data is mostly concentrated in the simplex while the precision shows the extent of its diffusion.

Learning the parameters of a Dirichlet distribution from a set of exemplar data requires an iterative approximation because the Maximum Likelihood solution does not exist in a closed form. The Newton Raphson iterative method can be used to approximate a maximum likelihood solution for the parameters of a Dirichlet distribution [39]; for approximating the parameter $\alpha$, a Newton Raphson step takes the general form:

$$\alpha_{t+1} = \alpha_t - H^{-1}(\alpha_t) g(\alpha_t) \qquad (6)$$

Where $H^{-1}$ is the inverse Hessian (matrix of second derivatives) and $g$ is the gradient vector for $\alpha$ at step $t$. Since the objective function (the Dirichlet distribution) is convex and unimodal, it can be lower bound and the bound optimised to the maximum point of the likelihood. The log probability of a Dirichlet distribution with hyperparameter $\alpha$ is:

$$\log P(X; \alpha) = N \log \Gamma \left( \sum_D \alpha_d \right) \qquad (7)$$
$$- N \sum_D \log \Gamma(\alpha_d)$$
$$+ N \sum_D (\alpha_d - 1) \log \bar{x}_d$$

The observed sufficient statistics are given by:

$$\bar{x}_k = \frac{1}{N} \sum_D \log x_{id} \tag{8}$$

The gradient is the first derivative of the log likelihood with respect to $\alpha$:

$$g = \frac{d \log P(X; \alpha)}{d\alpha} = \Psi \left( \sum_{i=1}^{K} \alpha \right) - \Psi(\alpha_k) + \bar{x}_k \tag{9}$$

Where $\Psi$ denotes the digamma function (the first derivative of the log gamma function). The need for an explicit inverted Hessian was bypassed in [39] with the following:

$$(H^{-1}g)_k = \frac{1}{q_{kk}} \left[ g_k - \frac{\sum \frac{g_j}{q_{jj}}}{\frac{1}{z} + \sum \frac{1}{q_{ij}}} \right] \tag{10}$$

where:

$$q_{jk} = -N\Psi'(\alpha_k)\delta(j - k) \tag{11}$$

and:

$$z = N\Psi' \left( \sum_{i=1}^{K} \alpha_i \right) \tag{12}$$

where $\Psi'$ is the trigamma function (the second derivative of the log gamma function). Using this approach, Dirichlet distributions were estimated from the same set of PD pulse spectral defect compositions used to generate the decision tree in the previous section.

The contour density plot in Fig. 4 shows how the distribution of the data is captured by class-conditional Dirichlet distributions. Although broadly correct, the inflexible covariance structure of the Dirichlet distribution inadequately represents the shape of the defect class envelope leading to a poor fit in some areas and overlap between two of the classes.

## VII. Log Ratio Transform

Although the Dirichlet distribution is intended for use with compositional data it is restricted in the shapes it can represent. To give an impression of the true form of the distribution of PD spectral energy in $S^D$, a closer approximation to the density function may be found by a Gaussian density estimate. Fig. 5 shows that the Gaussian distribution has a more flexible covariance structure but is not suited to use within the simplex owing to its linear distance assumptions. This limits the clusters to ellipses, whereas the underlying compositions that they represent are not necessarily elliptic. In Fig. 5, this also results in a cluster entering an undefined space outside of the simplex.

The statistics typically used in $R^D$ are geared around the absolute magnitudes of variables rather than the relative magnitudes in $S^D$ [36]. Although these statistical measures cannot be readily applied in $S^D$, [37] showed that using the log ratio transformation, points in $S^D$ could be mapped into $R^{D-1}$ space where they could then be used with statistics that employed Euclidean measures of distance. If $x$ is a point
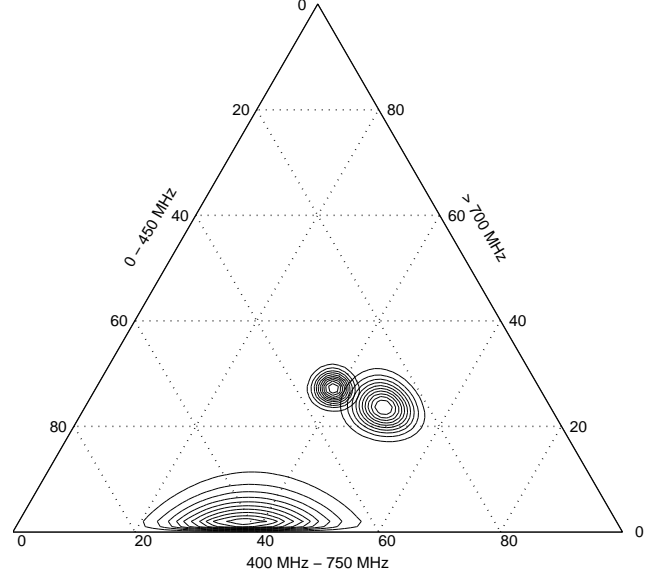


