

Fault Testing and Diagnosis in Analog Circuits Using Fault Dictionary Techniques

Mohammad Yoosof^{1*}

Abstract

The present study aims to test and diagnose fault in analog-based equipment including biological and medical laboratory equipment. The diagnosis has been performed using fault dictionary techniques by modern classification methods. The modern classification methods include probabilistic neural networks and wavelets. It is known that analog circuits have difficulties compared to the digital circuits due to lack of efficient fault modeling, tolerance, nonlinearity, parametric effects and other effects. In this study, the fault diagnosis by simulation before-test approach in order to obtain a very high diagnostic degree was performed.

Keywords: Analog-based equipment; Biological laboratories; Medical laboratories; Fault diagnosis; Fault dictionary techniques; Probabilistic neural network; Wavelets.

Introduction

Since the middle of the 1970s, analog fault diagnosis attracted much attention from test engineers and academic researchers. A detailed summary of the research efforts for the pre-1990 period can be seen in [1,2]. With today's rapid development of analog very large-scale integrated circuits and mixed-signal systems, analog fault diagnosis gains more interests. This research interests are fully justified since analog testing lags far

behind the mature digital testing techniques due to its inherited bottlenecks [3-5].

A fault means a physical or intellectual imperfection or impairment reaching some standard of perfection in disposition, action, or habit. In a circuit, it is said to be any change in the value of an element's parameter with respect to its nominal value. It can cause the whole circuit performance to decrease or totally fail. Testing is to apply a procedure or

FIEAust, IEng, BEngHons, MBA, IQA, Doha, Qatar

*Corresponding Author: Mohammad Yoosof, Former Assistant Professor in Electronics Engineering, FIEAust, IEng, BEngHons, MBA, IQA, Doha, Qatar.

Received Date: 05-31-2022

Accepted Date: 06-24-2022

Published Date: 08-09-2022

Copyright© 2022 by Yoosof M. All rights reserved. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

reaction used to identify or characterize a constituent as a means of analysis or diagnosis. In a circuit, it is said to be a reaction to identify performance of the circuit under test (CUT). Diagnosis is to analyze or investigate the cause or nature of a condition or problem. In a circuit, it is said to analyze the parameter(s) causing the fault after a test [6].

Faults are categorized into catastrophic (hard) faults and Parametric (deviation or soft) faults. The catastrophic fault exists if the faulty element produces either short or open circuit. The parametric fault exists if the faulty element parameter value deviates referred to its nominal value -outside its tolerance box- without reaching extreme bounds. The parametric faults exist mainly due to intrinsic nature of the production process of the faulty element and on-the-field deviations such as those related to aging, parasitic, thermal effects, and other physical/chemical processes. The catastrophic faults can occur mainly due to parametric effects and sudden impacts leading fault element parameter to diverge outside extreme bounds. Practically, the catastrophic faults are recovered by around 70% of actual faults.

Faults can be divided according to number of faults in the whole circuit as single fault or multiple faults. In single fault, just one parameter causes the fault. In multiple faults, more than one parameter causes the fault(s) simultaneously. Practically, the single fault mostly occurs in catastrophic types and many techniques address this case. Multiple faults mostly occur in the parametric types and modern techniques are used to address this case.

Testing is categorized according to its objective as specification tests or production tests. The Specification tests can either be designed categorization specification (which is based on specifying if the design meets standards), or diagnostic specification (which is based on finding the cause of a fault). The Production tests are based on testing large numbers of linear/mixed-signal circuits used in production [7].

The main problems in analog testing can be divided into four stages: Fault detection, Fault location and identification, Fault evaluation and fault prediction. The Fault detection is the minimum requirement for fault identification as the decision to say circuit under test (CUT) is faulty or not. The fault location and identification are to locate faulty sub circuit address then identify the faulty parameter. The fault evaluation is to say how much the parameter deviation referred to the nominal. The fault prediction is to identify whether fault is about to occur concerning to replace elements before an actual failure occurs and with minimum loss in lifetime of the replaced element [8-10].

The difficulties in testing and diagnosis of analog circuits are much more than digital circuits due to lack of efficient fault modeling, limited accessibility, tolerance, nonlinearity, and parametric effects (see parametric faults). Another difficulty in diagnosis exists when more than one fault has the same signatures. In such cases, a fault doesn't satisfy a good diagnostic degree for isolation owing to more cost in testing each element alone [1,11].

The analog fault identification and diagnosis (in specification testing after assuming

success detection) are classified as simulation-before-test approach and Simulation-after-test approach. The main steps in both approaches are shown briefly in Figure 1.

The simulation-after-test (SAT) approach assumes that enough information is available in the measurements and that measurements are mutually independent, so fault isolation is obtained by estimating the circuit parameters from the CUT outputs. The method suffers from several drawbacks associated with the identification procedure if the system is nonlinear or local minima issues arise. In addition, SAT techniques are time consuming when applied to large circuits. The SAT approach can either use parametric identification techniques (after some sufficient measurement techniques), fault verification techniques (after some limited

measurement techniques), or approximation techniques (after optimization-based techniques and limited measurements techniques too) [12-15].

The simulation-before-test (SBT) approach is based on the comparison of the circuit responses associated with predefined test stimuli with those induced by different fault conditions, so the fault detection and isolation must be solved by a subsequent classification. Here, SBT techniques ensures a reduced test time also when complex circuits are envisaged. The SBT approach can either use fault dictionary techniques, or approximation techniques (after probabilistic based techniques). The approximation techniques use probabilistic methods (to characterize circuit statistically), artificial intelligence techniques, and testability different measures [13,15,16].

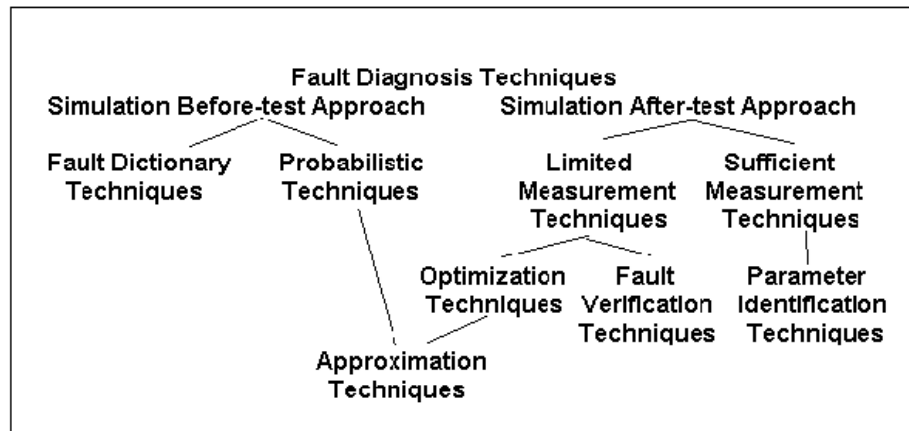


Figure 1: A summary of diagnosis techniques.

The Fault dictionary techniques approach is the most common test techniques in analog fault diagnosis used with SBT approach. It is based on pattern recognition i.e., comparing failed-board output levels with a set of pre stored outputs. In other words, comparing

signatures measured at the test nodes with those contained in a fault 'dictionary' before diagnosis phase [17,18]. In the fault dictionary techniques, many offline steps are required. These steps include good circuit and fault definitions, simulation, feature extraction,

fault dictionary construction, and finally a classification procedure to classify faults achieving the desired diagnostic degree.

Those requirements are represented by block diagrams in Figure 2.

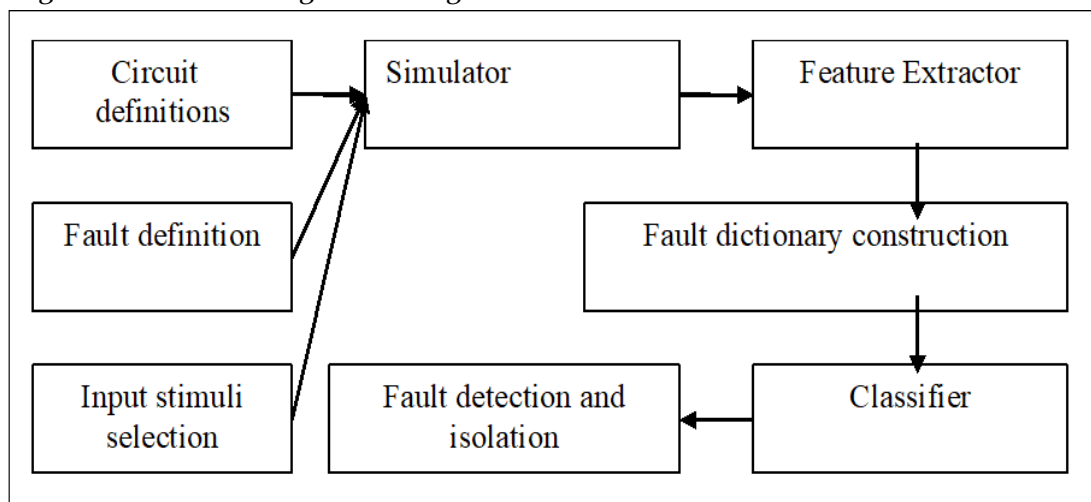


Figure 2: The main offline steps in fault dictionary SBT approach.

Three most important issues related to good circuit and fault definitions. The first is the selection of the most controllable and observable 'test' nodes in the CUT. The second is the selection of the most appropriate 'test' stimuli able to excite the CUT so that the faulty-induced effect propagates to an observable node. The third is the selection of the most likely faults. The greater the number of faults, the larger the size of the constructed dictionary.

The CUT simulation is to simulate the circuit with a certain stimulus which can either be dc or ac stimuli. The analysis can be done either in time or frequency domains. The frequency-domain simulation has many approaches such as: Sechu and Waxman approach, Bi-linear transformation approach, and sparse matrix recognition approach. The time-domain simulation has many approaches too, such as: Pseudo-noise signal approach and Test signal design approach [1,7,19,20]. The simulator tool could either be Workbench

simulator, P-spice simulators or many other different simulators.

The feature extraction is to extract a set of features from the simulated signals measured at the test nodes. The selected features must achieve the desired diagnostic degree (Detection and isolation of faults). The main objective is to reduce the amount of stored data (features' signatures) in the fault dictionary for easy fast processing and small storage. This will achieve the desired diagnostic degree with higher performance and lower cost.

Different optimum measurement selection techniques of feature extraction are used such as: Binary logical manipulation techniques, Heuristic techniques, Probabilistic theoretic techniques, and other statistical techniques. The feature extraction could either be done direct in time-domain, using Fourier's transforms, or wavelet transforms [1,20]. The fault dictionary consists of a lookup table

stored in it signatures of features, each of which is labeled by the indication of the faulty unit. So, the dictionary contains information associated with fault-free and faulty situations.

The CUT fault classification is to classify the circuit signatures by exploiting the knowledge present in the fault dictionary into sets of stimuli (inputs) and responses (outputs). The selected sets must detect and isolate the faults with the desired degree. The classifier tool could be using K-mean distance algorithm or artificial intelligent classifiers such as neural networks.

In the on-line fault identification (the fault identification for diagnosis), the faulty CUT will be excited by same stimuli used in dictionary construction. An efficient fault simulation technique is used. The obtained response signatures are compared with those stored in dictionary. Finally, a fault isolation criterion is implemented (either heuristic or statistic-based) to identify the faulty CUT to one of the pre stored faults or to an ambiguity set that contains a set of possible faults i.e., it's done by comparing implementation with dictionary entries.

Fault dictionary techniques approach has many advantages. It's suitable to linear and nonlinear analog circuits. It's also suitable to both catastrophic and parametric faults, with very high performance in single faults. Another advantage is the requirement of very limited number of accessible test nodes. The on-line computational requirements are minimal in contrast of the offline computational requirements. Finally, the most important advantage for this approach

is that it can achieve a very high diagnostic degree when diagnosis is limited to ambiguity set level. These advantages have made the fault dictionary techniques the most logical for practical considerations.

Two major disadvantages for the fault dictionary techniques approach. The first is the high off-line computational requirements due to need of considering large number of fault cases, but this is solved by using decomposition in dictionary construction. The second is the weak robustness i.e., degradation in diagnostic degree due to deviation of non-faulty elements-inside the tolerance region [1].

Circuit simulation using PSPICE software

Simulation is a very powerful and cost-effective tool to use during design and development of electronic circuits. It is one of the main steps for fault testing and diagnosis. The response of circuit under test (CUT) is usually simulated when a proper is an applied at its input. Two simulation types are usually done for the CUT namely, direct current (DC) analysis and alternating current (AC). (AC) analysis is either done in time or frequency domains. (AC) analysis gives a complete view for a signal in real time.

