further recessed and the HEMTs can be changed from depletion-mode to enhancement-mode. The 0.25V shift in V_{th} corresponds to about a 4nm shift in the position of the gate metal front. The ratio of the absorption depth (4nm) to the physical thickness (2.8nm) of the reacted Pt is ~1.4, consistent with that reported in [6].

As illustrated in Fig. 1*b*, the absorption of Pt into the InAlAs barrier does not change the electron density in the channel region between the source and gate electrodes. Therefore, R_S remained unchanged during the annealing process. R_S is ~0.2Ωmm obtained by TLM measurement. Owing to the small R_S obtained, the maximum g_m is not degraded but increased further to 1170mS/mm as a result of the reduced distance between the gate metal and the 2DEG channel. The maximum intrinsic transconductance g_i is estimated to be 1520mS/mm.

To further demonstrate the improvement in R_S gained by the buried gate approach, E-HEMTs were also fabricated using only wet recess etching. As shown in Fig. 1*a*, in addition to the region under the gate metal, the channel region around the gate periphery was also depleted due to the side-etching effect and a large R_S results. This larger R_S results in a degraded maximum g_m of 500mS/mm for a 0.5μm gate E-HEMT. Using the calculated intrinsic transconductance of 1520mS/mm, the total source resistance R_S in this case is estimated to be ~1.5Ωmm, about 8 times larger than that achieved using the buried-gate approach.

Conclusion: Pt-based buried-gate technology can effectively reduce the source resistance in InP-based enhancement-mode HEMTs. As a result, an excellent transconductance of 1170mS/mm was achieved in a 0.5μm gate device.

Acknowledgment: The authors would like to thank N. Suzuki for MBE growth. We are also grateful to Y. Imamura and Y. Ishii for their continuous encouragement.

7 April 1995

Electronics Letters Online No: 19950603

K.J. Chen, K. Maezawa, K. Arai, M. Yamamoto and T. Enoki (*NTT LSI Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa Pref., 243-01, Japan*)

References

1 ABE, M.: 'Ultrahigh-speed InGaAs-based HEMT technology'. Proc. 4th Int. Conf. InP and Related Materials, 1991, pp. 2–5

2 ABE, M., *et al.*: 'Ultrahigh-speed HEMT integrated circuits', *in* WILLARDSON, R.K., BEER, A.C., and DINGLE, R. (Eds.): 'Semiconductors and semimetals' (Academic, New York, 1987), pp. 249–278

3 DUH, K.H., PAO, P.C., LIU, S.M.J., HO, P., KAO, M.Y., and BALLINGALL, J.M.: 'A super low-noise 0.1μm T-gate InAlAs-InGaAs-InP HEMT', *IEEE Microw. Guid. Wave Lett.*, 1991, **1**, pp. 114–116

4 TAN, K., STREIT, D.C., CHOW, P.D., DIA, R.M., HAN, A.C., LIU, P.H., GARSKE, D., and LAI, R.: '140GHz 0.1μm gate length pseudomorphic InAlAs/InGaAs/InP HEMT'. IEDM Tech. Dig., Washington, DC, 1991, pp. 239–242

5 ITOH, T., BROWN, A.S., CAMNITZ, L.H., WICKS, G.W., BERRY, J.D., and EASTMAN, L.F.: 'Depletion- and enhancement-mode $Al_{0.48}In_{0.52}As/Ga_{0.47}In_{0.53}As$ modulation-doped field-effect transistors with a recessed gate structure'. Int. Symp. GaAs and Related Compounds, Karuizawa, Japan, 1985, pp. 571–576 (Inst. Phys. Conf. Ser. No. 79)

6 HARADA, N., KURODA, S., KATAKAMI, T., HIKOSAKA, K., MIMURA, T., and ABE, M.: 'Pt-based gate enhancement-mode InAlAs/InGaAs HEMTs for large-scale integration'. Proc. 3rd Int. Conf. InP and related materials, 1991, pp. 377–380

7 ENOKI, T., KOBAYASHI, T., and ISHII, Y.: 'Device technologies for InP-based HEMTs and their application to ICs'. Tech. Dig. IEEE GaAs IC Symp., 1994, pp. 337–340

8 BAHL, S.R., and DEL ALAMO, J.A.: 'Elimination of mesa-sidewall gate-leakage in InAlAs/InGaAs HFETs by selective side-wall etching', *IEEE Electron Device Lett.*, 1992, **13**, pp. 195–197

Low-stress nitride as oxidation mask for submicrometre LOCOS isolation

H.W. van Zeijl, L.K. Nanver and P.J. French

Indexing terms: Oxidation, Semiconductor technology

A large reduction in the LOCOS bird's beak is obtained by using a thick silicon nitride layer as an oxidation mask. Stress induced damage of the devices is avoided by using low-stress silicon-rich nitride.