Fig. 4. Contour probability density plot of Dirichlet distribution learned in the simplex
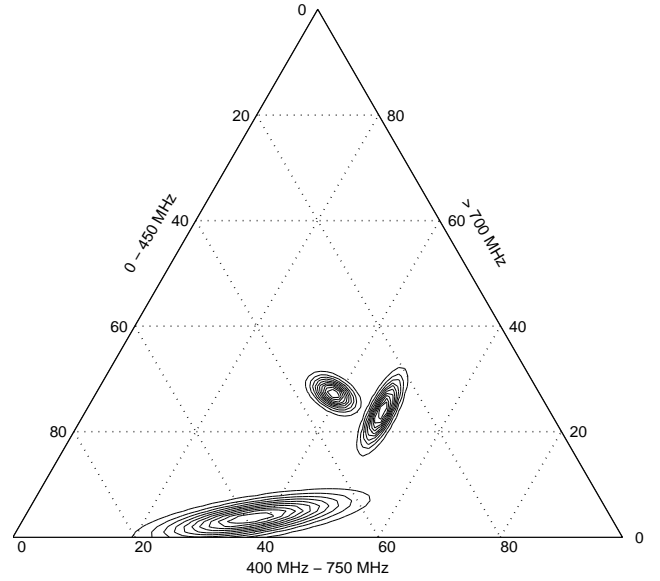


Fig. 5. Contour probability density plot of Gaussian distribution learned in the simplex. The largest cluster enters an undefined space outside the simplex

in $S^D$, to transform it into a variable $y$ in $R^{D-1}$ space, the following mapping is used:

$$y_i = \log \left( \frac{x_i}{x_D} \right) \quad i \neq D \tag{13}$$

The inverse transform can be used to map useful statistics (such as the location of a mean) back onto the simplex:

$$x_i = \frac{y_i}{\sum_{d=1}^{D-1} \exp \{y_d\} + 1} \tag{14}$$

Fig. 6 shows the PD pulse spectral data transformed into 2D log ratio space, where an increased separability between
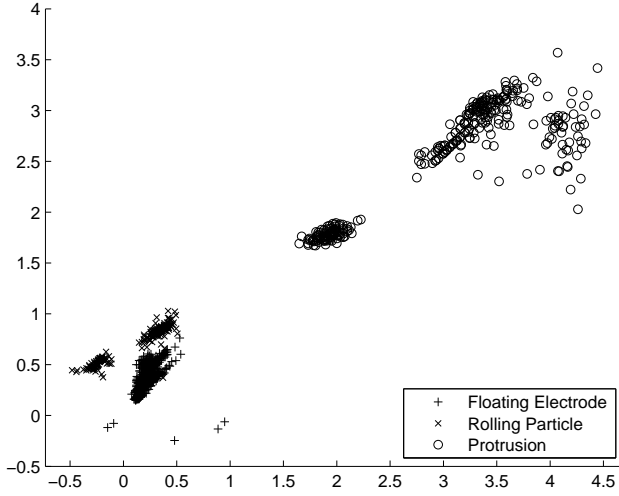
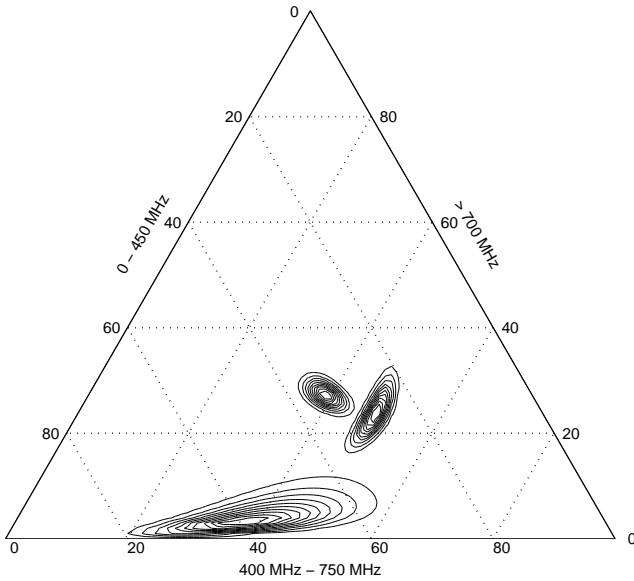Fig. 6.  Scatter plot of the 3-simplex transformed into 2D log ratio space