There are different programs that can be used for simulation (Electronic Work Bench and PSPICE). (PSPICE) program was used because it is a simple and do what is easier than other programs.

PSPICE software program

SPICE stands for (Simulation Program with Integrated Circuit Emphasis) It was developed at the university of California

Berkeley. A commercial version of SPICE called PSPICE is developed by Microsim Corporation [21].

One of the program advantages is that it helps user to simulate the circuit design graphically on the computer before building a physical circuit. Hence, the designer can make any necessary changes on the prototype without modifying any hardware. As soon as the test design is completed, PSPICE can help in circuit check before deciding to build a

hardware model. Hence, PSPICE allows user to check the operability of the circuit model in real life simulations to validate its viability. Since all the tests, designs and modifications are made over a terminal; the designer can save a lot of money that would have otherwise been spent on the building of models and modifying them [22].

A sample version of PSPICE is used in our project. It first describes the CUT in a text file having the extension (.cir).

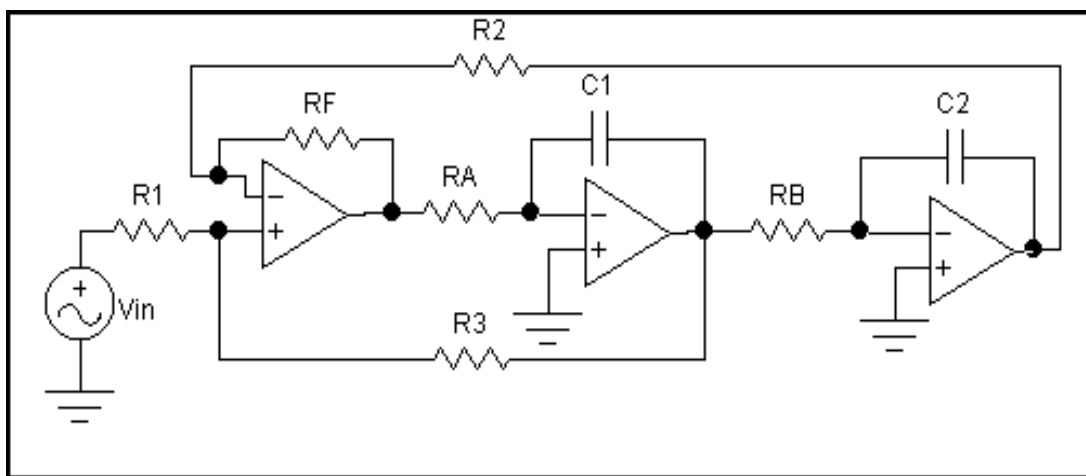


Figure 3: The shown circuit is identified by the following PSPICE circuit file.

Feature extraction

The module of feature extraction operates on the simulated waveforms at selected test nodes. It aims to reduce the dimensionality problem by replacing time-domain measurements by a set of features capable of highlighting faults. The selected features must have minimum number of data (features) achieving the desired diagnostic degree. The main objective is to store the signatures of those features giving minimum amount of data in the fault dictionary. This leads to easy fast processing and small storage area which means lower cost and higher

performance of test to achieve the desired diagnostic degree.

Different optimum measurement selection techniques of feature extraction are used such as: Binary logical manipulation techniques, Heuristic techniques, Probabilistic theoretic techniques, and other statistical techniques. The extractor tool could either be direct time-domain analysis, Fourier's transforms and frequency-domain analysis, or wavelet transforms and analysis [1,20].

In signal applications, one is always concerned with transmitting, processing, and

using information in any of the four areas (instrumentation, control, communication, and computers). In analog information systems, information can be represented by electrical voltage waves whose amplitude is analogous to the quantity being represented. In digital information systems, such information can be represented by digital quantities numbers which greatly reduces design and production quality problems. Such digital devices are well suited to computer systems and similar data processing systems and storage [20].

In our fault identification and diagnosis problem, digital information was used. It has one advantage of perfect noiseless transmission that can be required if communicating to diagnose the CUT and exchange data between two far points in manufacturing and industry. A second advantage is the perfect facilitate processing

that will lead us to high performance computation that is required especially in computing microelectronic circuits to have a high-resolution feature. A third advantage is the perfect binary storage and fast accessing its memory units.

In the feature extraction process, one can either extract features by having the effective samples of time-domain or having the effective samples of frequency-domain-spectrum- of the time-domain signal. Some other modern techniques can be used such as wavelet analysis, frequency-scale analysis, and short time-frequency analysis.

Spectrum analyzers

There are laboratory instruments called spectrum analyzers which are designed to calculate and display the frequency spectrum of a voltage signal. Internally, an FFT is calculated Figure 4 [26].

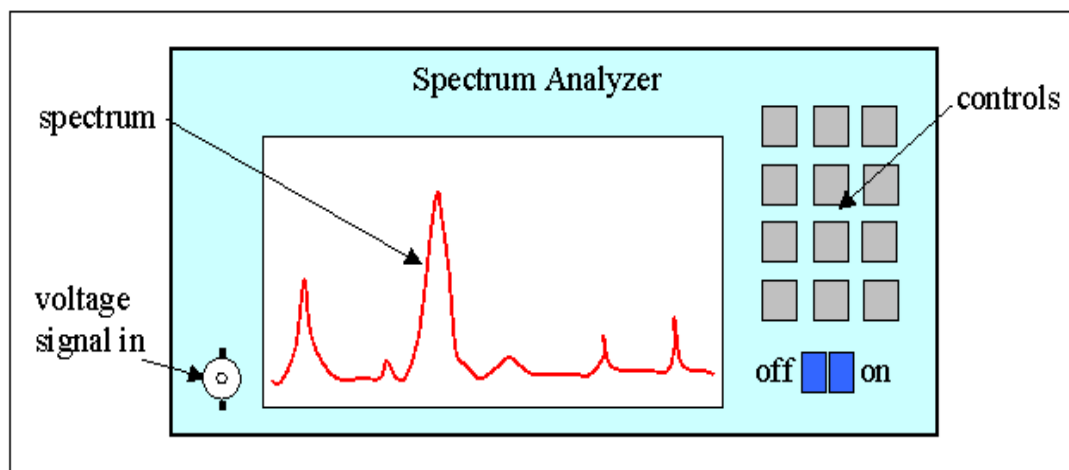


Figure 4: A spectrum analyzer usually actuates with FFT.

Wavelet analysis

The wavelet feature extraction is based on having sum overall time of the signal

multiplied by scaled and shifted versions of a certain standard wavelet function. Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly,

wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet. A wavelet is a waveform

of effectively limited duration that has an average value of zero.

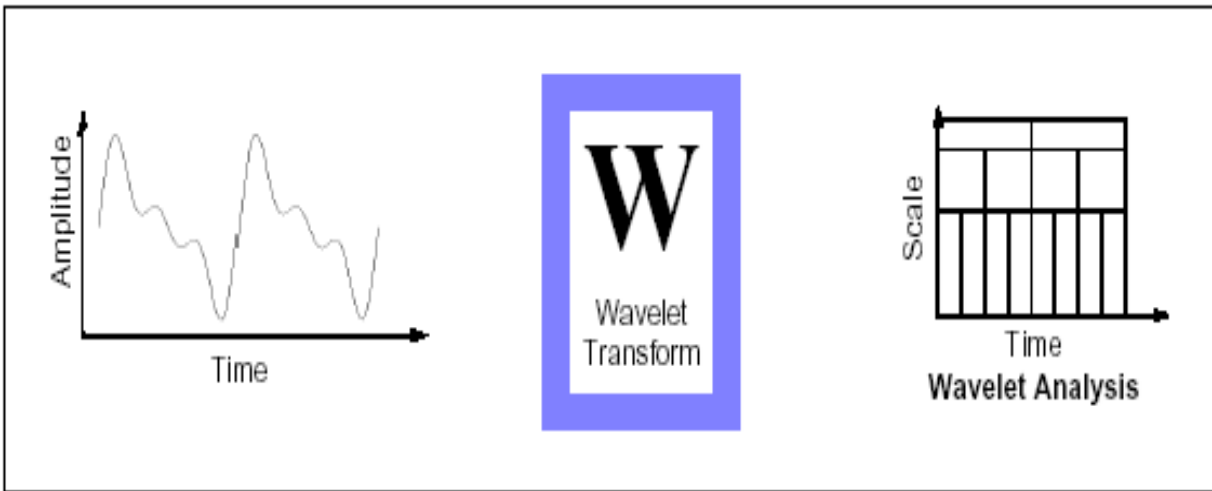


Figure 5: A representation for the wavelet transform.

The advantages of this technique that it is capable of revealing aspects of data that another signal analysis techniques miss, aspects like trends, breakdown points, discontinuities in higher derivatives, and self-similarity.

Furthermore, because it affords a different view of data than those presented by traditional techniques, wavelet analysis can often compress or de-noise a signal without appreciable degradation. Indeed, in their brief history within the signal processing field, wavelets have already proven themselves to be an indispensable addition to the analyst's collection of tools and continue to enjoy a burgeoning popularity today [24]. The only disadvantage for this type is the requirement to some computational transformations. The wavelet scale aspects are more applied in biology for cell membrane recognition, metallurgy for the characterization of rough surfaces, finance for

detecting the properties of quick variation of values, and in Internet traffic description for designing the services size. Thus far, as one-dimensional data, which encompasses most ordinary signals, wavelet analysis can be applied to two-dimensional data (images) and, in principle, to higher dimensional data.

In analog circuit identification and diagnosis, due to these advantages a very high diagnostic degree in our field can be achieved.

This type is applied in our experiments giving a very high accuracy values after classification process compared to the time-domain analysis (Circuit 2).

Classification

Classification means a way of grouping objects, on the basis of some relationship between them. The term classification refers to the placing of objects into groups in such a way that the members of the groups bear a

closer relationship to other members of the same group than they do to members of other group.

Classification process has two main stages, in the first stage, the number and nature of the categories are determined, while in the second stage every unknown element is assigned to one of the categories according to its level of resemblance (or similarity). These stages are often called classification and identification, respectively.

K-Mean algorithm

This algorithm tries to select an optimal set of points which are placed at the centroids of clusters of training data. Given K radial units, it adjusts the positions of the centers such that each training point "belongs to" a cluster center and is nearer to this center than to any other center. Each cluster center is the centroid of the training points which belong to it. Once centers are assigned, deviations are set.

The size of the deviation (also known as a smoothing factor) determines how "spiky" the Gaussian functions are. If the Gaussians are too spiky, the network will not interpolate between known points and the network loses the ability to generalize.

If the Gaussians are very broad, the network loses fine detail. This is actually another way to show the fitting dilemma, deviations should typically be chosen so that Gaussians overlap with "a few" nearby centers. Methods available are Explicit and isotropic, Explicit allows choosing deviation while for isotropic, the deviation is the same for all units. It is selected heuristically to reflect the number of

centers and the volume of space they occupy [27].

Iterative partitioning clustering algorithm

The algorithm has four basic steps. In Step 1, an initial partition with K clusters is selected. An initial partition can be formed by first specifying a set of K seed points which can be the first K patterns or K patterns chosen randomly from the pattern matrix. In Step 2, a new partition is generated by assigning each pattern to its closest cluster center.

In Step 3, new cluster centers are computed as the centroids of clusters, A set of K patterns that are well separated from each other can be obtained by taking the centroid of the data as the first seed point and selection successive seed points which are at least a certain distance away from the seed points already chosen, Partitions are updated by reassigning patterns to clusters in an attempt to reduce the square-error.

The Euclidean metric is the most common metric for computing the distance between pattern and a cluster center. In Step 4, Adjust the number of clusters by merging and splitting existing clusters or by removing small, or outlier, clusters, some clustering algorithms can create new clusters or merge existing clusters if certain conditions are met [3].

A reference signature is obtained for each class by computing an average of its neighbor patterns. The input patterns are considered to belong to the signature having minimum distance to these patterns.

