DIRECT AND INDIRECT CLASSIFICATION OF HIGH FREQUENCY LNA GAIN PERFORMANCE – A COMPARISON BETWEEN SVMS AND MLPS

Peter C. Hung, Seán F. McLoone, Ronan Farrell

Institute of Microelectronics and Wireless Systems, Department of Electronic Engineering, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland,
{phung, sean.mcloone, rfarrell}@eeng.nuim.ie, http://imws.eeng.nuim.ie/

Abstract: *The task of determining low noise amplifier (LNA) high-frequency performance in functional testing is as challenging as designing the circuit itself due to the difficulties associated with bringing high frequency signals off-chip. One possible strategy for circumventing these difficulties is to inferentially estimate the high frequency performance measures from measurements taken at lower, more accessible, frequencies. This paper investigates the effectiveness of this strategy for classifying the high frequency gain of the amplifier, a key LNA performance parameter. An indirect Multilayer Perceptron (MLP) and direct support vector machine (SVM) classification strategy are considered. Extensive Monte-Carlo simulations show promising results with both methods, with the indirect MLP classifiers marginally outperforming SVMs.*

Keywords: *LNA, Functional testing, Classification, Support Vector Machines, Multilayer Perceptrons.*

1. INTRODUCTION

Functional testing of radio frequency integrated circuits (RFIC) is becoming an increasingly difficult problem for IC manufacturers, especially for RF components that operate in the multi-gigahertz range. Two main problems exist when working at these frequencies: relaying the high-frequency signals to and from the automatic test equipment (ATE) without affecting the performance of the circuits under test; building RF production testers that are not prohibitively expensive. While advances in technology and market requirements have seen rapid growth in high-frequency and high integration RFIC designs, testing practice has not followed suit. Indeed, reliable high-frequency testing has become the dominant factor in the cost and time-to-market of novel wireless products [1]. Consequently, developing cost-efficient testing solutions is becoming an increasingly important research topic [2-4].

Some of the proposed schemes for RFIC testing are based on an end-to-end strategy in which the output of the transmitter and the input of the receiver are linked through a loop-back connection. In this configuration, the testing of the complete system is carried out without any external stimulus by employing the on-chip digital hardware available. Unfortunately, this solution is not always applicable to all kinds of RF components. Other recent proposals for RF system testing have focused on the development of methodologies and algorithms for automated test, and Design for Testability (DfT). One approach, Built-In-Test (BIT) uses additional circuitry that allows high frequency tests to be performed on-chip and then evaluated using lower frequency or DC external testers [3-5]. An example of this approach by Bhattacharya and Chatterjee used a CMOS RMS detector for testing wireless transceivers [6, 7]. By employing a three-stage rectifier, they converted the high frequency circuit output into a DC signal which was then correlated with the equivalent test specification values of the circuit-under-test (CUT). In this manner, the actual performance of the system, under certain conditions, could be estimated. However, when considering BIT testing, issues such as the on-chip area and power consumption required for the additional embedded circuitry can add significantly to the cost of the design.

In this paper a different approach is considered. As many RFICs show strong correlation between their responses to circuit parameter variation at different frequencies, it is hypothesised that knowledge of the responses at lower frequencies may provide sufficient information to allow classification of responses at higher frequencies.

To investigate this hypothesis, the testing requirement of a low noise amplifier (LNA) is used as a case study. LNAs are key components in modern wireless systems as they provide the initial amplification for weak signals received from antennae. For our study, a standard LNA design utilising United Microelectronics Corporation's (UMC) 0.18 μm silicon process technology and designed to operate at 2.4 GHz was selected [8]. The LNA circuit consisted of 2 bias transistors (0.18 μm channel length), 4 RF transistors (0.5 μm channel width), 4 resistors, 3 capacitors and 4 inductors and was deemed to be functioning correctly if the value of S21 @ 2.4 GHz was in the range 14.7 dB to 17.2 dB and faulty otherwise. S21 is the forward transmission coefficient or forward voltage gain (hereafter referred to as 'gain') of the amplifier and is a critical RF circuit performance measure. The '@' notation is used to indicate the circuit excitation frequency at which the measurement is taken.

Random circuit parameter perturbations representative of typical manufacturing process variations were generated and the value of S21 recorded at different frequencies. Fig.1 shows a plot of the correlation that exists between variations in gain (S21 @ 2.4 GHz) and variations in the same parameter computed at other frequencies. There is a strong, but decreasing, correlation evident as the circuit excitation frequency moves away from the operating frequency.