Introduction: The conventional local oxidation of silicon (LOCOS), routinely used in IC processing, is characterised by excellent electrical properties such as low diode leakage and high gate oxide quality. However, lateral oxide encroachment ('bird's beak' formation) under the nitride, reduces the device dimensions, making conventional LOCOS unsuitable for deep submicrometre geometries. Advanced isolation techniques that reduce the bird's beak typically rely on either sidewall masking, interface sealing or pad-oxide thinning [1]. The tradeoffs are an increase in process complexity and the need for more critical control of defect formation.

In this Letter, a very straightforward technique is presented for effectively reducing the LOCOS bird's beak. A relatively thick low-stress silicon-rich nitride film, SiN_x, is deposited by standard LPCVD techniques and this film acts as a rigid clamp, suppressing the oxide encroachment. The increased pressure under the nitride edge will cause the oxide to flow away from the edge and simultaneously the diffusion and solubility of the oxidant species are decreased [2]. This low-stress nitride, frequently used for the fabrication of thin membranes in sensor applications [3], preserves the properties of interest in stoichiometric nitride while alleviating the stress related problems. Normally, with thick nitride layers, the stress becomes too high and defects are generated in the underlying silicon substrate. The relationship between the stress in the nitride and defect generation has previously been examined by using SiN_x for surface isolation in a 15 GHz washed-emitter-base (WEB) bipolar NPN process [4]. For strain levels below 1.8×10^{-3}, no emitter-collector (E-C) shorts were detected, while with strains from $2 \times 10^{-3} - 3 \times 10^{-3}$, the frequency of shorts increased rapidly with the strain. A WEB NPN with a LOCOS isolated emitter was developed to detect any stress induced damage related to the processing of the 'bird's beak free' LOCOS. Again the number of E-C shorts was correlated to the mechanical stress and damage free characteristics were obtained by using low-stress nitride.

Processing procedures: Silicon nitride films were deposited using a standard LPCVD system. Throughout the experiments, the temperature and pressure were maintained at 850°C and 150mtorr, respectively. The gases used were $SiCl_2H_2$ (DCS) and NH_3, and the ratio of these two gases was used as a process variable, with the total gas flow being maintained at 200sccm. The mechanical strain in the silicon nitride films was determined by using a micromachined strain measurement structure [5]. The high dependence of film strain on the gas flow ratio, and thus the ratio of silicon to nitrogen in the fillm, is shown in Fig. 1.

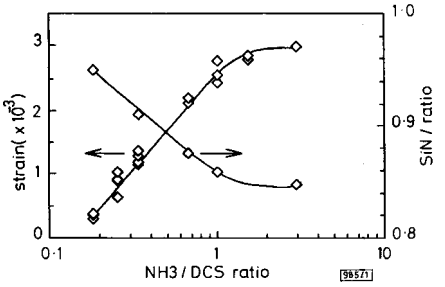

Fig. 1 *Film strain and ratio of silicon to nitrogen in film against NH_3/DCS gas flow ratio*

The basic WEB NPN process flow with a LOCOS isolated emitter is shown in Fig. 2. The substrate is (100) *n*-type with a resistivity of 2 – 5Ωcm. After a wet etch step in HF (1%) to remove any native oxide, a 300nm SiN_x film is deposited on the silicon. The emitter area is then defined and outside these regions the nitride is removed by plasma etching. A 0.4μm thick field oxide is thermally grown at 1000°C in a wet ambient. The extrinsic base is implanted through the oxide with a resist mask protecting the emitter regions. The remaining nitride is then removed by wet etching, and, using the field oxide as a mask, the emitter and intrinsic base are implanted to conform with the standard WEB process flow. Contact windows to the extrinsic base are etched in the field oxide. All dopants are activated by a thermal anneal for 30min at 950°C and the windows are contacted by sputtering Al/1%Si.