Fig. 7.  Log ratio Gaussian densities plotted in the 3-simplex. The density functions representing each defect class are capable of assuming more complex forms than either of the alternatives shown in Figs. 4 and 5

defect classes can be seen. Fig. 7 demonstrates how a Gaussian distribution can be applied in log ratio space and translated back to the 3-simplex for visualisation. This results in more complex and more accurate density functions which better match the shape of the clusters made up by individual data points shown in Fig. 3. It is evident that the densities in log ratio space provide a more accurate representation of the defect conditional probability mass than their real space equivalents.

## VIII. Linear Classification in Log Ratio Space

In order to quantify the benefit of using the Log ratio transform, one of the simplest classifiers in pattern recognition is employed: the General Linear Model (GLM) [35]; this affords a computationally simple diagnostic capability and one which can be implemented in a practical setting with minimal

TABLE II
ACCURACIES FOR CLASSIFICATION ON TRAINING AND TESTING SETS FOR
GENERAL LINEAR MODEL (GLM) IN SIMPLEX AND LOG RATIO SPACES

| Feature Space | Accuracy / % | |
|---|---|---|
| | Training | Testing |
| Simplex | 94.74 | 95.94 |
| Real (log ratio) | 99.20 | 99.28 |

TABLE III
TRAINING SET CONFUSION MATRIX FOR THE GENERAL LINEAR MODEL
IN THE SIMPLEX, FOR FLOATING ELECTRODE (FE), ROLLING PARTICLE
(RP) AND PROTRUSION (PRO) DEFECTS

| | | Classified as | | |
|---|---|---|---|---|
| | | FE | RP | PRO |
| Actual | FE | 683 | 66 | 0 |
| | RP | 0 | 179 | 0 |
| | PRO | 0 | 0 | 326 |

risk. The simple GLM is based on the assumption that classes are linearly separable with discrimination based on Euclidean distance from the decision boundary. The performance of this classifier is demonstrated on an exemplar PD set both in the 3-simplex and in the corresponding 2D log-ratio space.

A PD defect data set was randomly partitioned into a 1254 record training set and a held-out set of 419 records for validation purposes, both with 3 PD classes: 1) Floating Electrode; 2) Rolling Particle; and 3) Protrusion in $SF_6$.

Table II shows the accuracies for classification on the training and test sets for both spaces. The GLM training set confusion matrix, in Table III, shows that the GLM performs worst on training data in the simplex, misclassifying 66 rolling particle instances as floating electrode instances. The reason is the simple decision surface coupled with the inappropriate Euclidean assumption made on non-Euclidean problem space. Outperforming the Simplex-based training and testing sets, the log ratio transformed testing set results can be seen in Table IV. Significant gains in accuracy are achieved by using the GLM with the adoption of the log ratio transform. The explanation for this is transformation into Euclidean space, which augments linear separation.

## IX. Discussion

Analysing the PD composition further, there may be circumstances where a simplex is inhabited by multiple subpopulations which are distributed heterogeneously. This is illustrated

TABLE IV
TESTING SET CONFUSION MATRIX FOR THE GENERAL LINEAR MODEL IN
THE SIMPLEX, FOR FLOATING ELECTRODE (FE), ROLLING PARTICLE
(RP) AND PROTRUSION (PRO) DEFECTS

| | | Classified as | | |
|---|---|---|---|---|
| | | FE | RP | PRO |
| Actual | FE | 213 | 1 | 2 |
| | RP | 0 | 66 | 0 |
| | PRO | 0 | 0 | 137 |

in Fig. 6 where the Protrusion defect can be seen to have multiple distinct subpopulations which were the result of varying the orientation of the PD defect to identify the effect on the measured PD frequency spectra. The resulting sample space is non-stationary, but non-stationarity can be approximated in a piecewise manner using a linear combination of parametric densities by using a finite mixture model. In [40, 41] the Dirichlet distribution was used within a mixture modelling framework, allowing an arbitrarily complex distribution on the simplex to be represented. The finite mixture model assumes that the true data likelihood function can be expressed using a finite number of distributions of a known parametric form:

