LNA from 1 to 2GHz are given in Fig. 3. The simulated and typical measured performances of the LNA are shown in Fig. 4.

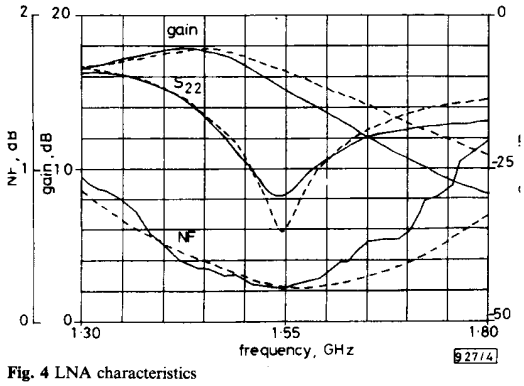

Fig. 4 LNA characteristics

—— Measured
------ Simulated

Three LNA stages were constructed to prove consistency and repeatability. The measured noise figure and associated gain at 1.55GHz are 0.22dB, 0.25dB, 0.28dB and 14.9dB, 15dB, 14.8dB, respectively. The unconditionally stable characteristics of the LNA are illustrated by measurement as well as by analysis. The measured S parameters over 10–18.0GHz are used to calculate the Rollett stability factor. The results are shown in Fig. 5.

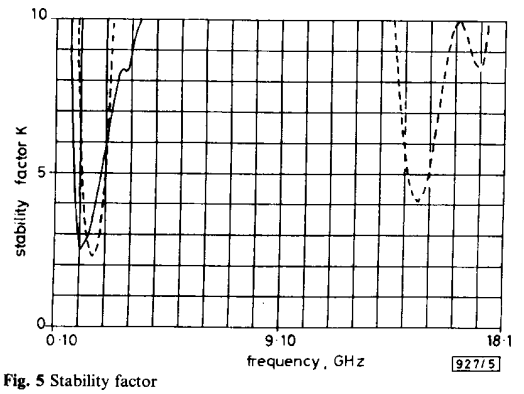

Fig. 5 Stability factor

------ k Simulated
—— k Measured

Conclusions: An LNA miniaturisation technique which does not compromise electrical performance is outlined. The distinct feature of the design topology is the assignment of multiple electrical functions to the lumped components of the matching networks. Application of the design technique is illustrated with an L-band LNA design realised in MIC technology with a total surface area not exceeding 4.8 × 8.8mm², and with outstanding noise figure performance of 0.25dB. This is believed to be the first report of an extensively miniaturised L-band LNA with a noise figure significantly better than 0.5dB. The same design technique has been used at S-band and a similar electrical performance has been achieved.

26 August 1993

Electronics Letters Online No: 19931103

E. Bayar (Com Dev Europe, Aylesbury, Bucks, United Kingdom)

A. H. Aghvami (King's College, University of London, London, United Kingdom)

References

1 BAYAR, E., ROBERTSON, I.D., and AGHVAMI, A.H.: 'An advanced technology L-band light weight receiver front end'. European Gallium Arsenide Application Symp., April 1992, Paper ESA WPP-027

2 IMAI, Y., TOKUMITSU, M. and MINAKAWA, A.: 'Design and performance of low current GaAs MMIC's for L-band front-end applications', IEEE Trans. February 1991, **MTT-32**, pp. 209–215

Planar MESFET transmission wave amplifier

T. Mader, J. Schoenberg, L. Harmon and Z.B. Popović

Indexing terms: Microwave amplifiers, Active antennas, MESFETs

The authors present a quasioptical power combining transmission amplifier for increasing the power level available from solid-state circuits. Receiving and transmitting arrays of patch antennas, input/output isolation, MESFETs, bias and matching circuitry are contained on a single substrate, making monolithic millimetre-wave integration possible. The flexibility of selecting input polarisation with respect to the output while maintaining amplifier stability is demonstrated. A 24-MESFET patch antenna amplifier array is presented.