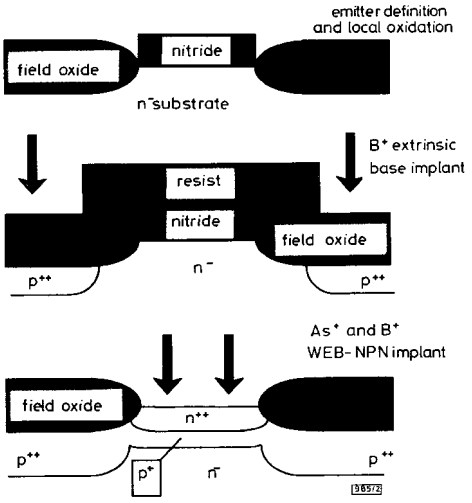

Fig. 2 *Basic WEB NPN process flow with LOCOS isolated emitter*

Results: The bird's beak formation has been studied for thin (97 nm) and thick (300 nm) low-stress nitride (NH_3/DCS ratio = 0.18) combined with pad oxides from 0 – 27nm thick. In these tests, a field oxide of 0.8 μm was grown. SEM cross-sections of the six combinations are shown in Figs. 3 and 4. With the thin nitride layer, comparable to conventional LOCOS processing, the bird's beak extends over the entire emitter window. Some suppression of the bird's beak is achieved by omitting the pad oxide and thus reducing the lateral oxidation path. However, the use of a thick nitride layer proves to be much more efficient in all cases.

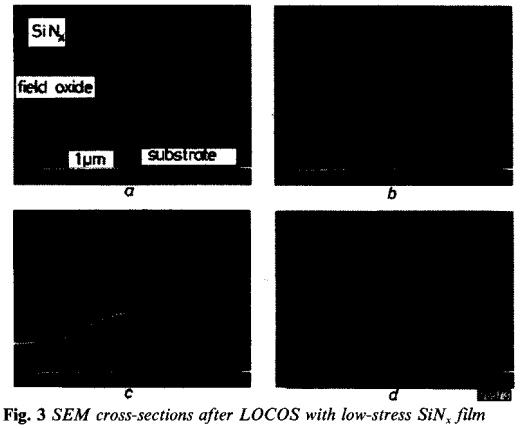

Fig. 3 *SEM cross-sections after LOCOS with low-stress SiN_x film*

a Wafer A, no pad oxide, film thickness = 97nm
b Wafer B, 9nm pad oxide, film thickness = 97nm
c Wafer C, 27nm pad oxide, film thickness = 97nm
d Wafer D, no pad oxide, film thickness = 300nm

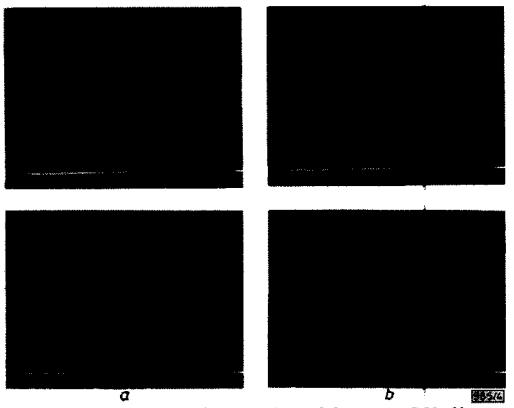

Fig. 4 *SEM cross-sections after LOCOS with low-stress SiN_x film*

a Wafer E, 9nm pad oxide, film thickness = 300nm
b Wafer F, 27nm pad oxide, film thickness = 300nm

Wafers with 300nm thick SiN_x with strains of either 0.28×10^{-3}, 2.0×10^{-3} and 3.0×10^{-3} were prepared and processed with WEB NPNs with LOCOS isolated emitters. After the field oxidation, a visual evaluation of the stress induced damage was performed by dipping a set of stripped wafers in Wright etch for 2min. No defects were observed for the lowest strain level, whereas for the higher stress levels the number of defects increased with the stress. Over the wafer, NPNs with several emitter geometries were measured. With an emitter area of $1 \times 2\mu m^2$, no E-C shorts were detected, but for larger geometries the frequency of E-C shorts directly correlates to the SiN_x strain and the perimeter of the emitter (Table 1). Thus both the visual and electrical inspections confirm the absence of stress induced damage in the samples with low-stress nitride.