Input patterns shortest distance reference signature

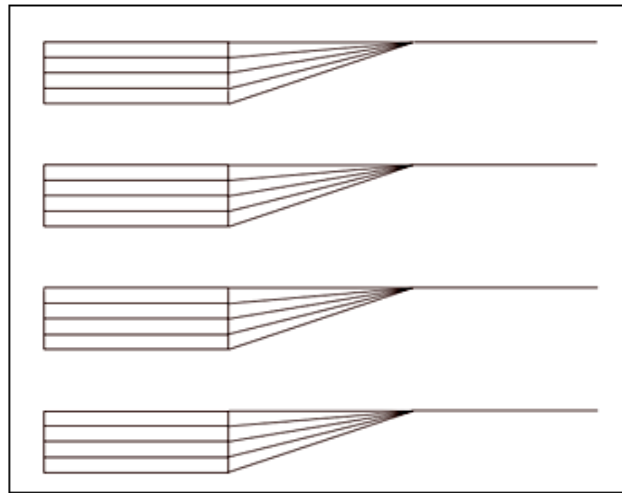


Figure 6: Euclidian distance between the input patterns and the reference signature.

Artificial Neural Networks

Artificial neural networks (ANN) are collections of mathematical models that emulate some of the observed properties of biological nervous systems. The key element of the ANN paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements that are analogous to neurons and are tied together with weights [4].

Brains store information as patterns. Some of these patterns are very complicated and allow us the ability to recognize individual faces from many different angles. This process of storing information as patterns, utilizing those patterns, and then solving problems encompasses a new field in computing. This field does not utilize traditional programming but involves the creation of massively parallel networks and the training of those networks to solve specific problems. This field also utilizes words very different from traditional

computing, words like behave, react, self-organize, learn, generalize, and forget.

Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons; this is true of ANN. Learning typically occurs by example through training, or exposure to a training set of input/output data where the training algorithm iteratively adjusts the connection weights (synapses). These connection weights store the knowledge necessary to solve specific problems.

The advantage of ANN lies in their resilience against distortions in the input data and their capability of learning. They are often good at solving problems that are too complex for conventional technologies (e.g., problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found) and are often well suited to problems that people are good at solving, but for which traditional methods are not.

There are multitudes of different types of ANN. Some ANN are classified as feed forward while others are recurrent (i.e., implement feedback) depending on how data is processed through the network. Another way of classifying ANN types is by their method of learning (or training). ANN can be implemented in software or in specialized hardware [4].

Traditional computing methods work well for problems that can be well characterized, for example balancing checkbooks, networks results, protect equipments, these applications do not need the special characteristics of neural networks. Neural networks offer a different way to analyze data,

and to recognize patterns within that data, than traditional computing methods. However, they are not a solution for all computing problems. ANN has found many applications, in sensor processing, speech recognition, pattern recognition and text recognition [28].

Structure of neural network

Neuron model

A neuron with a single R-element input vector is shown below, here the individual element inputs are multiplied by weights and the weighted values are fed to the summing junction.

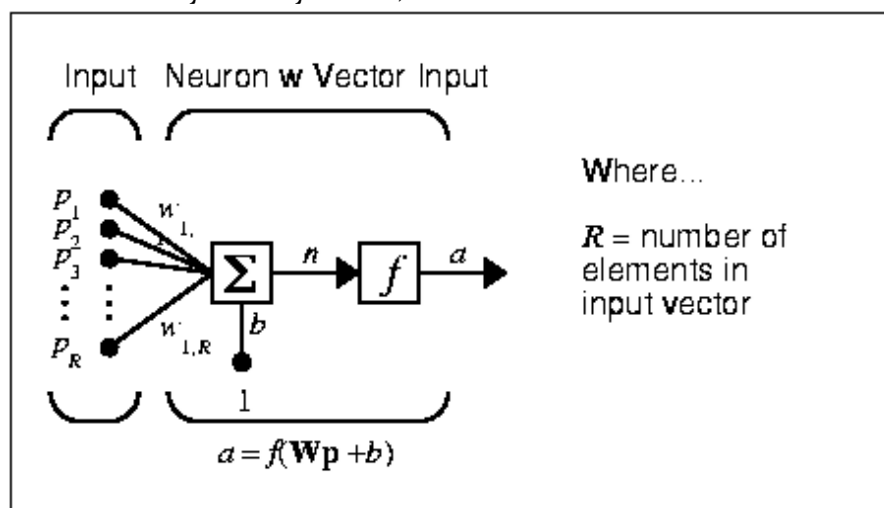


Figure 7: Neuron model.

The central idea of neural networks is that such parameters can be adjusted so that the network exhibits some desired or interesting behavior.

Thus, one can train the network to do a particular job by adjusting the weight or bias parameters, or perhaps the network itself will

adjust these parameters to achieve some desired end [6].

Transfer functions

In the transfer function the summation total can be compared with some threshold to determine the neural output. If the sum is greater than the threshold value, the processing element generates a signal. If the

sum of the input and weight is less than the threshold, no signal (or some inhibitory

signal) is generated. Both types of response are significant.

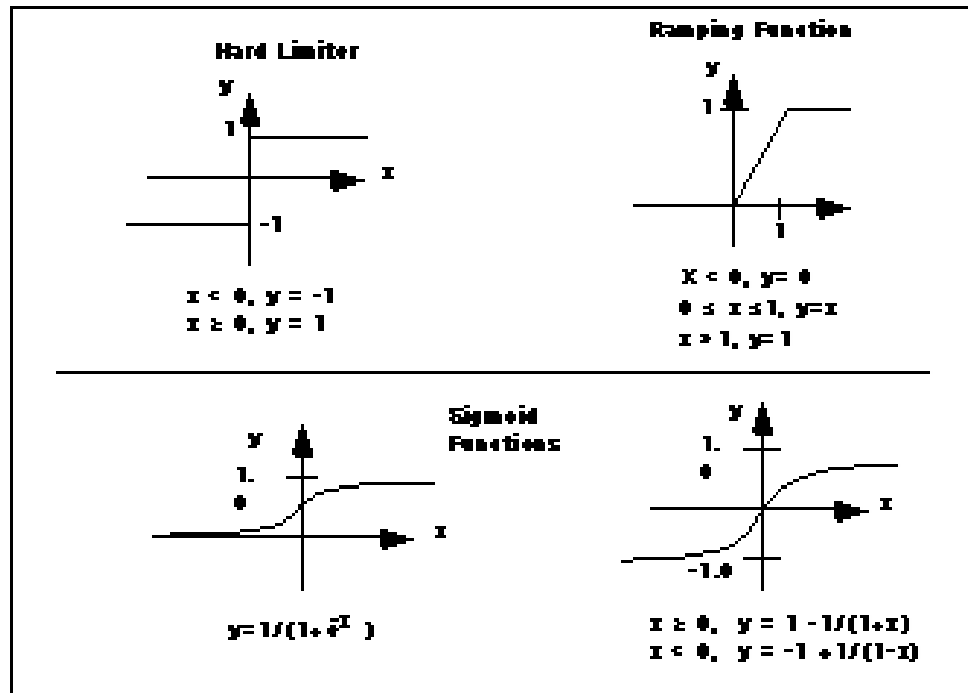


Figure 8: Transfer functions.

1. The hard limiter transfer function
2. The sigmoid transfer functions
3. The linear transfer functions
4. The ramping transfer function

Artificial network topology

Currently, neural networks are the simple clustering of the primitive artificial neurons. This clustering occurs by creating layers which are then connected to one another. Basically, all artificial neural networks have a similar structure or topology as shown in Figure 9.

In that structure there is input, hidden, and output layer. The layer of input neurons receives the data either from input files or directly from electronic sensors. The output

layer sends information directly to the outside world, between these two layers can be many hidden layers. These internal layers contain many of the neurons in various interconnected structures. The inputs and outputs of each of these hidden neurons simply go to other neurons.

Number of neurons in each layer depends on number of features, for input layer, number of output classes for output layer, trial and error for hidden layers.

In most networks each neuron in a hidden layer receives the signals from all of the neurons in a layer above it, after a neuron performs its function, it passes its output to all of the neurons in the layer below it, providing a feed forward path to the output.

These lines of communication from one neuron to another are important aspects of neural networks, they are the glue to the system, and they are the connections which provide a variable strength to an input. There are two types of these connections. One causes the summing mechanism of the next neuron to add while the other causes it to subtract. In more human terms one excites while the other inhibits. Some networks want

a neuron to inhibit the other neurons in the same layer, this is called lateral inhibition. The most common use of this is in the output layer. It can do that with lateral inhibition. This concept is also called competition. Another type of connection is feedback; this is where the output of one-layer routes back to a previous layer. An example is shown in Figure 10.

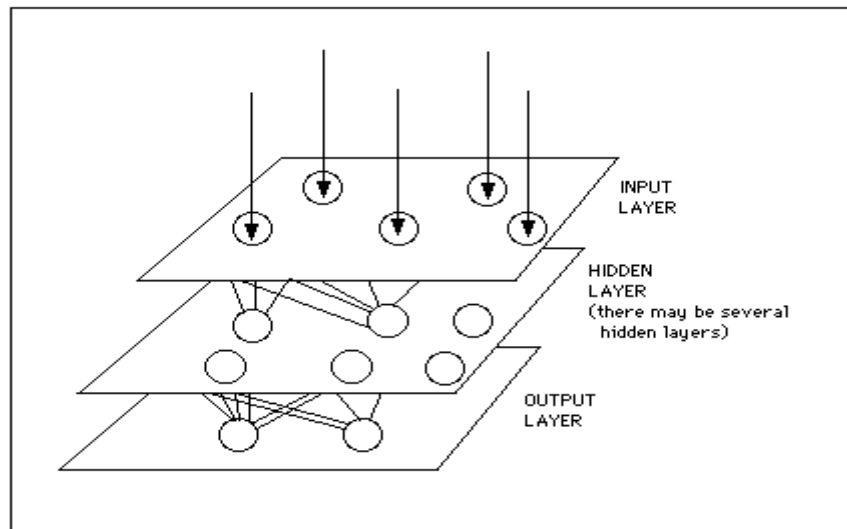


Figure 9: A simple NN diagram.

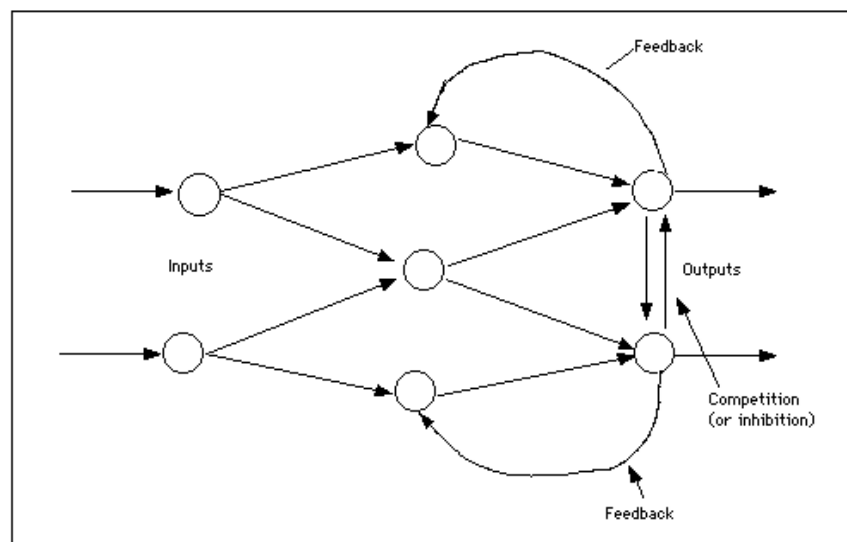


Figure 10: Simple networks with feedback and competition.

Training an artificial neural network

The training process is started by randomly choosing weights. Then, the training, or learning, begins. There are two approaches for training-supervised and unsupervised. Supervised training involves a mechanism of providing the network with the desired output either by manually "grading" the network's performance or by providing desired outputs with the inputs; supervised training is analogous to a student guided by an instructor.

In this mode, the actual output of a neural network is compared to the desired output. Weights, which are usually randomly set to begin with, are then adjusted by the network so that the next iteration, or cycle, will produce a closer match between the desired and the actual output. The learning method tries to minimize the current errors of all processing elements.

This global error reduction is created over time by continuously modifying the input weights until acceptable network accuracy is reached. After a supervised network performs well on the training data, then it is important to see what it can do with data it has not seen before. If a system does not give reasonable outputs for this test set, the training period is not over. Indeed, this testing is critical to ensure that the network has not simply memorized a given set of data but has learned the general patterns involved.

Unsupervised training is where the network has to make sense of the inputs without outside help. Unsupervised algorithms essentially perform clustering of the data into

similar groups based on the measured attributes or features serving as inputs to the algorithms. This is analogous to a student who derives the lesson totally on his or her own. (Artificial network technology) [28].