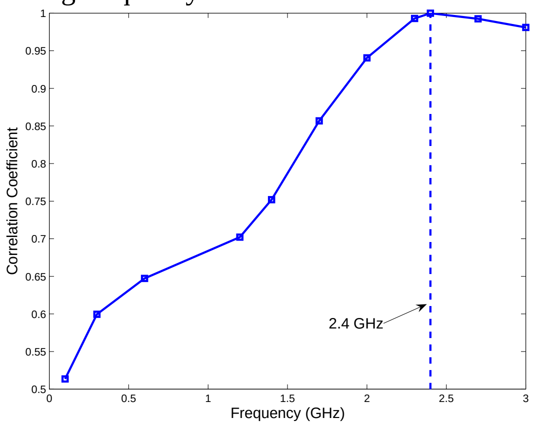

Fig. 1 – Correlation between variations in S_{21} computed at different frequencies and the variations in S_{21} computed at the target frequency (2.4 GHz).

Even when the correlation is high, circuit performance classification on the basis of these lower frequencies is not straightforward, as demonstrated in Fig.2. This shows the relationship between S21 @ 2.4 GHz and the values computed at: (a) 2.0 GHz and (b) 0.1 GHz, respectively, and highlights the fact that even when the correlation is greater than 90% it is not possible to discriminate between the 'good' and 'bad' circuits effectively. In fact, simple thresholding on the basis of S21 @ 2.0 GHz leads to a misclassification rate of greater than 20%. The misclassification rate increases rapidly as the frequency is reduced and reaches 43.6% for S21 @ 0.1 GHz. The rapid deterioration in performance is a result of the localised influence of some parameter variations and the complex nonlinear interaction between circuit components.

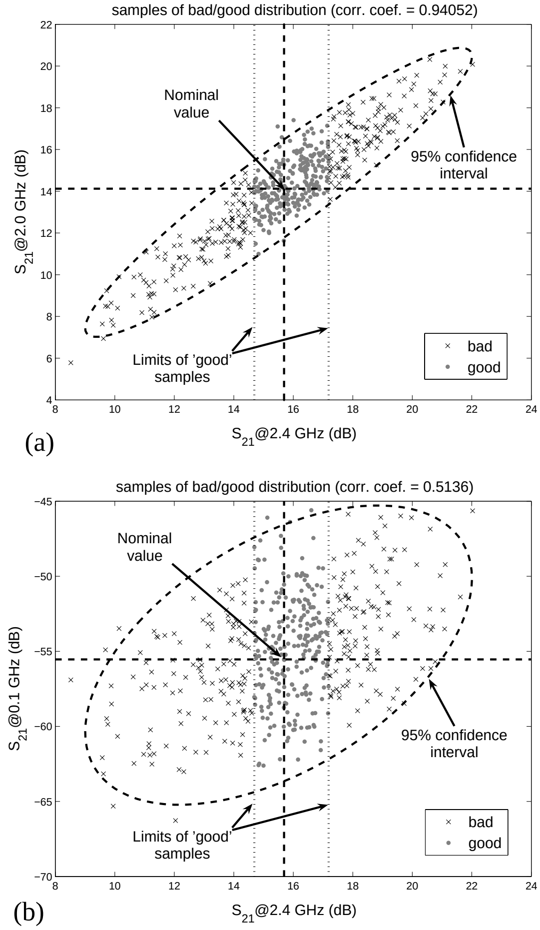

Fig. 2 – S_{21} parameter relationships for the 2.4 GHz LNA model used in circuit simulations: (a) S_{21} @ 2.0 GHz and (b) S_{21} @ 0.1 GHz plotted against S_{21} at the operating frequency.

Since the shape of the frequency response of an LNA is a deterministic nonlinear function of its component parameters, better classification performance can be expected if the information from several low-frequency measurements can be combined. To that end, this paper considers the possibility of classifying circuit performance using machine learning techniques. Two strategies are investigated. In the first, an Artificial Neural Network (ANN) is trained to predict the value of S_{21} @ 2.4 GHz from the values measured at other frequencies and then a thresholding rule is applied to this prediction to perform the circuit classification, while in the second, a Support Vector Machine (SVM) is trained to directly classify circuit performance on the basis of the low-frequency S_{21} measurements. These two machine learning techniques and the LNA classification methodology are introduced in Section 2. The simulation study is then described in Section 3 followed by the results in Section 4. Finally, the conclusions of the study are presented in Section 5.