Table 1: Strain levels and percentage of devices with shorts for several emitter areas (A_e) at different NH_3/DCS ratios

Wafer	NH_3/DCS ratio	Strain ($\times 10^{-3}$)	Percentage of E-C shorts		
			54 transistor arrays $A_e = 12.000 \times 2 \times 1\mu m^2$	108 transistors $A_e = 10 \times 40\mu m^2$	54 transistors $A_e = 40 \times 40\mu m^2$
1	0.18	0.28	0	0	0
2	0.58	2.00	0	2	8
3	3.00	3.00	0	26	30

Conclusions: The use of a thick low-stress nitride layer as an oxidation mask in LOCOS is very effective in reducing the bird's beak without introducing complex processing. The damage in the silicon has been directly correlated to the strain of the SiN_x layer. Defect free devices have been produced by using a low-stress material with a strain of 0.28×10^{-3}.

Acknowledgments: The authors are grateful to C.I.M. Beenakker for his continual support and encouragement, and to R. Mallee and W. Verveer for their contributions to the experimental material.

18 April 1995
Electronics Letters Online No: 19950626

H.W. van Zeijl, L.K. Nanver and P.J. French (*Delft University of Technology, PO Box 5053, NL-2600 GB Delft, The Netherlands*)

References

1 KOOI, E.: 'The invention of LOCOS'. IEEE Case Histories of Achievement in Science and Technology, (IEEE, New York, 1991), Vol. 1, pp. 101–128,

2 KAO, D.B., MCVITTIE, J.P., NIX, W.D., and SARAWAT, K.C.: 'Two dimensional silicon oxidation experiments and theory'. IEDM Tech. Dig., 1985, pp. 388–391

3 FRENCH, P.J., WOLFFENBUTTEL, R.F., MALLEE, R., and SARRO, P.: 'Optimization of a low-stress silicon nitride process for surface micromachining applications'. Book of abstracts, Eurosensors VIII Conf., 1994, pp. 205
4 NANVER, L.K., FRENCH, P.J., GOUDENA, E.J.G., and VAN ZEIJL, H.W.: 'Low-stress nitride as surface isolation in bipolar transistors', to be published in *J. Mater. Sci.*, January 1995
5 VAN DRIEENHUIZEN, B.P., GOOSEN, J.F.L., FRENCH, P.J., and WOLFFENBUTTEL, R.F.: 'Comparison of techniques for measuring both compressive and tensile stress in thin films', *Sens. Actuators*, 1993, **A37–38**, pp. 756–765

Least-squares moment matching reduction methods

I.D. Smith and T.N. Lucas

Indexing term: Model order reduction

It is shown how the two apparently different approaches to modelling reduction by least-squares time moment matching give identical results. This property also enables a clear understanding of how the method actually approximates the system. Computational savings are seen to be made by adopting the two-stage approach.

Introduction: Since Shoji *et al.* [1] first proposed the idea of using a least-squares moment matching technique for reducing the order of linear, time-invariant systems, there has been much interest in this method. The method has been refined by Lucas and Beat [2] and extended to also include the use of Markov parameters [3, 4]. More recently it has been adopted by Lalonde [5] for use in discrete-time systems using Markov parameters only. However, the focus of the work so far appears to concentrate mainly on the basic idea of extending the fitting of system parameters so that a stable reduced model results. Little has been written to clarify how the technique actually approximates the full system.

The purpose of this Letter is first to clarify what the two different approaches to least-squares moment matching are, then to prove the equivalence of the resulting models. Finally, observations are made on the implications for understanding the least-squares reduction process and the sense in which it minimises the errors involved. The computational efficiency of the method is also seen to benefit from the analysis.

Least-squares Padé methods: It is well known that for continuous time systems [6] a kth order reduced transfer function $R(s)$ may be derived from the nth order full transfer function $R(s)$ using the exact Padé method which matches the first $2k$ time moments of $G(s)$ with those of $R(s)$, where the time moment proportionals, c_i (i = 0, 1, 2, ...), are defined by

$$G(s) = c_0 + c_1 s + c_2 s^2 + c_3 s^3 + \ldots \quad (1)$$

(hereafter the c_i will be referred to simply as the time moments). This approach leads to reduced order models which may give good steady-state approximations to the system response, but which can also lead to unstable models being derived from stable full systems.