The properties of ANN

ANN has the following properties:

- Ability to generalize in making dissections about imprecise input data
- Ability to deal with nonlinear problems
- Robust classification (ability to achieve good in presence of errors)
- Offering ideal solutions to a variety of classification problems such as speech, character and signal recognition, as well as functional prediction and system modeling where the physical processes are not understood or are highly complex.
- Operating as pattern recognition engines

Types and applications of neural networks

Because all artificial neural networks are based on the concept of neurons, connections and transfer functions, there is a similarity between the different neural networks. Most of the variations stems from the various learning rules and how those rules modify a network's typical topology.

Basically, most applications of neural networks fall into the following five categories:

Network Type	Networks	Use for Network
Prediction	Back-propagation Delta Bar Delta Extended Delta Bar Delta Directed Random Search Self-organizing map into Back-propagation	Use input values to predict some output (e.g., pick the best stocks in the market, predict weather, identify people with cancer risks etc.)
Classification	Learning Vector Quantization Counter-propagation Probabilistic Neural Networks	Use input values to determine the classification (e.g., is the input the letter A, is the blob of video data a plane and what kind of plane is it)
Data Association	Hopfield Boltzmann Machine Hamming Network Bidirectional associative Memory Spation-temporal Pattern Recognition	Like Classification but it also recognizes data that contains errors (e.g., not only identify the characters that were scanned but identify when the scanner isn't working properly)
Data Conceptualization	Adaptive Resonance Network Self-Organizing Map	Analyze the inputs so that grouping relationships can be inferred (e.g., extract from a database the names of those most likely to buy a particular product)
Data Filtering	Recirculation	Smooth an input signal (e.g., take the noise out of a telephone signal)

Table 1: The differences between the network categories and shows which of the more common network topologies belong to which primary category.

Probabilistic Neural Network (PNN)

The PNN was developed by Donald Specht. They are a kind of radial basis network. This network provides a general solution to pattern classification problems by following an approach developed in statistics, called Bayesian classifiers. Bayes theory developed in the 1950's considers the relative likelihood of events and uses a priori information to improve prediction. The network paradigm also uses Parzen Estimators which were developed to construct the probability density functions required by Bayes theory. The greatest advantage of PNN is the fact that the output is probabilistic (which makes interpretation of output easy), and the training speed. Training a PNN actually

consists mostly of copying training cases into the network, and so is as close to instantaneous as can be expected. The greatest disadvantage is network size, a PNN network actually contains the entire set of training cases and is therefore space-consuming and slow to execute. They are slower to operate because they use more computation than other kinds of networks to do their function approximation [29].

Learning of PNN

The PNN uses a supervised training set to develop distribution functions within a pattern layer. These functions, in the recall mode, are used to estimate the likelihood of an input feature vector being part of a learned category, or class. The learned patterns can

also be combined, or weighted, with the a priori probability, also called the relative frequency, of each category to determine the most likely class for a given input vector. If the relative frequency of the categories is unknown, then all categories can be assumed to be equally likely, and the determination of category is solely based on the closeness of the input feature vector to the distribution function of a class. An example of a probabilistic neural network is shown in Figure 11.

Probabilistic neural network structure and topology

This network has three layers. The network contains an input layer which has as many elements as there are separable parameters needed to describe the objects to be classified. It has a pattern layer, which organizes the training set such that each input vector is represented by an individual processing

element and finally, the network contains an output layer, called the summation layer, which has as many processing elements as there are classes to be recognized. Each element in this layer combines via processing elements within the pattern layer which relate to the same class and prepares that category for output. Sometimes a fourth layer is added to normalize the input vector if the inputs are not already normalized before they enter the network. The pattern layer represents a neural implementation of a version of a Bayes classifier, where the class dependent probability density functions are approximated using a Parzen estimator. This approach provides an optimum pattern classifier in terms of minimizing the expected risk of wrongly classifying an object. With the estimator, the approach gets closer to the true underlying class density functions as the number of training samples increases, so long as the training set is an adequate representation of the class distinctions.

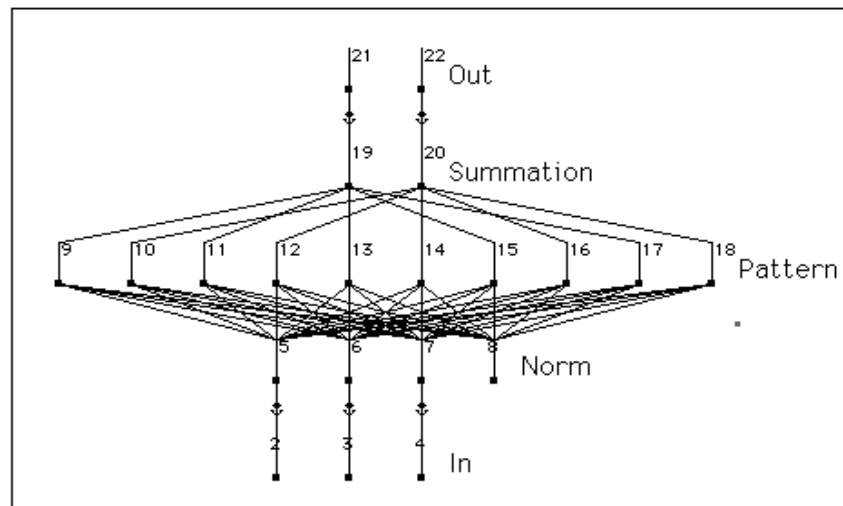


Figure 11: A PNN Example.

In the pattern layer, there is a processing element for each input vector in the training

set. Normally, there are equal amounts of processing elements for each output class.

Otherwise, one or more classes may be skewed incorrectly, and the network will generate poor results. Each processing element in the pattern layer is trained once. An element is trained to generate a high output value when an input vector matches the training vector. The training function may include a global smoothing factor to better generalize classification results. In any case, the training vectors do not have to be in any special order in the training set, since the category of a particular vector is specified by the desired output of the input. The learning function simply selects the first untrained processing element in the correct output class and modifies its weights to match the training vector.

The pattern layer operates competitively, where only the highest match to an input vector wins and generates an output. In this way, only one classification category is generated for any given input vector. If the input does not relate well to any patterns programmed into the pattern layer, no output is generated. The Parzen estimation can be added to the pattern layer to fine tune the classification of objects, this is done by adding the frequency of occurrence for each training pattern built into a processing element. Basically, the probability distribution of occurrence for each example in a class is multiplied into its respective training node. In this way, a more accurate expectation of an object is added to the features which make it recognizable as a class member.

Training of the probabilistic neural network is much simpler than with back-propagation. However, the pattern layer can be quite huge

if the distinction between categories is varied and at the same time quite similar is special areas. There are many proponents for this type of network, since the groundwork for optimization is founded in well known, classical mathematics [28].

SPREAD determines the width of an area in the input space to which each neuron responds. If SPREAD is 4, then each radbas neuron will respond with 0.5 or more to any input vectors within a vector distance of 4 from their weight vector. Each bias in the first layer is set to $0.8326/\text{SPREAD}$. This gives radial basis functions that cross 0.5 at weighted inputs of $\pm \text{SPREAD}$. SPREAD should be large enough that neurons respond strongly to overlapping regions of the input space. If spread is near zero, the network will act as a nearest neighbor classifier. As spread becomes larger, the designed network will consider several nearby design vectors. SPREAD is large enough so that the active input regions of the radbas neurons overlap enough so that several radbas neurons always have fairly large outputs at any given moment. This makes the network function smoother and results in better generalization for new input vectors occurring [24]. Here the net input to the radbas transfer function is the Vector distance between its weight vector w and the input vector p , multiplied by the bias b . The box in this figure accepts the input vector p and the single row input Weight matrix and produces the dot product of the two. And then the output is multiplied by bias in the first layer that is set to $0.8326/\text{SPREAD}$. This gives an output n that will enter the transfer function.

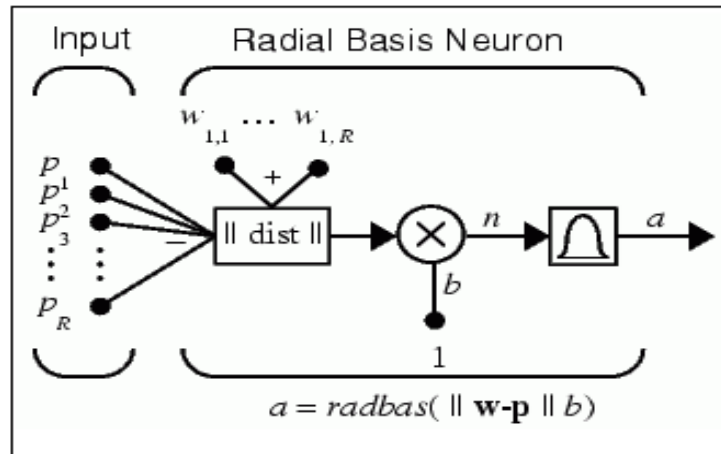


Figure 12: The transfer function for a radial basis neuron.

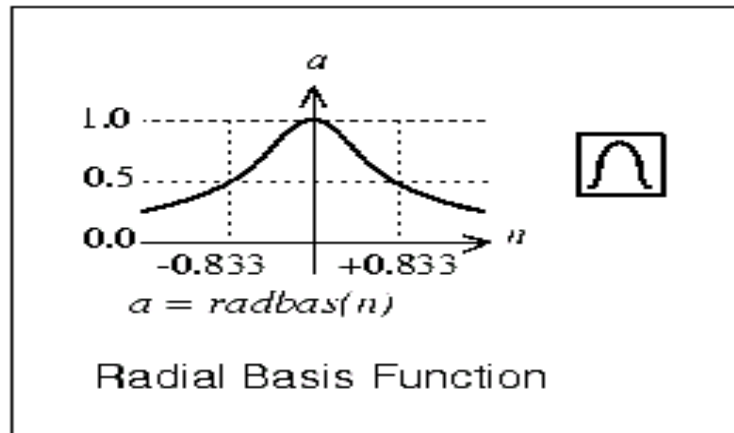


Figure 13: The transfer function for a radial basis neuron.

Experimental results

Our project considers the problem of hard fault classification. For each faulty element, we consider two hard faults short circuit (S.C.) and open circuit (O.C.). For resistors, S.C. fault is simulated by setting the value of resistance to $1\ \Omega$, an O.C. fault is simulated by setting the value of resistance to $100\text{M}\Omega$. For capacitors, S.C. fault is simulated by setting the value of capacitance to $100\mu\text{F}$, O.C. fault is simulated by setting the value of capacitance to 1pF . The first issues related to good circuit and fault definitions is the selection of the

most appropriate 'test' stimuli able to excite the circuit under test (CUT) so that the faulty-induced effect propagates to an observable node, the sensitivity method was used. The second is the selection of the most controllable and observable 'test' nodes in the CUT. The third is the selection of the most likely faults. The simulation is based on ac time domain analysis. The tool used is the PSPICE circuit simulator. The feature extraction is based on discrete time-domain, frequency-domain and wavelet analysis. The tools used are direct time extraction, Fourier Fast Transform (FFT), and Db Wavelet

transform. The classification is based on Euclidean distance and artificial neural network. The tools used are the Euclidean Distance K-mean Algorithm and Probabilistic Neural Networks (PNN) classifiers. Three active circuit examples are used to demonstrate our work. The first example considers a three-stage low-band-high pass filter [30]. The second a band-stop filter [31]. The third is a Sallen-Key filter [32]. The

objective is to calc. classification accuracy (CA).

Example 1: Low-band-high Pass (LBH) active filter circuit

The circuit's outputs at nodes 4, 6, and 8 are respectively the high-pass, band-pass, and low-pass filters' outputs as seen Figure 14. The frequency response of the circuit is seen in Figure 15. It consists of 3 operational amplifiers, 6 resistors and 2 capacitors.

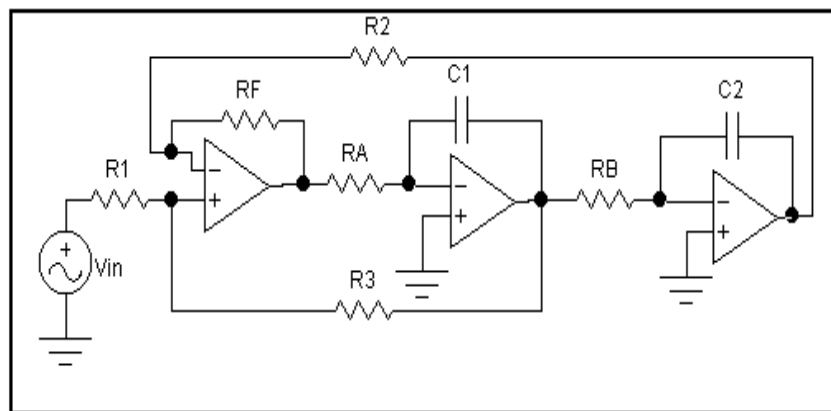


Figure 14: Low-band-high Pass (LBH) Active Filter Circuit.