2. MACHINE LEARNING LNA PERFORMANCE CLASSIFICATION

Defining the set of N low-frequency S_{21} measurements of the i^{th} LNA circuit as the feature vector $\mathbf{x}_i$ (row vector) and the corresponding class label y_i, with $y_i = +1$ indicating 'good' and $y_i = -1$ indicating 'bad', we can generate a set of L training data examples,

$$(\mathbf{X}, \mathbf{y}) = (\mathbf{x}_1, y_1), \ldots (\mathbf{x}_i, y_i), \ldots (\mathbf{x}_L, y_L) \in \Re^N, \quad (1)$$

with which to train an LNA circuit classifier to estimate a decision function $f(\mathbf{x})$

$$f(\mathbf{x}) : \Re^N \to \{\pm 1\}, \quad (2)$$

that can then be used to classify new circuits. Here, 'good' and 'bad' are determined by a threshold function, z_{th} applied to S_{21} @ 2.4 GHz, that is

$$z_{\text{th}}(x) = \begin{cases} +1 & \text{if} \quad 14.7 < x < 17.2 \\ -1 & \text{otherwise} \end{cases} \quad (3)$$

The decision function in (2) can be estimated directly from the training data by using, for example, a SVM classifier. Alternatively it can be estimated indirectly by first predicting the value of S_{21} @ 2.4 GHz from the feature vector,

$$g(\mathbf{x}) \to \text{S}_{21}\,@\,2.4\text{GHz}, \quad (4)$$

and then using a threshold function to perform the classification, that is

$$f(\mathbf{x}) : z_{\text{th}}(g(\mathbf{x})) \to \{\pm 1\}, \quad (5)$$

The difference between the two implementations is illustrated graphically in Fig.3.

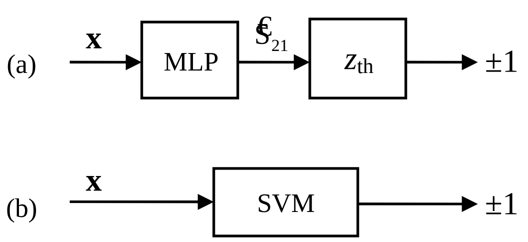

Fig. 3 – Two classifier implementations: (a) indirect MLP; and (b) direct SVM.

An artificial neural network such as a Multilayer Perceptron (MLP) can be used to learn the nonlinear mapping represented by (4).

2.1 SUPPORT VECTOR MACHINES

SVMs, first proposed by Vladimir Vapnik in 1963 [9], are a supervised linear learning technique widely used for classification problems. They are known to perform binary classification well in many practical applications. Consider a separating hyperplane that divides two classes of data:

$$\mathbf{w} \cdot \mathbf{x} - b = 0, \quad \mathbf{w} \in \Re^N, b \in \Re, \quad (6)$$

where $\mathbf{w}$ and b are unknown coefficients, and two additional hyperplanes that are parallel to the separating hyperplane:

$$\begin{aligned} \mathbf{w} \cdot \mathbf{x} - b &= 1 \\ \mathbf{w} \cdot \mathbf{x} - b &= -1 \end{aligned} \quad (7)$$

Defining the margin as the perpendicular distance between the parallel hyperplanes, the optimal hyperplane is the one which results in the maximum margin of separation between the two classes.

Mathematically the problem can be expressed as

$$\max_{\mathbf{w}} \frac{2}{|\mathbf{w}|} \equiv \min_{\mathbf{w}} (\mathbf{w} \cdot \mathbf{w}), \quad (8)$$

subject to

$$y_i(\mathbf{w} \cdot \mathbf{x}_i - b) \geq 1, \quad i = 1, 2, \ldots, L. \quad (9)$$

This is a constrained quadratic optimisation problem whose solution w has an expansion [10]

$$\mathbf{w} = \sum_i v_i \mathbf{x}_i \quad (10)$$

where $\mathbf{x}_i$ are the subset of the training data, referred to as support vectors, located on the parallel hyperplanes, and v_i are the corresponding weighting factors. The linear SVM (LSVM) decision function is then given by

$$f_{\text{LSVM}}(\mathbf{x}) = z_{\text{SVM}}(\mathbf{w} \cdot \mathbf{x} - b), \quad (11)$$

where z_{SVM} is defined as

$$z_{\text{SVM}}(\phi) = \begin{cases} +1 & \text{if } \phi \geq 0 \\ -1 & \text{if } \phi < 0 \end{cases} \quad (12)$$

The decision function in (11) can be rewritten as

$$f_{\mathrm{LSVM}}(\mathbf{x}) = z_{\mathrm{SVM}}\left(\sum_i v_i(\mathbf{x}\cdot\mathbf{x}_i) - b\right), \qquad (13)$$

with the result that it is only dependent on dot products between the test data vector, **x**, and the support vectors. This important property allows SVMs to be extended to problems where nonlinear partitions of data sets are required. This is achieved by replacing the dot products by a kernel function $k(.)$ which meets the Mercer's condition [11]:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)\cdot\Phi(\mathbf{x}_j). \qquad (14)$$