Introduction: A plane-wave amplifier was first demonstrated [1,2] in which the active devices are closely spaced with respect to the free-space wavelength. The input and output are polarisation isolated, where both the input and output waves are linearly polarised, so external polarising elements are required. The active grid power combining technique was demonstrated earlier in a 100 element MESFET oscillator [3]. Patch antenna oscillator power combiners with Gunn diode and transistor active elements have also been demonstrated [4,5]. Here we present a MESFET patch antenna amplifier array in which the input and output wave polarisations may be arbitrarily selected. This amplifier, shown in Fig. 1, has an array of patch antennas on input and output sides of the substrate. Each common-source MESFET amplifier is coupled and matched to its input and output patch elements with microstrip lines. A substrate via connects the output of each amplifier to its radiating patch printed on the opposite side of the substrate. Alternating ground planes effectively isolate input and output sections of the amplifying structure. The ground planes are connected by vias periodically spaced at quarter-wavelength intervals. High impedance gate and drain bias lines run through the voltage nulls of the patches on input and output surfaces, respectively.

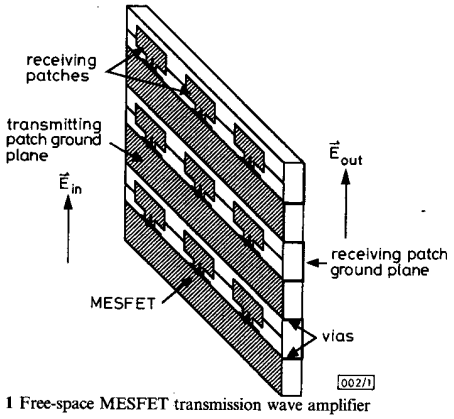

Fig. 1 Free-space MESFET transmission wave amplifier

The input wave is received by patch antennas connected to the gates of the transistors. The drains are connected to transmitting patches on the other side of the substrate. The ground plane isolates the input from the output.

Single-MESFET, dual-patch amplifier element: A 10GHz prototype amplifier element using Avantek ATF13484 GaAs MESFETs was built on a 0.508mm thick substrate with $\epsilon_r = 2.2$. To save space, minimise parasitics, and maximise bandwidth, a single transmission line matching section is used. The patch antennas were designed using an analysis and synthesis program [6], and are fed at their radiating edges. The use of a frequency-dependent one-port equivalent circuit for the patch is essential to ensure stability of the amplifier over the frequency range in which the MESFET has gain. A bilateral design at 10GHz was performed with a 100Ω stabilising resistor between the gate and source of the MESFET. The DC and RF voltages across the gate-to-source terminals of the MESFET are smallest in magnitude, so that efficiency is maximised and heat dissipation minimised. This design is stable and the maximum gain is 8.4dB. The final design of a single-MESFET dual-patch amplifier is shown in Fig. 2a. The chip resistor is soldered directly between the transistor gate and source leads.

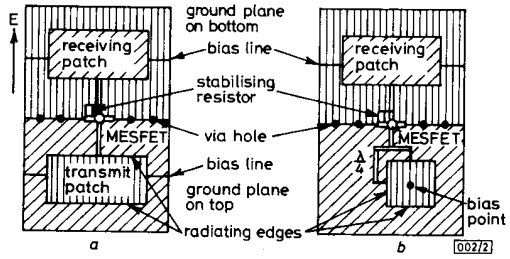

Fig. 2 Single-MESFET dual-patch free space amplifier with resistive loading for stability and single-FET amplifier with linearly polarised input and circularly polarised output wave

a Single-MESFET amplifier
b Single-FET amplifier

A free-space measurement technique similar to that in [1] determines the gain contributed by the MESFETs, as well as the frequency response of the array. Our technique uses a known gain value of the patch element and the theoretical directivity of a uniform amplitude, equiphase array. Fig. 3 shows the experimental setup in which an absorbing screen at the array plane prevents diffraction path measurement error. Standard gain horns G_1 and G_2 are connected to an HP83640A synthesised sweeper and HP71500A/HP70820A microwave transition analyser, respectively. First the path loss between the horn antennas is measured. The free-space transmission amplifier is then placed between the horns, and the path loss is again measured. The ratio between these two measured powers is then given by