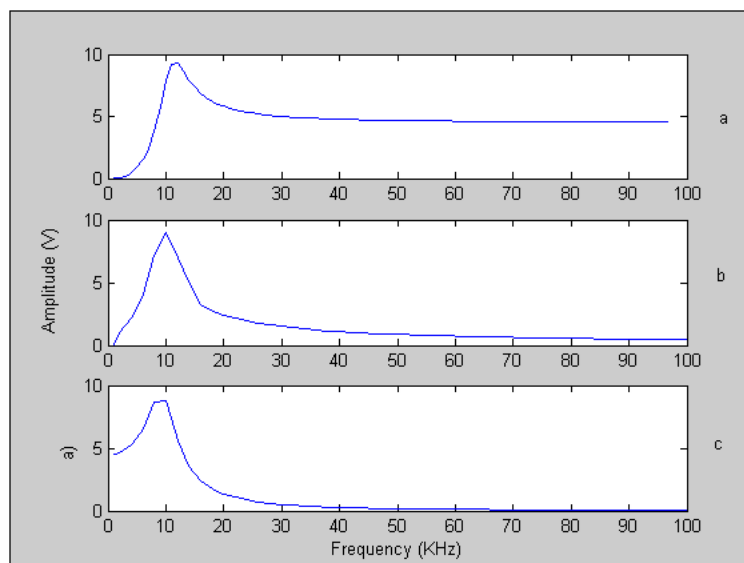


Figure 15: The frequency response of the circuit 1(a) at node 4(b) at node 6(c) at node 8.

Single excitation

Simulation

The CUT is simulated at fault-free and different faulty conditions using PSPICE simulator. All circuit capacitors are nominally ∞ and are assumed to be fault free for simplicity. The method selected was the stimuli frequency is based in a classic method of picking a frequency falling around cutoff frequencies in the most effective nodes manually (by try and error every test). The frequency response of the circuit at nominal is shown in Figure 15. Here, the selected test stimulus is a single sinusoidal signal ac 10KHz without any input resistances. The selected test nodes are 2, 3, 4, 6, and 8. The nominal

values of the circuit resistors and a list of 12 hard faults (6 faults with S.C. resistance and 6 with O.C. resistance) are shown in Table 2.

The selected sampling frequency must be greater than 40KHz. 200KHz which means a sampling period of 5 μ s is simulated during an interval of 0.1ms (which means each waveform is simulated by 21 points). For each of the 13 (nominal +12 faults) fault Scenario, 100 Monte Carlo (MC) simulations are carried out. During each MC simulation, fault free elements are tolerated within 5% around its nominal values. The number of features will be 21. So, for each fault Scenario, simulated data are collected in a matrix (1300*21) each node.

Element	Nominal value	Fault class/ value	Fault index	Fault class/ value	Fault index
R1	10 K Ω	S.C. (1 Ω)	1	O.C. (100M)	7
R2	10 K Ω	S.C.	2	O.C.	8
R3	30 K Ω	S.C.	3	O.C.	9
Ra	15.9 K Ω	S.C.	4	O.C.	10
Rb	15.9 K Ω	S.C.	5	O.C.	11
Rf	10 K Ω	S.C.	6	O.C.	12

Table 2: The selected hard faults and its description.

Feature extraction

Time-domain

The total number of patterns obtained for each of the selected nodes, are equal to the number of fault scenarios multiplied by the number of MC simulations to be 1300 patterns. Each pattern will have 21 features each node. Calculated features are collected in a matrix (1300*21) that's average matrix (13*21) each node. A sample of an average of

100 MC patterns for fault scenario 6 at node 6 is seen in Figure 16(a).

Frequency-domain

For each pattern (21 samples) in time domain, 16 Fourier Transform (FT) coefficients are computed. It represents the DC and the magnitude of the first 15 harmonics. The second half of the FT coefficients is redundant. Calculated features are collected in a matrix (1300x16) that's average matrix

(13*16) each node. A sample of the average signal of 100 FFT patterns for fault scenario 7

at combination of nodes (2,6) is seen in Figure 16(b).

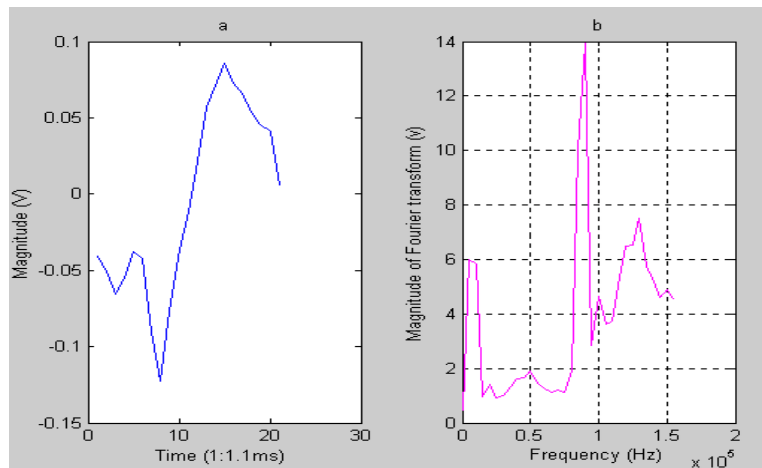


Figure 16: (a) The average signal of 100 MC for fault scenario 6 at the node 6. (b) The average signal of 100 FFT patterns for fault scenario 7 at combination of nodes (2,6).

Classification

There are several classification algorithms used in fault separation.

The K-mean algorithm (based on the Euclidean distance) and the PNN more accurate classifier was used.

Single Node	CA %	Dual-Nodes	CA %	Triple-Nodes	CA %
2	71	2,3	69.2308	2,3,4	72.2308
3	69	2,4	72.3077	2,3,6	75.5385
4	69.3077	2,6	76.3846	2,3,8	74.2308
6	76.3846	2,8	74.0769	2,4,6	73.8462
8	74.1538	3,4	72.2308	2,4,8	72.3077
		3,6	75.5385	2,6,8	74.6923
		3,8	74.2308	3,4,6	73.7692
		4,6	73.8462	3,4,8	72.3077
		4,8	72.3077	3,6,8	74.7692
		6,8	74.6923	4,6,8	73.4615

Table 3: The total CAs due to Euclidean distance classification on time-domain.

Classification matrix concluded that: 1) Fault scenarios 0,3,4,5,8,10,11 can be considered to be uniquely identified. 2) Fault scenarios (1,12), (2,9), and (6,7) can be collected in 3 ambiguity groups.

Taking these into considerations, each ambiguity group has weight of 100 added to the true weights of uniquely identified faults (from Table 2). The classification CA will rise to 96.9%.

Single Node	CA%	Dual-Nodes	CA%	Triple-Nodes	CA%
2	70.5385	2,3	68.7692	2,3,4	68.7692
3	70.1538	2,4	77	2,3,6	68.7692
4	77.9231	2,6	77.7692	2,3,8	68.7692
6	78.4615	2,8	75.7692	2,4,6	77
8	77.2308	3,4	77.1538	2,4,8	77
		3,6	76.4615	2,6,8	77.7692
		3,8	75.7692	3,4,6	77.1538
		4,6	77.4615	3,4,8	77.1538
		4,8	76.0769	3,6,8	76.4615
		6,8	76.2308	4,6,8	77.4615

Table 4: The total CAs due frequency-domain feature extraction and K-mean classification.

By selecting some features that's most effective to reduce the features from 16 to only 6 features, the dc component and 5 harmonic components having maximum CA at node 6 is 76.0769% for set of features {1,2,3,6,7,13} was taken. Classification matrix, concluded

that: 1) Fault scenarios 0,3,4,5,8,10,11 can be considered to be uniquely identified. 2) Fault scenarios (1,12), (2,9), and (6,7) can be collected in 3 ambiguity groups see Table 5. Taking these into considerations, the classification CA will rise to 99.4%.

Element	V (2,6)		V (6)	
R1	1	7	1	7
R2	2	8	2	8
R3	3	9	3	9
Ra	4	10	4	10
Rb	5	11	5	11
Rf	6	12	6	12

Table 5: Ambiguity groups for single node 6 and dual node (2,6).

Artificial neural network classification on frequency-domain features

Since obtained results using K-mean classifier, are rather poor, PNN as a classifier was used. Collected patterns (1300) patterns are divided into two parts, 650 patterns used as training data and 650 patterns are used as testing data. The execution result of classification by applying the PNN on the samples after calculating FFT is train

classification matrix and test classification matrix of (13*13).

Different values of NN parameters (spread) are tried to get best CA. The total CAs allowing best single and combined nodes seen in the past section (single node 6, dual nodes 2,6) are printed depending on spread values with some random permutations shown Table 6. The highest CA at single node 6 is 85.07% (at spread 0.07) was observed.

Spread	Single node 6 CA%		Dual nodes 2,6 CA%	
	Training	Total	Training	Total
1	82	79.78	82.3	79.84
0.7	83.08	80.54	86	81.69
0.5	87.54	82.46	87.53	82.53
0.3	90.46	83.54	91.38	84.92
0.1	98.15	83.31	99.84	84.85
0.09	99.39	84.46	99.84	84.23
0.07	99.54	85.07	100	85.31
0.05	99.84	84.23	100	83.61
0.03	100	80.61	100	81.23
0.01	99.84	75.23	100	74.69

Table 6: The training and total CAs using PNN classifier in frequency-domain (with spreads).

By selecting classes from classification matrix of K-mean algorithm (at frequency domain) for each of the single and dual nodes (Table 6), the ambiguity groups are (1,12), (2,9), and

(6,7) can be collected in 3 ambiguity groups. Taking these into considerations, the classification CA will raise to 99.92% (at spread 0.9). This is seen in Table 7.

Spread	Single node 6 CA%		Dual nodes 2,6 CA%	
	Training	Total	Training	Total
1	99.84	99.92	100.00	99.85
0.9	99.85	99.92	100.00	100
0.7	99.85	99.77	100.00	100
0.5	99.85	99.61	100.00	99.85
0.3	99.85	99.38	100.00	99.07
0.1	99.85	96.07	100.00	96.07
0.09	100	96	100.00	96.07
0.07	100	96.15	100.00	96.15
0.05	99.85	95.38	100.00	95.13

Table 7: The training and total CAs using PNN classifier in frequency-domain (with spreads) after selecting classes according to ambiguity groups.

Multiple excitations

Simulation

The CUT is simulated at fault-free and different faulty conditions using PSPICE simulator. All circuit capacitors are nominally 1nF and are assumed to be fault free for

simplicity. The selected test stimuli are three sinusoidal signals ac 10KHz, 25KHz, 35KHz with input resistances of 50Ohms. The selected test nodes are 2,3,4,6, and 8. The nominal values of the circuit resistors and a list of 12 hard faults (6 faults with S.C. resistance and 6 with O.C. resistance) are shown in Table 1. The selected sampling

frequency must be greater than 140KHz. 200KHz which means a sampling period of 5 us is simulated during an interval of 0.155ms (which means each waveform is simulated by 32 points) was selected. For each of the 13 (nominal +12 faults) fault Scenario, 100 Monte Carlo (MC) simulations are carried out. During each MC simulation, fault free elements are tolerated within 5% around its nominal values. The number of features will be 32. So, for each fault Scenario, simulated data are collected in a matrix (1300*32) each node.

Feature extraction

Time-domain: The total number of patterns obtained for each of the selected nodes, are equal to the number of faults multiplied by the number of MC simulations to be 1300 patterns. Each pattern will have 32 features each node. Calculated features are collected in a matrix (1300*32) that's average matrix (13*32) each node.

Frequency-domain: For each pattern (32 samples) in time domain, 16 Fourier Transform (FT) coefficients are computed. It represents the DC and the magnitude of the

first 15 harmonics. The second half of the FT coefficients is redundant. Calculated features are collected in a matrix (1300x16) that's average matrix (13*16) each node.

Classification

Here, the Euclidean distance K-mean algorithm as a simple classifier too (based on the Euclidean distance calculations) and the PNN as a more accurate classifier was used.