This maps the data into a higher dimension feature space where linear SVM classification can be performed. Note the resulting decision function, in the original data space, will be nonlinear and takes the form

$$f_{\mathrm{SVM}}(\mathbf{x}) = z_{\mathrm{SVM}}\left(\sum_i^L v_i.k(\mathbf{x}, \mathbf{x}_i) - b\right). \qquad (15)$$

In non-separable problems where different classes of data overlap, slack variables can be introduced so that a certain amount of training error or data residing within the margin is permitted. This gives rise to a 'soft margin' optimisation function [11, 12]. To give users the ability to adjust the amount of training error allowed in the optimisation, a smoothing parameter C is incorporated into the soft margin function, with a larger C corresponding to assigning a larger penalty to errors.

The Gaussian radial basis function (RBF), defined as

$$k_{\mathrm{RBF}}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\left|\mathbf{x}_i - \mathbf{x}_j\right|^2}{2\sigma^2}\right) \qquad (16)$$

is a popular choice of SVM kernel and the one selected for this application. The parameter, σ, controls the width of the kernel and is determined as part of the classifier training process.

2.2 ARTIFICIAL NEURAL NETWORKS

Neural networks [13] are one of the best known and most commonly-used machine learning techniques. There are various configurations and structures of NNs, but all contain an array of neurons that are linked together, usually in multiple layers. In this application a single hidden layer Multilayer Perceptron (MLP) topology is chosen because of its universal function approximation capabilities, good generalisation properties and the availability of robust efficient training algorithms [14].

The output of a single hidden layer MLP can be written as a linear combination of sigmoid functions (i.e. neurons),

$$g(\mathbf{x}, \mathbf{w}_{\mathrm{NN}}) = b^h + \sum_i w_i^h sig_i(\mathbf{x}), \qquad (17)$$

where

$$sig_i(\mathbf{x}) = \frac{1}{1 + \exp(\mathbf{w}_i^u \cdot \mathbf{x} + b_i^u)}. \qquad (18)$$

Here, w_i^h, $\mathbf{w}_i^u$, b_i^u ,($i = 1,2,\cdots,M$) and b^h are weights and biases which collectively form the network weights vector, $\mathbf{w}_{\mathrm{NN}}$. Defining a Mean Squared Error (MSE) cost function over the training data

$$E(\mathbf{w}_{\mathrm{NN}}) = \frac{1}{L}\sum_{p=1}^{L}(g(\mathbf{x}_p, \mathbf{w}_{\mathrm{NN}}) - d_p)^2, \qquad (19)$$

with d_p corresponding to the desired network output for the p^{th} training pattern (i.e. S_{21} @ 2.4 GHz), the optimum weights can be determined using gradient based optimisation techniques.

3. SIMULATION STUDY

Circuit simulations were required to evaluate the potential for employing multiple low-frequency S_{21} measurements to classify LNA S_{21} performance at high frequency, and to compare the performance of the proposed machine learning classifiers. Hence, a Monte Carlo simulation study was undertaken using a 2.4 GHz LNA model implemented in Eldo® [8]. Uniform random variations were introduced into 38 of the model parameters and 10,000 circuit simulations were performed. These parameter variations were chosen to emulate typical industrial LNA manufacturing process variations, such as the lithographic dimensions that affect component resistance and capacitance. While in practice circuit parameters might be expected to vary normally around their nominal values, uniform distributions were chosen to give an even coverage of the LNA parameter space. Catastrophic failures, such as short-circuits, were not considered as these can be identified relatively easily using existing IC testing techniques.

For each circuit the S_{21} performance parameter was recorded at 0.1, 0.3, 0.6, 1.2, 1.4, 1.7 and 2.0 GHz and also at the operating frequency (2.4 GHz). This data was then normalised to have zero mean and unit variance and divided into training and test data sets, each containing 5,000 samples.

As the study focused on the viability of performance classification using low frequency S_{21} measurements, two different feature vectors were considered in the study, containing measurements from 0.1 to 1.4 GHz and 0.1 to 2.0 GHz, respectively, that is:

$$\mathbf{x}_{1.4} = [S_{21}^{0.1}, S_{21}^{0.3}, S_{21}^{0.6}, S_{21}^{1.2}, S_{21}^{1.4}], \qquad (20)$$

and

$$\mathbf{x}_{2.0} = [S_{21}^{0.1}, S_{21}^{0.3}, S_{21}^{0.6}, S_{21}^{1.2}, S_{21}^{1.4}, S_{21}^{1.7}, S_{21}^{2.0}]. \qquad (21)$$

Here, S_{21}^{f} denotes the value of S_{21} at f GHz. In each case the target MLP model output is $S_{21}^{2.4}$ while the target labels for the SVM classifier are given by $z_{\text{th}}(S_{21}^{2.4})$.