$$P(x) = \sum_{m=1}^{K} P(\theta_m)P(x;\theta_m) \tag{15}$$

Each distribution has its own parameter vector $\theta$ and is scaled by a mixing proportion $P(\theta_m)$ which is the probability of the observation being drawn from the associated distribution. The mixture parameters may be learned by Monte Carlo methods / stochastic approximation or by iterative ML approach such as Expectation Maximisation (EM) [42, 43]. In [44, 45], the authors note that the simple linear mixture will not capture hierarchical structure in the data. The possibility of a hierarchical Dirichlet model which embeds layers of covariance complexity to an arbitrary fit was examined in [46]. Considering the effects of defect geometry on observed frequency spectra, exploring this hierarchical model further has potential as a diagnostic tool as it may reveal multiple simultaneous defects of the same type, albeit in different locations.

The transforms described in this paper may be applicable to the analysis of dissolved gases in transformer oil. Duval's triangle [47], a commonly used dissolved gas analysis (DGA) technique also uses a 3-simplex for compositional analysis, relying upon linear classification boundaries. The results presented in this paper suggest that Duval's and other compositional data diagnostic methods may be developed by applying the techniques described.

## X. CONCLUSION

This paper has shown that using the pulse spectral method can reduce the PD monitoring problem to a frequency composition represented in a low dimensional simplex. The Dirichlet distribution, an established probability distribution over compositional data, has been demonstrated as lacking flexibility to adequately describe defect class conditional probabilities, while using Gaussian distributions in an inherently non-Euclidean feature space was demonstrated to be unsuitable. The approach presented here involved using the log ratio transform to map the PD frequency compositions from the 3-simplex into a 2D Euclidean space, allowing a computationally simple linear classifier to be employed. Higher rates of accuracy were demonstrated using a simple classifier, outperforming the same models on untransformed spaces.

Future work we intend to pursue in this area will be in the areas of more sophisticated PD modelling for enhanced decision support, proof of concept hardware implementations for field tests, and investigations into other applications where simplicial modelling and transformation may be applied to compositional data to simplify and increase the accuracy of simple diagnostic classifiers, such as Duval's method for DGA.

## REFERENCES

[1] J. Nelson and J. Stein, "A field assessment of PD and EMI methodology applied to large utility generators," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 17, no. 5, pp. 1411–1427, Oct. 2010.

[2] R.-N. Wu and C.-K. Chang, "The Use of Partial Discharges as an Online Monitoring System for Underground Cable Joints," *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1585–1591, Jul. 2011.

[3] W. Gao, D. Ding, and W. Liu, "Research on the Typical Partial Discharge Using the UHF Detection Method for GIS," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2621–2629, Oct. 2011.

[4] I. E. Portugues, P. J. Moore, I. A. Glover, C. Johnstone, R. H. McKosky, M. B. Goff, and L. van der Zel, "RF-Based Partial Discharge Early Warning System for Air-Insulated Substations," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 20–29, Jan. 2009.

[5] S. Markalous, S. Tenbohlen, and K. Feser, "Detection and location of partial discharges in power transformers using acoustic and electromagnetic signals," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 15, no. 6, pp. 1576–1583, Dec. 2008.

[6] M. D. Judd, "Experience with UHF partial discharge detection and location in power transformers," in *2011 Electrical Insulation Conference (EIC)*., IEEE, Jun. 2011, pp. 201–205.

[7] CIGRE Working Group A2.27, "Recommendations for condition monitoring and condition assessment facilities for transformers," *Electra*, vol. 237, pp. 48–57, Apr. 2008.

[8] P. C. Baker, M. D. Judd, and S. McArthur, "A frequency-based RF partial discharge detector for low-power wireless sensing," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 17, no. 1, pp. 133–140, Feb. 2010.

[9] EPRI, "Wireless Connectivity for Electric Substations," EPRI, Palo Alto, CA, Tech. Rep. 1016145, 2008.

[10] B. Akyol, H Kirkham, S. Clements, and M. Hadley, "A Survey of Wireless Communications for the Electric Power System," US Department of Energy, Tech. Rep. PNNL-19084, 2010.

[11] EPRI, "Automation in Power Plants and Wireless Technology Assessments," EPRI, Tech. Rep. 1010468, 2005.

[12] D. Chen, M. Nixon, and A. Mok, *WirelessHART - Real-Time Mesh Network for Industrial Automation*. New York, NY: Springer Verlag, 2010.