K-mean algorithm classification on Time-domain features

The execution result of classification by applying the K-mean algorithm directly on time samples is a classification matrix (13*13). The total CA worst and best cases allowing nodes' combinations are seen in Table 8. Node 3 is the best test node in the time-domain tests was observed. From the classification matrix of node 3, concluded that: 1) Fault scenarios 0,3,4,5,8,10,11 can be considered to be uniquely identified. 2) Fault scenarios (1,12), (2,9), and (6,7) can be collected in 3 ambiguity groups. Taking in considerations, the classification CA will rise to 99.2%.

Case	Single node	CA%	Double-node	CA%	Triple-node	CA%
Worst	6	71.3846	6,8	73	2,6,8	73
Best	3	78.8462	2,3	79.0769	2,3,4	77.1538

Table 8: Total worst and best CAs after classification in time-domain feature.

K-mean Algorithm classification on Frequency-domain features

The execution result of classification by applying the K-mean algorithm on the samples after calculating FFT is a

classification matrix (13*13). The total CA worst and best cases allowing nodes' combinations are seen in Table 9.

The node 8 is the best test node in frequency-domain tests was observed.

By selecting some features that's most effective to reduce the features from 16 to only 7 features, the dc component and 6 harmonic components having maximum CA at node 8 is 76.9231% for set of features {1,3,4,5,6,7,8} was taken. From classification matrix of node

8, concluded that: 1) Fault scenarios 0,3,4,5,8,10,11 can be considered to be uniquely identified. 2) Fault scenarios (1,12), (2,9), and (6,7) can be collected in 3 ambiguity groups. Taking these into considerations, the classification CA will rise to 99.85%.

Case	Single node	CA%	Double-node	CA%	Triple-Node	CA%
Worst	3	69.3846	3,4-3,6	69.3846	3,4,6 - 3,4,8	69.3846
Best	8	76.7692	6,8	76	4,6,8	75.3077

Table 9: Total worst and best CAs after frequency-domain feature extraction and classification.

Artificial neural network classification on frequency-domain features

Since obtained results using K-mean classifier, are poor, PNN as a classifier was used. Collected 1300 patterns are divided into two parts, 650 patterns used as training data and 650 used as testing data.

The result of classification by applying the PNN on the samples after calculating FFT is train classification matrix and test classification matrix of (13*13). Different values of NN parameters (spread) are tried to get best CA. The total CAs allowing best

single node seen in the past section (single node 8) are printed depending on spread values with some random permutations shown Table 10.

The highest CA at single node 8 is 86.30 % (at spread 0.1) was observed. By selecting classes from classification matrix of K-mean algorithm (at frequency domain) for the single node 8, the ambiguity groups are (1,12), (2,9), and (6,7) can be collected in 3 ambiguity groups. Taking these into considerations, the CA will rise to 99.90% (at spread 1.1) Table 11. The difference after selecting class to before is plotted in Figure 17.

Spread	Single node 8 CA %	
	Training	Total
1	84.3	79.61
0.9	87.69	81.61
0.7	89.07	81.61
0.5	91.84	82.15
0.3	95.38	83.46
0.1	99.84	86.3
0.09	100	84.46
0.07	100	84.35

Table 10: The training and total CAs using PNN classifier in frequency-domain (with spreads).

Spread	Single node 8 CA %	
	Training	Total
1.2	100	99.7
1.1	100	99.9
1	100	99.84
0.9	100	99.38
0.7	100	98.84
0.5	100	97.61
0.3	100	97
0.1	100	96

Table 11: The training and total CAs using PNN classifier in frequency-domain (with spreads) after selecting classes according to ambiguity groups.

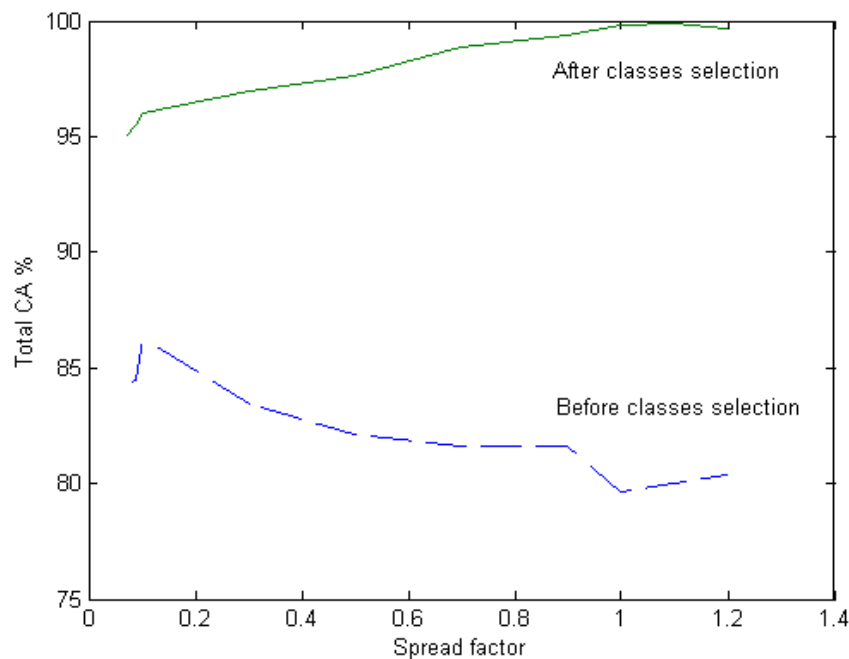


Figure 17: The difference between the total CA before and after classes' selections.

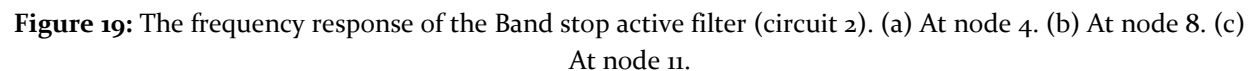
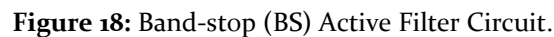
Example 2: Band-stop (bs) active filter circuit

Circuit is shown in Figure 18. The outputs at nodes 4,8, and 11 are respectively the low-pass, high-pass, and band-stop filters' outputs. The frequency response of the circuit is seen in Figure 19. It consists of 3 op-amps, 12 resistors and 4 capacitors.

Single excitation

Simulation

The CUT is simulated at fault-free and different faulty conditions using PSPICE simulator. All circuit capacitors are nominally 10nf and are assumed to be fault free for simplicity.



greater than 4 KHz. The frequency selected 200KHz which means a sampling period of 5 μ s is simulated during an interval of 0.155ms (which means each waveform is simulated by 32 points). For each of the 25 (nominal +24 faults) fault Scenario, 100 MC simulations are carried out. During each MC simulation, fault free elements are tolerated within 5% around its nominal values. The number of features will be 32. So, for each fault Scenario, simulated data are collected in a matrix (2500*32) each node.

Element	Nominal value	Fault class/ value	Fault index	Fault class/ value	Fault index
R1	15.8 K Ω	S.C. (1 Ω)	1	O.C. (100M Ω)	13
R2	15.8 K Ω	S.C.	2	O.C.	14
R3	10 K Ω	S.C.	3	O.C.	15
R4	6.56 K Ω	S.C.	4	O.C.	16
R5	31.8 K Ω	S.C.	5	O.C.	17
R6	31.8K Ω	S.C.	6	O.C.	18
R7	10 K Ω	S.C.	7	O.C.	19
R8	5.65 K Ω	S.C.	8	O.C.	20
R9	10 K Ω	S.C.	9	O.C.	21
R10	10 K Ω	S.C.	10	O.C.	22
R11	10 K Ω	S.C.	11	O.C.	23
R12	1 K Ω	S.C.	12	O.C.	24

Table 12: The selected hard faults and their description.

Feature extraction

Frequency-domain

The total number of patterns obtained for each of the selected nodes, are equal to the number of faults multiplied by the number of MC simulations to be 2500 patterns. Each pattern will have 32 features each node. For each pattern (32 samples) in time domain, 16 Fourier Transform (FT) coefficients are computed.

It represents the DC and the magnitude of the first 15 harmonics. The second half of the FT coefficients is redundant. Calculated features are collected in a matrix (2500x16) that's average matrix (25*16) each node.

Classification

Here, the Euclidean distance K-mean algorithm as a simple classifier (based on the Euclidean distance calculations) and the PNN as a more accurate classifier was used.

K-mean Algorithm Classification on Frequency-domain Features

The execution result of classification by applying the K-mean algorithm on the samples after calculating FFT is a classification matrix (25*25). The total CA worst and best cases allowing nodes' combinations are seen in Table 13. The combined node 8,11 give the best test node in frequency-domain tests by this classifier was observed.

Case	Single node	CA%	Dual-Node	CA%	Triple-Nodes	CA%
Worst	4	28.52	4,11	50.24	4,8,11	59.16
Best	8	47.56	8,11	52.52		

Table 13: Total worst and best CAs after K-mean algorithm classification on frequency-domain.

By selecting some features that's most effective to reduce the features to be less than 16, the max CA decreases a bit. Anyway, all achieved CAs are very poor. From classification matrix of node 8, concluded that: 1) Fault scenarios 0,1,2,5,6,10,12,13,14,17,18,21,22,23 and 24 can be considered to be uniquely identified. 2) Fault scenarios (3,16), (4,15), (7,20), (8,19) and (9,11) can be collected in 5 ambiguity groups. The CA will not raise much compared to past.

Artificial Neural Network Classification on Frequency-domain Features

Since obtained results using K-mean classifier, are rather poor, PNN as a classifier was used. Collected patterns (2500) patterns are divided into two parts, 1250 patterns used as training data and 1250 patterns are used as

testing data. The execution result of classification by applying the PNN on the samples after calculating FFT is train classification matrix and test classification matrix of (25*25). Different values of NN parameters (spread) are tried to get best CA. The total CAs allowing best single and combined nodes seen in the past section (single node 8, dual nodes 8,11) are printed depending on spread values with some random permutations shown Table 14. By selecting classes from classification matrix of K-mean algorithm (at frequency domain) for the single node 8, the ambiguity groups are (3,16), (4,15), (7,20), (8,19) and (9,11) can be collected in 5 ambiguity groups. If we take these to considerations, the CA will not raise much too.

Spread	Single node 8 CA%		Dual nodes 8,11 CA	
	Training	Total	Training	Total
1	51.92	50.52	63.12	60.64
0.5	55.2	52.32	67.04	62.56
0.3	61.12	55.48	74.32	65.08
0.1	72.48	60.04	87.6	70.88
0.05	83.6	64.96	93.2	71.28
0.04	83.2	65.04	93.64	70.48
0.03	85.68	65.52	95.6	70.24
0.02	86.16	64.2	96.56	65.76
0.01	87.28	61.44	98.72	60.32

Table 14: The training and total CAs using PNN classifier in frequency-domain.

Multiple excitations

Simulation

The CUT is simulated at fault-free and different faulty conditions using PSPICE simulator. All circuit capacitors are nominally

10nf and are assumed to be fault free for simplicity. The method applied to select the test stimuli frequencies is the sensitivity calculations. The selected test stimuli are three sinusoidal signals ac 500 Hz, 800 Hz, 2000 Hz with input resistances of 50 Ω . The selected test nodes are 4,8, and 11.

The nominal values of the circuit resistors and a list of 24 hard faults (12 faults with S.C. resistance and 12 with O.C. resistance) are shown in Table 15. The selected sampling frequency must be greater than 8000 Hz (8KHz). Selected frequency was 10KHz which means a sampling period of 0.1ms is simulated during an interval of 3.1ms (which means each waveform is simulated by 32 points). For each of the 25 (nominal +24 faults) fault Scenario, 100 MC simulations are carried out. During each MC simulation, fault free elements are tolerated within 5% around its nominal values. The number of features will be 32. So, for each fault Scenario, simulated data are collected in a matrix (2500*32) each node.

The sensitivity calculations are done in the following steps:

1. For each fault scenario, simulate the voltage at selected test node on a wide frequency range as matrix of faults patterns*frequency points.
2. Calculate the absolute voltage difference between each fault scenario and the nominal case.
3. Calculate the frequency band (min/max) for which the difference is greater or equal to 70% of the maximum of each absolute voltage difference.
4. Compare those sets of frequencies for all faults and select a frequency that is commonly found in all sets. Any uncommon set of frequencies requires us to select more frequencies. Those selected frequencies are the test frequencies that's most regarded.