$$P' = \frac{P}{P_0} = G_{in} G_{out} G_{FET} \frac{(r_1 + r_2)^2}{r_1^2 r_2^2} \left(\frac{\lambda}{4\pi}\right)^2 \quad (1)$$

where G_{in} and G_{out} are the effective input and output antenna gains of the amplifier array, respectively, G_{FET} is the gain added by the MESFETs, and r_1 and r_2 are the distances from the horn antennas to the array. Distances r_1 and r_2 are selected to be equal and in the far field of both the amplifier array and the horns.

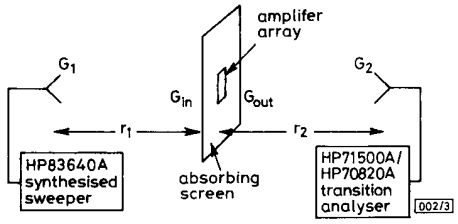

Fig. 3 Experimental setup for measuring transistor-added gain for free-space patch antenna amplifier

Once P' is measured, the only unknown parameter in the above equation is G_{FET}, which is expected to be the same regardless of array size. The transistor-added gain of the single-MESFET dual-patch amplifier was measured from 9 to 11GHz using the method outlined above, and the results are plotted as solid lines in Fig. 4. When the FETs are not biased, there should be little power transmitted through the amplifier, as the alternating ground planes provide good input-to-output isolation. When the FETs are biased (V_{DS} = 2.5V, I_{DS} = 20mA), a maximum gain of 7.1dB is measured at 10GHz. The crosspolarisation ratio for the output of the amplifier is 23.4dB. In Fig. 4, the patch is assumed to have 5dB gain over the entire 9–11GHz range. As this is not the case, the graph gives insight into the degradation of the entire system gain due to the limited bandwidth of the patches and matching circuits.

The structure of the patch antenna amplifier array allows for flexibility in the polarities of both the input and output waves. A single patch antenna amplifier element was designed to receive a linearly polarised wave, and transmit a circularly polarised wave, as shown in Fig. 2b. The transmitting square patch is fed on two adjacent sides with 300Ω edge impedance and a 90° separation in phase between the two feeds. The drain of the MESFET is biased through a via to the back of the square patch at a voltage null. The transistor-added gain of the single-MESFET dual-patch circularly polarised amplifier was measured from 9 to 11GHz using the method outlined above, and the results are plotted as dotted lines in Fig. 4. A maximum gain of 6.8dB and an output axial ratio within 0.3dB were measured at 10GHz.

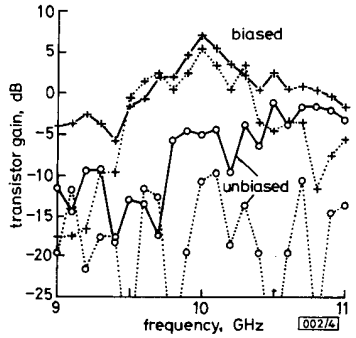

Fig. 4 Measured gain against frequency for the single-MESFET linearly polarised wave amplifier and for wave amplifier with circularly polarised output wave

+ Biased measurements
○ unbiased measurements
— linearly polarised
···· circularly polarised

Patch antenna amplifier array: The single patch antenna amplifier element was placed into a 4x6 array, so that the power of 24 MESFETs could be combined in free space as shown in Fig. 1. The horizontal spacing between elements is 2λ/3 and the vertical spacing between elements is 4λ/3. It is assumed that this spacing is large enough so that the individual amplifiers do not parasitically couple to each other. The measured gain for the 24-MESFET array at 10GHz was 21dB higher than that of the single element amplifier.

Acknowledgments: We thank K.C. Gupta for the use of the patch design program. This work was funded by the Naval Weapons Research Center under Contract N60530-92-C-0238. J. Schoenberg holds an AFIT Fellowship, and L. Harmon is funded by the National Science Foundation REU program.