[13] A. Nasipuri, R. Cox, J. Conrad, L. Van der Zel, R. Zel, and M. Bienvenido Rodriguez, "Design considerations for a large-scale wireless sensor network for substation monitoring," in *5th IEEE International Workshop on Practical Issues in Building Sensor Network Applications*, Denver, Colorado: IEEE, 2010, pp. 866–873.

[14] M Zhu, P. C. Baker, N. M. Roscoe, M. D. Judd, and J Fitch, "Alternative Power Sources for Autonomous Sensors in High Voltage Plant," in *Proceedings of the 2009 IEEE Electrical Insulation Conference (EIC), Montreal Canada*, Jun. 2009, pp. 36–40.

[15] IEC, *IEC 60270: High-Voltage Test Techniques - Partial Discharge Measurements*. 2000.

[16] A Cavallini, A. Contin, G. Montanari, and F. Puletti, "Advanced PD inference in on-field measurements. I. Noise rejection," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 10, no. 2, pp. 216–224, 2003.

[17] E Howells and E. Norton, "Detection of partial discharges in transformers using acoustic emission techniques," *Power Apparatus and Systems, IEEE*, no. 5, pp. 1538–1549, 1978.

[18] M. Judd, L Yang, and I. Hunter, "Partial discharge monitoring of power transformers using UHF sensors. Part I: sensors and signal interpretation," *Electrical Insulation Magazine, IEEE*, vol. 21, no. 2, pp. 5–14, 2005.

[19] J. S. Pearson, O Farish, B. F. Hampton, M. D. Judd, D Templeton, B. M. Pryor, and I. M. Welsh, "Partial discharge diagnostics for gas insulated substations," *IEEE Transactions on Dielectrics and Electrical Insulation*, no. 5, pp. 893–905,

[20] N. H. Ahmed and N. N. Srinivas, "On-line Partial Discharge Detection in Cables," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 5, no. April, pp. 181–188, 1998.

[21] A Contin, A. Cavallini, G. Montanari, G. Pasini, and F. Puletti, "Digital detection and fuzzy classification of partial discharge signals," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 9, no. 3, pp. 335–348, 2002.

[22] L Hao, P. L. Lewin, and S. G. Swingler, "Identification of multiple partial discharge sources," in *Condition Monitoring and Diagnosis (CMD). International Conference on*, Apr. 2008, pp. 118–121.

[23] S. Strachan, S. Rudd, S. McArthur, M. Judd, S. Meijer, and E. Gulski, "Knowledge-based diagnosis of partial discharges in power transformers," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 15, no. 1, pp. 259–268, 2008.

[24] R. Patsch and F. Berton, "Pulse sequence analysis-a diagnostic tool based on the physics behind partial discharges," *Journal of Physics D: Applied Physics*, vol. 35, p. 25, 2002.

[25] M. Hoof, B. Freisleben, and R. Patsch, "PD source identification with novel discharge parameters using counterpropagation neural networks," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 4, no. 1, pp. 17–32, 1997.

[26] A. Reid, M. Judd, B. Stewart, and R. Fouracre, "Frequency distribution of RF energy from PD sources and its application in combined RF and IEC60270 measurements," in *Electrical Insulation and Dielectric Phenomena, 2006 IEEE Conference on*, IEEE, 2007, pp. 640–643.

[27] S. Meijer, "Partial discharge diagnosis of high-voltage gas-insulated systems," PhD thesis, Technische Universiteit Delft, Nov. 2001.

[28] Y. H. Md Thayoob, P. Ghosh, and A. B. A. Ghani, "Frequency spectral analysis of electrical partial discharge signals in XLPE cable under various soil conditions," in *International Power and Energy Conference*, IEEE, Dec. 2008, pp. 1528–1531.

[29] T Ishikura, H Muto, K Wada, and N Hosokawa, "Distinction of the partial discharge source in oil by two frequency correlation method," in *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, vol. 2, Beijing, 2008, pp. 42 –45.

[30] T. Babnik, R. K. Aggarwal, S. Member, and P. J. Moore, "Principal Component and Hierarchical Cluster Analyses as Applied to Transformer Partial Discharge Data With Particular Reference to Transformer Condition Monitoring," *Power Delivery, IEEE Transactions on*, vol. 23, no. 4, pp. 2008–2016, Oct. 2008.