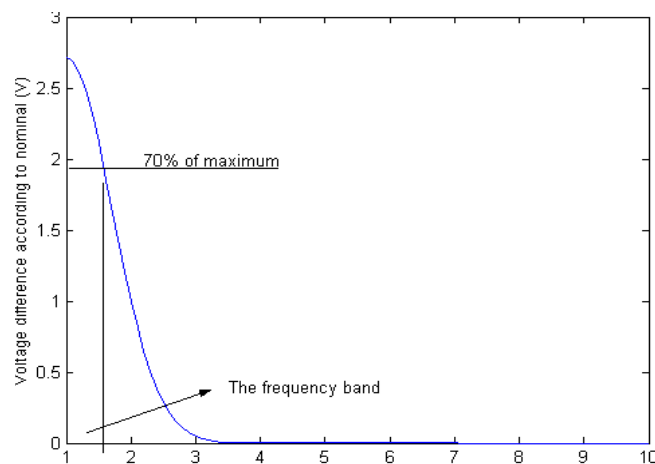


Figure 20: Sensitivity's selected set of frequencies.

Feature extraction

Frequency-domain

The total number of patterns obtained for each of the selected nodes, are equal to the number of faults multiplied by the number of

MC simulations to be 2500 patterns. Each pattern will have 32 features each node. For each pattern (32 samples) in time domain, 16 Fourier Transform (FT) coefficients are computed. It represents the DC and the magnitude of the first 15 harmonics. The

second half of the FT coefficients is redundant. Calculated features are collected in a matrix (2500x16) that's average matrix (25*16) each node.

Wavelet analysis

The wavelet analysis is recently utilized for feature extraction in fault diagnosis of analog circuits [33-35]. The total number of patterns obtained for each of the selected nodes, are equal to the number of nominal and faults multiplied by the number of MC simulations to be 2500 patterns. Each pattern will have 32 features each node. For each pattern (32 samples) in time domain, 4 Wavelet Transform (WT) coefficients are computed. It represents 4 approximate coefficients. Calculated features are collected in a matrix (2500x4) that's average matrix (25*4) each node.

Classification

Here, the Euclidean distance K-mean algorithm as a simple classifier based on the Euclidean distance calculations was used. Then, the PNN classifier which must give better performance of classification.

K-mean algorithm classification on frequency-domain features

The execution result of classification by applying the K-mean algorithm on the samples after calculating FFT is a classification matrix (25*25). The total CA worst and best cases allowing nodes' combinations are seen in Table 15. The combined node 4,11 give the best test node in frequency-domain tests by this classifier.

Case	Single node	CA%	Dual-Node	CA%	Triple-Nodes	CA%
Worst	8	48.5	4,8	50	4,8,11	67
Best	11	63.72	4,11	67		

Table 15: Total worst and best CAs after K-mean algorithm classification on frequency-domain.

By selecting some features that's most effective to reduce the features to be less than 16, the max CA at single-node is 62.20% (node 11 for features set {1:8}) and at dual-nodes is 64.12% (combined nodes 4, 11 for set of features {1,2,3,18,19,23,24}).

From classification matrix of node 11, concluded that: 1) Fault scenarios 0,1,2,5,16,17 and 18 can be uniquely identified. 2) Fault scenarios (3,16), (4,15), (6,13,14), (7,20), (8,19), (9,22), (10,21), (11,24) and (12,23) can be

collected in 9 ambiguity groups. The CA will raise to 99.46%.

Artificial neural network classification on frequency-domain features

Since obtained results using K-mean classifier, are rather poor, PNN as a classifier was used. Collected patterns (2500) patterns are divided into two parts, 1250 patterns used as training data and 1250 patterns are used as testing data. The execution result of classification by applying the PNN on the

samples after calculating FFT is train classification matrix and test classification matrix of (25*25). Different values of NN parameters (spread) are tried to get best CA. The total CAs allowing best single and

combined nodes seen in the past section (single node 11, dual nodes 4, 11) are printed depending on spread values with some random permutations. The CA for both single node and dual node is shown in Table 16.

Spread	Single node 11 CA %		Dual nodes 4, 11 CA %	
	Training	Total	Training	Total
1	72.08	67.28	75.68	70.2
0.7	74.72	68.4	78	70.48
0.3	81.6	71.4	88.8	75.88
0.1	90	74.2	98.24	78.8
0.09	95.44	76.88	99.2	80.32
0.07	97.76	77.4	99.33	80.25
0.05	99.04	77.12	99.4	78.4
0.03	99.52	76.80	100	75.92

Table 16: The train and total CAs using PNN classifier in frequency-domain.

By selecting classes from classification matrix of K-mean algorithm (for each of the single and dual nodes), the ambiguity groups can have three best classes. At the first groups for

instance, the ambiguity groups are (3,16), (4,15), (6,13, and 14), (7,20), (8,19), (9,22), (10,21), (11,24), and (12,23). They are shown in Table 17.

Element	V (11) has a triple group		V (11) has dual groups		V (4, 11)	
R1	1	13	1	13	1	13
R2	2	14	2	14	2	14
R3	3	15	3	15	3	15
R4	4	16	4	16	4	16
R5	5	17	5	17	5	17
R6	6	18	6	18	6	18
R7	7	19	7	19	7	19
R8	8	20	8	20	8	20
R9	9	21	9	21	9	21
R10	10	22	10	22	10	22
R11	11	23	11	23	11	23
R12	12	24	12	24	12	24

Table 17: Ambiguity groups of faults are seen in different classes' selections.

The cross-linked arrows show a normal ambiguity group. Practically, that's true where R3 at S.C. always has same

characteristic as with R4 at O.C all in the node 11 (and vice-versa) for instance. The total CAs allowing best single and combined nodes

(single node 11, dual nodes 4,11) after selecting classes are printed depending on spread

values with some random permutations. This is seen in Table 18.

Spread	Single node 11 with a triple ambiguity CA %		Single node 11 dual ambiguities CA %		Dual nodes 4, 11 CA %	
	Training	Total	Training	Total	Training	Total
1	99.2	98.96	95.04	94.96	99.94	99.48
0.9	99.36	99.36	95.36	95.32	99.68	99.4
0.8	99.44	99.32	95.6	95.36	99.84	99.56
0.7	99.84	99.28	95.28	95.28	99.92	99.64
0.6	99.6	99.28	95.76	95.32	100	99.56
0.5	100	99.24	95.92	95.08	100	99.44
0.4	100	99.12	95.84	95.2	100	99.28
0.3	99.84	99.84	96	94.8	100	98.76
0.2	100	98.52	96.96	95.48	100	98.56
0.1	100	96.68	99.84	95.08	100	96.76
0.08	100	95	100	94.24	100	95.52

Table 18: The training and total CAs using PNN classifier in frequency-domain (with spreads) after selecting classes according to ambiguity groups.

After selecting classes of ambiguity groups that the achieved CA is very rich. The single node with triple ambiguity gives the highest CA but a likely fault from the triple ambiguities will always be considered as three faults simultaneously. The dual node has a very high CA, but this requires more test time which is done in two nodes 4, and 11. Selected test to be done in the single node 11 that's triple ambiguity group which is applicable in the manufacturing lines with the CA of 98.84% (at spread 0.3).

Artificial neural network classification on wavelet features

Since PNN classification with wavelet features is most recently used due its higher accuracy and a smaller number of features. Collected patterns (2500) patterns are divided

into two parts, 1250 patterns used as training data and 1250 patterns are used as testing data. The result of classification by applying the PNN on the samples after calculating the discrete wavelet transform (DWT) is train classification matrix and test classification matrix of (25*25). Different values of NN parameters (spread) are tried to get best CA. The totals CAs depending on spread values with some random permutations in Table 19. By selecting classes from classification matrix of K-mean algorithm (for the single node 11), The ambiguity groups are (3,16), (4,15), (6,13,14), (7,20), (8,19), (9,22), (10,21), (11,24), and (12,23). The total CAs (at nodes 11) after selecting classes is printed depending on spread values with some random permutations shown in Table 20. Nearly values to past classification.

Spread	Single node 11 CA %		Dual nodes 4,11 CA %	
	Training	Total	Training	Total
1	60.16	59.52	64.88	64.08
0.7	62.64	59.8	66.24	64.12
0.3	64.96	61.56	70.56	66.68
0.1	70.88	64.64	76.88	70.08
0.07	72.24	65.88	79.36	70.72
0.05	73.76	66.52	84.24	72.96
0.03	79.68	69.68	91.04	75.76
0.01	92.64	75.16	98.24	78.08
0.009	93.2	73.36	98.88	77.56
0.007	94.96	73.8	99.52	76.64

Table 19: The training and total CAs using PNN classifier in frequency-domain (with spreads).

Spread	Single node 11 with triple ambiguity CA%	
	Training	Total
1	61.44	60.96
0.9	70.8	70.68
0.7	76.8	78.04
0.5	87.76	87.12
0.3	94.8	95
0.1	97.6	97.36
0.07	98.08	97.4
0.05	98.32	97.88
0.03	99.36	98.32
0.01	100	94.64

Table 20: The training and total CAs using PNN classifier in frequency-domain (with spreads) after selecting classes according to ambiguity groups.

Example 3: Sallen-key band-pass (bp) active filter circuit

Circuit has a stage of band-pass filter seen in Figure 21. The output is at node 5 which has the frequency response shown in Figure 22. It consists of one operational amplifier, five resistors and two capacitors.

Simulation

The CUT is simulated at fault-free and different faulty conditions using PSPICE simulator. The method applied to select the test stimuli frequencies is the sensitivity calculations.

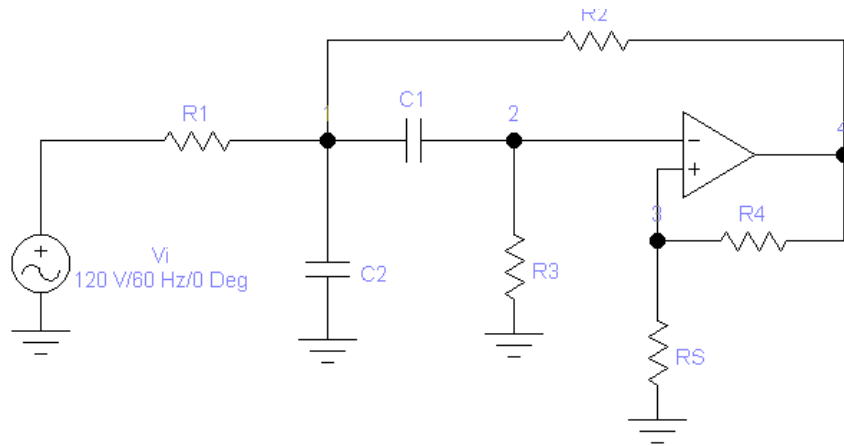


Figure 21: Sallen-Key Band-pass (BP) Active Filter Circuit.

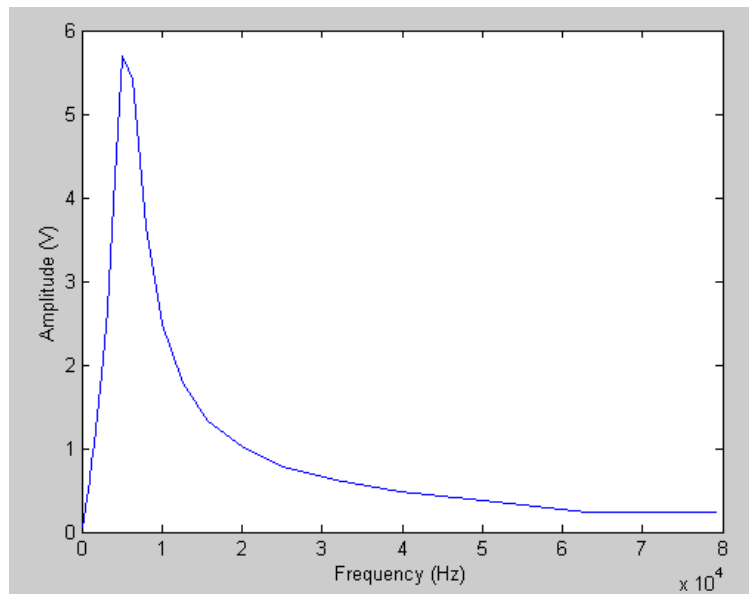


Figure 22: The frequency response of the circuit 1 at node 5.