3.1 MLP TRAINING

MLP training was performed using the hybrid BFGS training algorithm [14] with stopped minimisation used to prevent over-fitting [15]. The optimum number of neurons (M) was determined for each model by systematically evaluating different network sizes and selecting the network with the minimum MSE on the test data set. Training was repeated ten times for each network size to allow for random weight initialisations and the best set of weights recorded in each case.

The optimum network sizes and resulting model fit, measured in terms of the correlation with the true value of $S_{21}^{2.4}$, are summarised in Table 1. For comparison purposes, the correlation between $S_{21}^{2.4}$ and the single frequency measurements at 1.4 and 2.0 GHz are also given.

Table 1. Optimum MLP classifier model dimensions and resulting model fit (correlation coefficient)

Feature Vector	Network Dimensions (input, hidden, output)	Model Fit
$\mathbf{x}_{1.4}$	MLP(5,12,1)	0.9364
$\mathbf{x}_{2.0}$	MLP(7,15,1)	0.9979
$S_{21}^{1.4}$	-	0.7537
$S_{21}^{2.0}$	-	0.9408

As expected, the exploitation of multiple frequencies results in much better predictability of $S_{21}^{2.4}$ than using the measurement at a single frequency. Notably, the information provided by $S_{21}^{2.0}$ is still marginally greater than the combined information provided by all measurements up to 1.4 GHz. The classification performance of these networks, when employed in the indirect LNA classifier scheme, will be reported in Section 4.

3.2 SVM TRAINING

SVM training was performed using two Matlab® packages: LibSVM v2.84 [16] and simpleSVM [17]. The kernel width parameter σ and smoothing parameter C were fine tuned in LibSVM using a five-fold cross validation strategy.

Initial SVM results were quite poor despite expectations of superior performance to indirect classification using MLPs (results presented in Section 4). It was determined that this might be due to the bimodal distribution of the out-of-specification circuits forming the 'bad' class, i.e. it consists of two segments separated by the 'good' class, as shown in Fig.4.

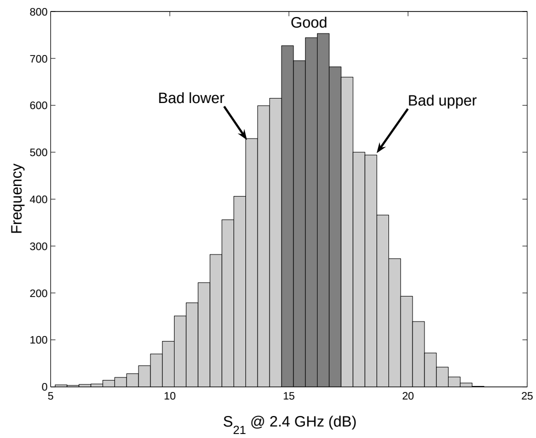

Fig. 4 – Histogram of 'good' and 'bad' circuits as defined by the S_{21} performance criteria.

The *a priori* knowledge of the distribution of the 'bad' circuits can be taken into account by splitting the 'bad' samples into 'bad lower' and 'bad upper' samples, thereby introducing 3 classes – 'bad lower', 'good' and 'bad upper'. SVM classification is then performed in two-stages as shown in Fig.5. Firstly, two binary SVMs are trained, one to classify LNAs as either 'bad lower' or 'not bad lower', and one to classify LNAs as either 'bad upper' or 'not bad upper'. Then the overall classification is obtained from a NOR logic combination of the individual decision functions, as shown in Table 2.

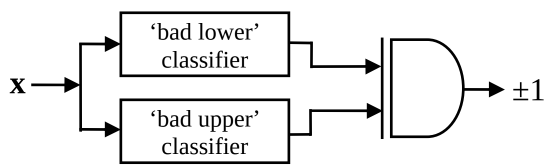

Fig. 5 – SVM classification for 3 classes.

This three-class SVM approach is denoted SVM3 while the original two-class SVM classifier will be referred to as SVM2.

Table 2. Logic table of two-stage SVM3 classification

Stage 1	Stage 2	Final decision
'bad lower' (S_{21} < 14.2 dB)?	**'bad upper' (S_{21} > 17.2 dB)?**	**Circuit 'good'/'bad'?**
no	no	'good'
yes	no	'bad'
no	yes	'bad'
yes	yes	assumed 'bad'*

* Though theoretically unattainable, in practice such situations arise due to incorrect SVM predictions.