[31] Q. Yu, A. Cavallini, and G. Montanari, "Frequency and Time-Domain Analysis of Partial Discharge Measurements in PWM Inverter-fed Induction Motors," in *Power Electronics and Motion Control Conference, 2004. IPEMC 2004. The 4th International*, vol. 2, Xi'an: IEEE, 2005, pp. 661–663.

[32] P. C. Baker, B Stephen, M. D. Judd, and S. D. J. McArthur, "Development of an Integrated Low-Power RF Partial Discharge Detector," in *Electrical Insulation Conference, 2009. EIC 2009. IEEE*, May 2009, pp. 273–277.

[33] S. Meijer, E. Gulski, J. Smit, F. J. Wester, T. Grun, and M. Turner, "Interpretation of PD in GIS using spectral analysis," in *High Voltage Engineering, 1999*, 1999, pp. 124–127.

[34] S. Meijer, "Partial Discharge Diagnosis of High-Voltage Gas-Insulated Systems," en, PhD thesis, TU Delft, Nov. 2001.

[35] P. C. Baker, "Enhancing substation condition monitoring through integrated diagnostics, wireless sensor networks and multi-agent systems," PhD thesis, University of Strathclyde, 2010.

[36] J Aitchison, "The Statistical Analysis of Compositional Data.," in *Monographs on statistics and applied probability*. London, UK: Chapman and Hall, 1986, p. 416.

[37] ——, "The Statistical Analysis of Compositional Data," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 44, no. 2, pp. 139–117, 1982.

[38] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA: Morgan Kaufmann, 1993.

[39] T. Minka, "The Dirichlet-tree distribution," MIT Media Lab Technical Note, Tech. Rep., 2004, pp. 1–6.

[40] K Sjölander, K Karplus, M Brown, R Hughey, a Krogh, I. S. Mian, and D Haussler, "Dirichlet mixtures: a method for improved detection of weak but significant protein sequence homology.," *Computer applications in the biosciences : CABIOS*, vol. 12, no. 4, pp. 327–45, Aug. 1996.

[41] N. Bouguila and D. Ziou, "Unsupervised selection of a finite Dirichlet mixture model: an MML-based approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 8, pp. 993–1009, Aug. 2006.

[42] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[43] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, 1997.

[44] D. MacKay and L. Peto, "A Hierarchical Dirichlet Language Model," *Natural language engineering*, vol. 1, no. 03, pp. 289–308, 1995.

[45] R. M. Neal, "Density Modeling and Clustering Using Dirichlet Diffusion Trees," *Bayesian Statistics 7*, pp. 619–629, 2003.

[46] S. Y. Dennis, "A Bayesian analysis of tree-structured statistical decision problems," *Journal of Statistical Planning and Inference*, vol. 53, no. 3, pp. 323–344, Aug. 1996.

[47] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE Electrical Insulation Magazine*, vol. 18, no. 3, pp. 8–17, 2002.

**Dr Pete Baker** (M'06) is a Research Associate within the Institute for Energy & Environment at the University of Strathclyde in Glasgow, Scotland, UK. He received his BEng (Hons) Degree from the University of Strathclyde in 2004, and his PhD from the University of Strathclyde in 2010. His research interests include substation condition monitoring, wireless sensor networks, energy harvesting, machine learning and multi-agent systems.

**Dr Bruce Stephen** (M'09) currently holds the post of Senior Research Fellow within the Institute for Energy and Environment at the University of Strathclyde. He received his B.Sc. from Glasgow University and M.Sc. and PhD degrees from the University of Strathclyde and is a Chartered Engineer. His research interests include Distributed Information Systems, Machine Learning applications in Power System and Animal Welfare Condition Monitoring / Asset Management.

**Dr Martin Judd** (M'02, SM'04) was born in Salford, England, in 1963. He graduated from the University of Hull in 1985 with a BSc (Hons) degree in Electronic Engineering. His subsequent industrial experience included 4 years working for Marconi Electronic Devices Ltd followed by 4 years with EEV Ltd, both in Lincoln, England. Martin received his PhD from the University of Strathclyde in 1996 for research into the excitation of UHF signals by partial discharges in gas insulated switchgear. From 1999 to 2004 he held an EPSRC Advanced Research Fellowship concerned with electrodynamics of electrical discharge phenomena. His fields of interest include high frequency electromagnetics, generation and measurement of fast transients, partial discharges and energy harvesting. Dr Judd holds the post of Reader in the Institute for Energy and Environment at the University of Strathclyde. He is a Chartered Engineer, a Member of the IET and Senior Member of the IEEE.