The selected test stimuli are three sinusoidal signals ac 100Hz, 5KHz, and 94.5KHz with input resistances of 50. The selected test node is 5 (the output). The nominal values of the circuit elements and a list of 14 hard faults (7 faults with S.C. and 7 with O.C.) shown in Table 21. The selected sampling frequency must be greater than 378KHz. The frequency selected was 500KHz which means a sampling period of 2 us is simulated during an interval

of 0.098ms (which means each waveform is simulated by 50 points). For each of the 15 (nominal +14 faults) fault scenario, 100 MC simulations are carried out. During each MC simulation, fault free elements are tolerated within 5% for resistance and 10% for capacitance (around its nominal values).

The number of features will be 50. So, for each fault Scenario, simulated data are collected in

a matrix (1500*50) each node. The sensitivity calculations are done in the following steps:

1. Each fault scenario, simulate the voltage at selected test node on a wide frequency range as matrix of faults patterns*frequency points.
2. Calculate the absolute voltage difference between each fault scenario and the nominal case.
3. Calculate the frequency band (min/max) for which the difference is

greater or equal to 70% of the maximum of each absolute voltage difference.

4. Compare those sets of frequencies for all faults and select a frequency that is commonly found in all sets. Any uncommon set of frequencies requires us to select more frequencies. Those are the regarded test frequencies.

A sample of a set of selected set of frequencies is shown in Figure 22.

Element	Nominal value	Fault class/ value	Fault index	Fault class/ value	Fault index
R1	10 KΩ	S.C. (1Ω)	1	O.C (100 M)	8
R2	5 KΩ	S.C.	2	O.C.	9
R3	10 KΩ	S.C.	3	O.C.	10
R4	4 KΩ	S.C.	4	O.C.	11
R5	4 KΩ	S.C.	5	O.C.	12
C1	5nF	S.C. (100μ)	6	O.C.(1pF)	13
C2	5nF	S.C.	7	O.C.	14

Table 21: The selected hard faults and its description.

Feature extraction

Frequency-domain

The total number of patterns obtained for each of the selected nodes, are equal to the number of nominal and faults multiplied by the number of MC simulations to be 1500 patterns. Each pattern will have 50 features each node. For each pattern (50 samples) in time domain, 16 Fourier Transform (FT) coefficients are computed. It represents the DC and the magnitude of the first 15 harmonics. The second half of the FT coefficients is redundant. Calculated features are collected in a matrix (1500x16) that's average matrix (15*16) each node.

Wavelet analysis

The wavelet analysis is recently utilized for feature extraction in fault diagnosis of analog circuits [33-35]. The total number of patterns obtained for each of the selected nodes, are equal to the number of nominal and faults multiplied by the number of MC simulations to be 1500 patterns. Each pattern will have 50 features each node. For each pattern (50 samples) in time domain, 7 Wavelet Transform (WT) coefficients are computed. It represents 7 approximate coefficients. The second half of the WT coefficients is redundant. Calculated features are collected in a matrix (1500x7) that's average matrix (15*7) each node.

Classification

Here, the Euclidean distance K-mean algorithm as a simple classifier (based on the Euclidean distance calculations) and the PNN as a much more accurate classifier was used.

K-mean algorithm classification on frequency-domain features

The execution result of classification by applying the K-mean algorithm on the samples after calculating FFT is a classification matrix (15*15). The total CA is 78.20%. By selecting some features that's most effective to reduce the features to be less than 16, the max CA at the node 5 is 78.20% (for features set {1:5,7}). From classification matrix of node 5, it is observed that: 1) Fault scenarios 0,1,2,3,4 and 5 can be considered to be uniquely identified. 2) Fault scenarios (4,12) and (5,11) can be collected in 2

ambiguity groups. The CA will raise to 84.06%.

Artificial neural network classification on frequency-domain features

Since obtained results using K-mean classifier, are poor, PNN as a classifier was used. Collected patterns (1500) patterns are divided into two parts, 750 patterns used as training data and 750 patterns are used as testing data.

The execution result of classification by applying the PNN on the samples after calculating FFT is train classification matrix and test classification matrix of (15*15). Different values of NN parameters (spread) are tried to get best classification CA (CA). The total CAs is printed depending on spread values with some random permutations. Training & total CA depend on spread variations are shown in Table 22.

Spread	Node 5 CA%	
	Training	Total
1	79.86	79.53
0.9	79.33	79.8
0.7	80.13	79.73
0.5	80	80
0.3	80.9	81.4
0.1	80.4	77
0.09	80.53	77.8
0.07	81	77
0.05	81.2	76.65

Table 22: The training and total CAs using PNN classifier in freq-domain (with spread).

By selecting classes from CA of K-mean algorithm (for node5), the ambiguity groups considered are (4,12) and (5,11). The total CAs

(at nodes 5) after selecting classes is printed depending on spread values with some random permutations shown in Table 23.

Spread	Single node 5 CA%	
	Training	Total
1	85.6	85.8
0.9	85.73	86.07
0.7	85.33	85.33
0.5	86.67	86.4
0.3	87.73	87.93
0.1	87.33	84.73
0.09	87.33	83.93
0.07	87.6	83.2
0.05	87.73	81.67
0.03	94.8	87.2
0.02	95.6	87.2
0.01	96.13	87.46
0.009	96.13	86.86
0.008	96.4	87.07
0.001	99.07	84.33
0.0005	99.9	78.46

Table 23: The training and total CAs using PNN classifier in frequency-domain (with spreads) after selecting classes according to ambiguity groups.

Artificial neural network classification on wavelet features

Since PNN classification with frequency-domain features gives not high CA, the wavelet features were used. Collected patterns (1500) patterns are divided into two parts, 750 patterns used as training data and 750 patterns are used as testing data. The execution result of classification by applying the PNN on the samples after calculating the discrete wavelet transform (DWT) is train classification matrix and test classification matrix of (15*15).

Different values of NN parameters (spread) are tried to get best classification CA (CA).

The total CAs is printed depending on spread values with some random permutations shown in Table 24. By selecting classes from CA of K-mean algorithm (for node5), the ambiguity groups considered are (4,12) and (5,11).

The total CAs (at nodes 5) after selecting classes is printed depending on spread values with some random permutations shown in Table 25.

Spread	Single node 5 CA%	
	Training	Total
1	78.13	77.8
0.9	77.46	77.4
0.7	78	77.6
0.5	77.33	77.53
0.3	78.8	78.2
0.1	80.67	79.07
0.09	81.07	79
0.08	80.26	79.13
0.07	81.2	78.73
0.01	82.5	74

Table 24: he training and total CAs using PNN classifier in frequency-domain (with spreads).

Spread	Single node 5 CA%	
	Training	Total
0.5	91.3	90.93
0.4	91.73	91.6
0.3	91.6	91.67
0.2	92.26	91.93
0.18	92.8	92.3
0.1	94.13	91.86
0.09	93.6	92.2
0.08	94.67	92.2
0.07	94.26	92.6
0.06	94.53	92.2
0.05	94	91.33
0.01	93.86	87.73

Table 25: The training and total CAs using PNN classifier in frequency-domain (with spreads) after selecting classes according to ambiguity groups.

The wavelet classification used gives much more CA compared to the past frequency-domain FT method with more less features extracted.

Conclusion

Based on finding of the current study, a very high diagnostic degree of fault diagnosis was achieved. Despite of the excellent results,

there is a need for high off-line computational requirements. In addition, it is realized that degradation in robustness due to the deviation of non-faulty elements inside the tolerance region. By applying artificial neural networks which are basically mathematical models that emulate observed properties of biological nervous systems, the data was analyzed and recognize patterns within the

data, in a different way than traditional computing methods. Furthermore, the probabilistic neural network classification with wavelets features was used.

The execution result of classification has raised the accuracy dramatically and with less features being extracted, compared to the frequency-domain methods.

References

1. Bandler JW, Salama AE. Fault diagnosis of analog circuits. *Proceedings of the IEEE*. 1985;73(8):1279-325. [CrossRef](#)
2. Liu RW. A circuit theoretic approach to analog fault diagnosis. In *Testing and Diagnosis of Analog Circuits and Systems* 1991:1-35. Springer, Boston, MA. [CrossRef](#)
3. Wang Z, Gielen G, Sansen W. A novel method for the fault detection of analog integrated circuits. In 1994 IEEE. ISCAS. 1994;1:347-350.
4. Slamani M, Kaminska B. Fault observability analysis of analog circuits in frequency domain. *IEEE transactions on circuits and systems II: Analog Integr Circuits Signal Process*. 1996;43(2):134-9. [CrossRef](#)
5. Robotycki A, Zielonko R. Some models and methods for fault diagnosis of analog piecewise linear circuits via verification technique. In *Proceedings of the 17th IEEE Instrumentation and Measurement Technology Conference [Cat. No. 00CH37066]* 2000;2:1050-1055. [CrossRef](#)
6. Farina A. Merriam-Webster, Webster's Ninth New Collegiate Dictionary (Book Review). *International Social Science Review*. 1990;65(1):45. [CrossRef](#)
7. <http://courses.cs.tamu.edu/cpsc680/walker/Slides>
8. L.Tung , R. Saeks, " An Expirement in fault prediction." 1977:249.
9. An Approach To Built-in Testing," *IEEE Aerospace Electronic Systems*. 1978;14.
10. Liberty SR, Tung L, Saeks R. Fault prediction—Toward a mathematical theory. In *Rational Fault Analysis* 1977:135-142.
11. Trick TN, Li Y. A sensitivity based algorithm for fault isolation in analog circuits. In *Proc. IEEE Int. Symp. Circuits and Systems* 1983:1098-1101.
12. Fedi G, Giomi R, Luchetta A, Manetti S, Piccirilli MC. On the application of symbolic techniques to the multiple fault location in low testability analog circuits. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*. 1998;10:1383-8.
13. Liu RW. *Analog Fault Diagnosis of Large-Scale Electronic Circuits*. notre dame univ in dept of electrical engineering.1983.
14. Tadeusiewicz M, Korzybski M. A method for fault diagnosis in linear electronic circuits. *International Int. J. Circuit Theory Appl.*. 2000;28(3):245-62. [CrossRef](#)
15. Starzyk J, Bandler J. Multiport approach to multiple-fault location in analog circuits. *IEEE transactions on circuits and systems*. 1983;30(10):762-5. [CrossRef](#)
16. Varghese KC, Williams JH, Towill DR. Computer aided feature selection for enhanced analogue system fault location. *Pattern Recog*. 1978;10(4):265-80. [CrossRef](#)
17. Becker PW, Thamdrup JE. Pattern recognition applied to automated testing. In *The Automation of Testing*. Inst Elect Eng Conf Publ. 1972:82-86.
18. <http://cadlab.ece.ucsb.edu>
19. Roadstrum WH, Wolaver DH. *Electrical engineering for all engineers*. John Wiley & Sons Incorporated.1994.
20. <http://power.ece.drexel.edu/coursesweb/ECEP352/PSpiceIntorduction.pdf/>
21. <http://people.deas.harvard.edu/~jones/es154/pages/PSpice/PSpice.html>
22. <http://www.coe.uncc.edu/~jcarter/pspice/>
23. Simulink," *MATLAB Help*," Documentation, Version 6.1.
24. Gabel RA, Roberts RA. *Signals and linear systems*. John Wiley & Sons.1991.
25. www.ecs.syr.edu.
26. <http://www.statsoft.com/textbook/stneunet.html#classification>

27. <http://www.dasc.dtic.mil/techs/neural/neural-toc>
28. <http://www.statsoft.com/textbook/stneunet.html#intro>
29. El-Yazeed MA, Mohsen AA. A preprocessor for analog circuit fault diagnosis based on Prony's method. AEU. Int J Electron Telecommun. 2003;57(1):16-22. [CrossRef](#)
30. Prasad VC, Babu NS. Selection of test nodes for analog fault diagnosis in dictionary approach. IEEE Trans Instrum Meas 2000;49(6):1289-97. [CrossRef](#)
31. Sedra AS, Smith KC, Carusone TC, Gaudet V. Microelectronic circuits. New York: Oxford university press. 2004.
32. Aminian F, Aminian M, Collins HW. Analog fault diagnosis of actual circuits using neural networks. IEEE Trans. Instrum Meas 2002;51(3):544-50. [CrossRef](#)
33. Aminian F, Aminian M. Fault diagnosis of nonlinear analog circuits using neural networks with wavelet and Fourier transforms as preprocessors. J Electron Test. 2001;17(6):471-81.
34. Aminian F, Aminian M. Fault diagnosis of analog circuits using Bayesian neural networks with wavelet transform as preprocessor. J Electron Test. 2001;17(1):29-36.