4. RESULTS

The performance of the MLP, SVM2 and SVM3 LNA classifiers was measured in terms of the following metrics computed on the test data set:

- **GPR**: good pass rate – Percentage of 'good' LNAs correctly classified as 'good'
- **BFR**: bad fail rate – Percentage of 'bad' LNAs correctly classified as 'bad'
- **FR**: failure rate – Percentage of LNAs classified as 'good' that are in fact 'bad'
- **MCR**: misclassification rate – Percentage of LNAs incorrectly classified

To provide robust estimates, the metrics were computed by averaging over 100 batches of LNAs generated from the test data set using sampling with replacement. Each batch consisted of 500 'good' and 500 'bad' circuits randomly selected (with replacement) from a total of 1,795 'good' and 3,205 'bad' examples in the test pool.

Since the good pass rate (GPR) and bad fail rate (BFR) of a classifier vary as a function of the classification threshold, with one increasing as the other decreases, the threshold can be adjusted to control one or other of these metrics. This is achieved by adding an offset, α , to the decision function applied to the classifier variable. In the case of SVMs, for example, the decision function $z_{\text{SVM}}(\phi)$ in (12) becomes $z_{\text{SVM}}(\phi - \alpha)$.

Tables 3 and 4 show the mean performance of the classifiers at 90% and 75% BFR, respectively. Here, the classifier offset was adjusted to achieve the target BFRs. This approach was adopted as it reflects the importance in the electronics industry of controlling the number of faulty components released to the market. The result for classification on the basis of single frequency measurements at 1.4 and 2.0 GHz are also included for comparison.

Table 3. Mean performance of the MLP and SVM LNA classifiers at 90% BFR

Feature Vector	Method	GPR (%)	FR (%)	MCR (%)
$\mathbf{x}_{1.4}$	MLP	57.11	14.88	26.45
	SVM2S*	46.61	17.61	31.69
	SVM2	53.48	15.72	28.26
	SVM3S*	46.87	17.55	31.56
	SVM3	54.26	15.53	27.87
$\mathbf{x}_{2.0}$	MLP	99.70	9.09	5.15
	SVM2S*	92.49	9.74	8.75
	SVM2	94.29	9.57	7.85
	SVM3S*	92.32	9.75	8.84
	SVM3	94.10	9.59	7.95
$S_{21}^{1.4}$	-	19.98	33.29	45.01
$S_{21}^{2.0}$	-	55.48	15.24	27.27

* From simpleSVM toolbox.

Table 4. Mean performance of the MLP and SVM LNA classifiers at 75% BFR

Feature Vector	Method	GPR (%)	FR (%)	MCR (%)
$\mathbf{x}_{1.4}$	MLP	81.03	23.55	21.99
	SVM2S*	77.40	24.39	23.80
	SVM2	81.74	23.40	21.63
	SVM3S*	78.44	24.15	23.28
	SVM3	82.12	23.32	21.44
$\mathbf{x}_{2.0}$	MLP	100.00	19.97	12.50
	SVM2S*	98.81	20.17	13.10
	SVM2	99.13	20.13	12.94
	SVM3S*	98.79	20.17	13.10
	SVM3	99.12	20.13	12.94
$S_{21}^{1.4}$	-	47.13	34.63	38.94
$S_{21}^{2.0}$	-	84.65	22.78	20.17

* From simpleSVM toolbox.

Comparing the different implementations of SVM, it is found that simpleSVM's performance is slightly inferior to LibSVM. This is further highlighted in Fig.6, which shows the operating curves of the classifiers obtained with each SVM package for the SVM3 approach. While the difference is marginal when using the $\mathbf{x}_{2.0}$ feature vector, it is significant when using $\mathbf{x}_{1.4}$, which corresponds to the more demanding of the classification problems. It is not clear why this should be the case, but it is believed that the simpleSVM solution may not be precise when there is a large degree of overlap between classes.

It is noted that high GPRs are always accompanied by correspondingly low values of FR and MCR. As expected, the performance of all classifiers deteriorates when only the lower frequency S_{21} measurements ($\mathbf{x}_{1.4}$) are considered, though SVM3 and MLP still provide comparable performance to classification using a single S_{21} measurement at 2.0 GHz.