3 August 1993

Electronics Letters Online No.: 19931085

T. Mader, J. Schoenberg, L. Harmon and Z. B. Popović (Department of Electrical and Computer Engineering University of Colorado, Boulder, CO 80309, USA)

References

1 KIM, M., ROSENBERG, J.J., SMITH, R.P., WEIKLE, R.M., HACKER, J.B., DELISIO, M.P., and RUTLEDGE, D.B.: 'A grid amplifier', IEEE Microw. Guid. Wave Lett., 1991, **1**, (11) pp. 322–324

2 WEIKLE, R.M., KIM, M., HACKER, J.B., DELISIO, M.P., POPOVIĆ Z.B., and RUTLEDGE, D.B.: 'Transistor oscillator and amplifier grids', Proc. IEEE, 1991, **11**, (11) pp. 1800–1809

3 POPOVIĆ Z.B., WEIKLE, R.M., KIM, M., and RUTLEDGE, D.B.: 'A 100-MESFET planar grid oscillator', IEEE Trans., 1991, **MTT-39**, (2) pp. 193–200

4 BIRKELAND, J., and ITOH, T.: 'An FET oscillator element for spatially injected locked arrays', Paper 3, IEEE MTT-S Int. Symp. Dig., 1992, pp. 1535–1538

5 YORK, R.A., and COMPTON, R.C.: 'Quasi-optical power combining using mutually synchronized oscillator arrays', IEEE Trans., 1991, **MTT-39**, (6), pp. 1000–1009

6 BENALLA, A., THNG, CHEOK, GUPTA, K.C.: in: 'CAD of microstrip patch antennas'. (MICROPATCH, Microstrip Designs, Inc., 1993)

Extended backpropagation algorithm

Y.F. Yam and T.W.S. Chow

Indexing terms: Backpropagation, Adaptive learning, Accelerated learning, Correlation coefficient

A fast learning algorithm for feedforward neural networks is described. This algorithm is based on modifications to the backpropagation algorithm. The learning rate and momentum coefficient are adapted according to the coefficient of correlation between the downhill gradient and the previous weight update.

Introduction: The Rumelhart backpropagation algorithm is perhaps the most widely used learning algorithm for feedforward neural networks [1]. It has been successfully applied to solve pattern recognition, signal processing and control problems [2,3]. The backpropagation algorithm is based on the technique of gradient descent, i.e. the weight update is performed in the opposite direction to the gradient of an error cost function. However, the convergence of gradient-descent learning is often too slow for practical applications. A very long training time is required to train medium sized networks for even the simplest problems.

A number of methods have been suggested in the literature for overcoming the problem. Most methods can significantly reduce the number of iterations to convergence. However, the computational complexity and storage requirement are greatly increased and the overall reduction in convergence time is small.

This Letter presents a new algorithm which is considerably faster than the backpropagation algorithm and has the additional advantage of being less affected by poor initial weight setting. The new algorithm is capable of varying the learning rate and momentum coefficient dynamically based on the correlation coefficient between the present downhill gradient and the last change in the weight.

Extended backpropagation algorithm: In the propagation algorithm, training is only based on the current downhill gradient of the error surface $-\nabla E(t)$ and last change in weight $\Delta W(t-1)$. E is the mean squared error, which is defined as $1/(2P)\Sigma_p\Sigma_k (t_{pk}-o_{pk})^2$. The two column vectors, $-\nabla E(t)$ and $\Delta W(t-1)$, are n dimensional vectors with elements $\delta E/\delta w_{ij}$ and Δw_{ij} respectively at the tth iteration. The local topography of error surface can be thoroughly evaluated for determining the optimum learning rate and momentum coefficient. The higher derivatives of the error function and line search technique are often used to find the optimum learning rate, however, a long computational time is required for these processes. In this newly developed algorithm, the learning rate adaptation is simply based on the correlation coefficient between the current downhill direction of the local gradient and the previous weight update. The correlation coefficient between the negative gradient and last weight update is given by