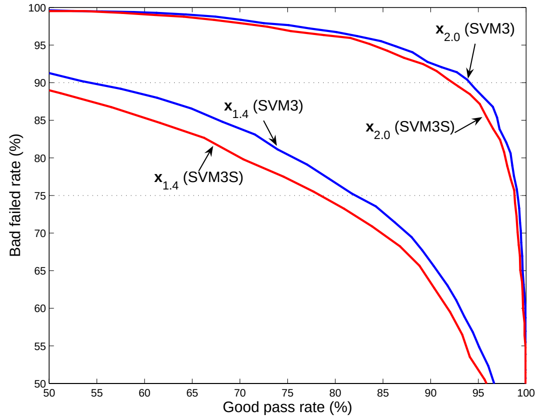

Fig. 6 – Operating curves (BFR vs. GPR) for $\mathbf{x}_{1.4}$ and $\mathbf{x}_{2.0}$ obtained with simpleSVM (SVM3S) and LibSVM (SVM3).

When comparing the MLP and SVM3 classifiers, it can be seen that the MLP gives the best results at almost all BFRs, as shown in Fig.7. When using $\mathbf{x}_{1.4}$ SVM3 is only marginally inferior to the MLP. However, the GPR gap increases to more than 5% when the higher frequency feature vector, $\mathbf{x}_{2.0}$, is considered at 90% BFR.

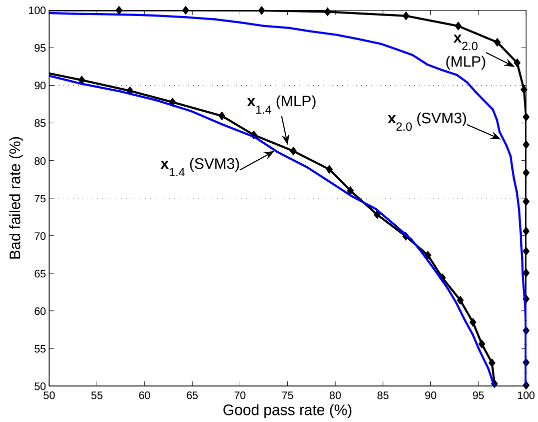

Fig. 7 – Operating curves (BFR vs. GPR) for MLP and SVM3 classifiers (for $\mathbf{x}_{1.4}$ and $\mathbf{x}_{2.0}$).

Interestingly, the direct SVM classifiers are not able to outperform the indirect MLP classifiers, even when the *a priori* knowledge of the bimodal distribution of the out-of-specification LNAs is taken into account. Indeed, the SVM2 and SVM3 classifiers give almost identical results indicating that SVMs can comfortably handle complex data distributions and that, therefore, exploitation of this knowledge is unnecessary. Consequently, it can be concluded that while, in general, SVMs may be the natural setting for classification tasks, the indirect classification approach can, in some instances, achieve better results. This may be a reflection of the fact that the indirect approach allows exploitation of more detailed information about the underlying mapping in its learning process.

In a manufacturing context the operating curves are a useful tool for assessing the trade-offs that can be obtained with each classifier. For example, for a target BFR of 90% a GPR of 55% is obtained with $\mathbf{x}_{1.4}$ as a feature vector. This increases to over 99% with $\mathbf{x}_{2.0}$. Alternatively, a 100% BFR can be obtained when using $\mathbf{x}_{2.0}$ if the GPR is dropped to 75%. The level of wastage at $\mathbf{x}_{1.4}$ may well be acceptable to manufacturers when the costs of generating the additional test signals needed for $\mathbf{x}_{2.0}$ are taken into account.

5. CONCLUSIONS

Functional testing of high-frequency LNAs is becoming a prohibitively expensive and time-consuming exercise, due to the difficulties with bringing such signals off-chip. This paper has investigated a novel testing strategy in which machine learning classifiers are used to predict high-frequency LNA gain performance by combining information from several lower frequency measurements. Results have been presented for both direct SVM and indirect MLP classifier implementations. These show that the proposed strategy has the potential to significantly extend the operating frequency range of existing ATE. For example, in the gain classification case study considered, the operating frequency range was extended by 20% at 2 GHz and 42% at 1.4 GHz.

Of the two classifier implementations considered the indirect MLP classifier yielded marginally superior performance and is the preferred approach for this application.

ACKNOWLEDGEMENTS

Research presented in this paper was funded by Enterprise Ireland Commercialisation Fund (EI CFTD/2003/304) under the National Development Plan. The authors gratefully acknowledge this support.

6. REFERENCES

[1] J. Ferrario, R. Wolf, H. Ding, Moving from mixed signal to RF test hardware development. *IEEE Int. Test Conference* (2001) pp. 948–956.

[2] W. Y. Lau, Measurement challenges for on-wafer RF-SOC test. *27th Annual IEEE/SEMI Int. Elect. Manufact. Tech. Symp.* (2002) pp. 353–359.

[3] M. Negreiros, L. Carro, A. Susin, Low cost on-line testing of RF circuits. *10th IEEE Int. On-Line Testing Symp.* (2004) pp. 73–78.

[4] D. C. Doskocil, Advanced RF built in test. *AUTOTESTCON '92 IEEE Sys. Readiness Tech. Conf.* (1992) pp. 213–217.

[5] M. E. Goff, C. A. Barratt, DC to 40 GHz MMIC power sensor. *Gallium Arsenide IC Symp.* (1990) pp. 105–108.

[6] S. Bhattacharya, A. Chatterjee, Use of embedded sensors for built-in-test RF circuits, *IEEE Int. Test Conf.* (2004) pp. 801 – 809.

[7] S. S. Akbay, A. Chatterjee, Feature extraction based built-in alternate test of RF components using a noise reference, *22nd IEEE VLSI Test Symp.* (2004) pp. 273 – 278.

[8] Mentor Graphics Corporation, *Eldo User Manual* (2005).

[9] V. Vapnik, A. Lerner, Pattern recognition using generalised portrait method. *Automation and Remote Control* 24 (1963).

[10] M.A. Hearst, SVMs – a practical consequence of learning theory. *IEEE Intelligent Sys.* (1998) pp. 18–21.

[11] C. J. Burges, A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2 (1998) pp. 121–167.

[12] C. Cortes, V. Vapnik, Support vector networks. *Machine Learning* 20 (1995) pp. 273–297

[13] S. Haykin, *Neural Networks: A comprehensive foundation.* 2nd edn. Prentice Hall, New Jersey (1998).

[14] S. McLoone, M. Brown, G. Irwin, G. Lightbody, A hybrid linear/nonlinear training algorithm for feedforward neural networks. *IEEE Trans. on Neural Networks* 9 (1998) pp. 669-684.

[15] J. Sjöberg, L. Ljung, Overtraining, regularization, and searching for minimum with application to neural networks. *Int. J. Control* 62 (1995) pp. 1391-1407.

[16] C.-C. Chang, C.-J. Lin, LibSVM: a library for support vector machines (2001). Software available at http://www.csie.ntu.edu.tw/~cjlin/-libsvm.

[17] S. V. N. Vishwanathan, A. J. Smola, M. N. Murty, SimpleSVM. *Proc. 20th Int. Conf. Machine Learning* (2003).

Peter C. Hung *obtained his M.Eng. in Electrical and Electronic Engineering at Queen's University Belfast, Northern Ireland in 2001. He graduated with a Ph.D. in Electrical and Electronic Engineering from the Virtual Engineering Centre, QUB in 2005. His thesis was on thermocouple sensor fusion and characterisation.*

Currently he is a post-doctoral research fellow in the National University of Ireland, Maynooth, Ireland.

His research interests include system identification, online and offline signal processing, fault diagnosis, pattern classification, and machine learning. His current projects are in the research fields of fast temperature measurement and integrated circuit testing.

Seán F. McLoone *received the M.Eng. degree in electrical and electronic engineering and the Ph.D. degree in control engineering from Queen's University Belfast, Belfast, U.K., in 1992 and 1996, respectively.*

He is a Senior Lecturer with the Department of Electronic Engineering, National University of Ireland Maynooth, Maynooth, Ireland.

His research interests are in the general area of data-based modeling and analysis of dynamical systems. This encompasses techniques ranging from classical system identification, fault diagnosis, and statistical process control to modern artificial-intelligence-inspired adaptive learning algorithms and optimization techniques. His research has a strong application focus, with many projects undertaken in collaboration with industry in areas such as process monitoring, control and optimization, time series prediction, and in-line sensor characterization.

Ronan Farrell *graduated from University College Dublin in 1993 with a B.E and returned later to receive his Ph.D. degree in 1998. After receiving his Ph.D. he joined Parthus Technologies as a mixed signal integrated circuit designer. In 2001, he left Parthus to join NUI Maynooth as a lecturer in the Department of Electronic Engineering. In 2004, he became an SFI theme leader for high frequency electronics research within the Centre for Telecommunications Value-Chain Driven Research (CTVR) which is the academic partner to Bell Labs Ireland. This Centre, with Bell Labs, remains the largest single investment (€79 million) by any research funding agency in Ireland. As of December 2005, Ronan was appointed director of the Institute of Microelectronics and Wireless Systems at NUI Maynooth, a multidisciplinary centre focused on applied research in wireless systems and their enabling technologies.*

Ronan's personal research interests include circuit design, microelectronics, and wireless technologies.