$$r = \frac{\sum(\nabla E_{ij} - \nabla \bar{E}_{ij})(\Delta W_{ij} - \Delta \bar{W}_{ij})}{\left[\sum(\nabla E_{ij} - \nabla \bar{E}_{ij})^2\right]^{\frac{1}{2}} \left[\sum(\nabla E_{ij} - \nabla \bar{E}_{ij})^2\right]^{\frac{1}{2}}}$$

where $\nabla E_{ij}=\Sigma \nabla E_{ij}/n$ and $\Delta \bar{w}_{ij}=\Sigma w_{ij}/n$. Three different conditions can be identified from this correlation coefficient:

(a) When the correlation coefficient is near to 1 there is almost no change in the direction of error minimisation and the change of weights is likely moving on the plateau. The learning rate can be increased to improve the convergence rate.

(b) When the correlation coefficient is near to –1, it implies an abrupt change in the direction of error minimisation which is likely to be moving along the wall of a ravine. The learning rate should then be reduced to prevent oscillation across both sides of the ravine.

(c) When there is no correlation between the negative gradient and previous weight update, the learning rate should be kept constant.

The following algorithm is suggested to alter the learning rate:

(i) r(0) = 0.0 and set the value of t = 1

(ii) **if** r(t) > 0.0
  if r(t–1) > 0.0 **then** $\eta(t) = \eta(t-1)[1.0 + 0.5r(t)]$
  else $\eta(t) = 1.0 + 0.5r(t)$
if r(t) < 0.0
  if r(t–1) > 0.0 then $\eta(t) = 1.0 + 0.5r(t)$
  else $\eta(t) = \eta(t-1)[1.0 + 0.5r(t)]$
if r(t) = 0.0 **then** $\eta(t) = \eta(t-1)$

(iii) set the value of t = t + 1, and repeat step (ii).

One of the significant features of this algorithm is the exponential increasing and decreasing of the learning rate. The learning rate can increase or decrease rapidly when the successive values of correlation coefficient remain the same sign. This feature enables the optimal learning rate to be found in a few learning iterations, and thus reduces the total output error rapidly. When the correlation coefficient changes sign, the algorithm can reset the excessive large value of learning rate and enhance the stability of this algorithm.

However, the convergence rate is not optimised with fixed momentum coefficient α. The momentum term has an accelerating effect only when the $-\nabla E(t)$ and $\Delta W(t-1)$ have the same direction. For the fixed momentum coefficient, the momentum term may override the gradient term when the $-\nabla E(t)$ and $\Delta W(t-1)$ are in the opposite direction. The momentum coefficient α is suggested to be

$$\alpha(t) = \lambda(t)\eta(t)\frac{\|-\nabla E(t)\|}{\|\Delta W(t-1)\|} \tag{3}$$

with

$$\lambda(t) = \eta(t) \tag{4}$$

The final algorithm becomes

$$\Delta W(t) - \eta(t)\|-\nabla E(t-1)\|[-\nabla E(t) + \eta(t)\Delta W(t-1)] \tag{5}$$

$-\nabla \hat{E}(t)$ and $\Delta \hat{W}(t)$ are the unit vectors of $-\nabla E(t)$ and $\Delta W(t)$, respectively. When $-\nabla \hat{E}(t)$ and $\Delta \hat{W}(t)$ are in opposite directions, $\eta(t)$ is less than 1.0, and hence the momentum term does not override the gradient term. When $-\nabla E(t)$ and $\Delta W(t-1)$ have the same direction, $\eta(t)$ is greater than 1.0, and enhance the accelerating effect of the momentum term. The fast increase of learning rate may drive the neurons to their extreme values in some neural network problems. This phenomenon can be avoided by the following strategy. If the present RMS error is greater than the previous one by 0.1, the last step is cancelled and $\eta(t)$ is reduced by half. This strategy gives a small preference to learning rate reduction and enhances the robustness of the training process.

Results and discussions: To compare the convergence performance of this learning algorithm with the conventional and adaptive backpropagation algorithm[4], we applied these algorithms to the