It was in an attempt to ameliorate this last-mentioned situation that Shoji *et al.* [1] first proposed extending the number of time moments included in the calculation until a stable reduced model is produced. This idea of introducing approximation, in the least-squares sense, into the derivation of the reduced model was further analysed by Lucas and Beat [2] and was also incorporated as a special case of a least-squares method proposed by Aguirre [4]. The former methods [1, 2] are what shall be referred to as the *partial* least-squares method, in that only the reduced denominator coefficients are approximated in the least-squares sense before the numerator coefficients are derived by exact moment-matching. On the other hand, the method of Aguirre [4] is what shall be termed the *full* least-squares method in that both the numerator and denominator are derived at the same time in a single generalised inverse operation. It will now be proved that, for least-squares Padé methods involving the use of time moments only, the full and partial methods are equivalent.

For a reduced kth order model given by

$$R(s) = \frac{d_0 + d_1 s + \ldots + d_{k-1} s^{k-1}}{e_0 + e_1 s + \ldots + e_{k-1} s^{k-1} + s^k} \quad (2)$$

which is derived from the first $2k + t$ time moments of $G(s)$, where t is the number of extra time moments used to extend the exact Padé method, the following set of equations are formed [4] by equating eqns. 1 and 2 :

$$d_i = c_i e_0 + c_{i-1} e_1 + \ldots + c_1 e_{i-1} + c_0 e_i \qquad i = 0, 1, \ldots, k-1 \quad (3)$$

and

$$0 = c_i e_0 + c_{i-1} e_1 + \ldots + c_{i-k+1} e_{k-1} + c_{i-k} \qquad i = k, \ldots, 2k+t-1 \quad (4)$$

where $e_j = c_j = 0$ for $j < 0$. These may be expressed in the matrix-vector form of

$$\mathbf{A}\underline{x} = \underline{b}$$

or in partitioned form as

$$\left[\begin{array}{c|c} \mathbf{I}_k & \mathbf{c}_0 \\ \hline \mathbf{\Phi} & \mathbf{c}_1 \end{array}\right] \left[\begin{array}{c} \underline{d} \\ \hline \underline{e} \end{array}\right] = \left[\begin{array}{c} \underline{0} \\ \hline \underline{q} \end{array}\right] \quad (5)$$

where $\mathbf{I}_k$ is the k by k identity matrix, Φ is the $(k+t)$ by k null matrix and

$$\underline{d} = [d_0\, d_1 \ldots d_{k-1}]^T \quad \underline{e} = [e_0\, e_1 \ldots e_{k-1}]^T$$

$$\underline{q} = [c_0\, c_1 \ldots c_{k+t-1}]^T$$

$$\mathbf{c}_0 = \begin{bmatrix} -c_0 & & & \\ -c_1 & -c_0 & & 0 \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot \\ -c_{k-1} & \cdot & \cdots & \cdot \; -c_0 \end{bmatrix}$$

$$\mathbf{c}_1 = \begin{bmatrix} -c_k & -c_{k-1} & \cdots & -c_1 \\ -c_{k+1} & -c_k & \cdots & -c_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ -c_{2k+t-1} & \cdot & \cdots & -c_{k+t} \end{bmatrix}$$

Here it is also noticed that the corresponding matrix-vector equation for the partial method [1, 2], to determine the denominator, can be written simply as

$$\mathbf{c}_1 \underline{e} = \underline{q} \quad (6)$$

Now the least-squares solution of eqn. 5 is obtained from

$$\mathbf{A}^T \mathbf{A} \underline{x} = \mathbf{A}^T \underline{b}$$

and notice that

$$\mathbf{A}^T\mathbf{A} = \left[\begin{array}{c|c} \mathbf{I}_k & \mathbf{\Phi}^T \\ \hline \mathbf{c}_0^T & \mathbf{c}_1^T \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_k & \mathbf{c}_0 \\ \hline \mathbf{\Phi} & \mathbf{c}_1 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{I}_k & \mathbf{c}_0 \\ \hline \mathbf{c}_0^T & \mathbf{c}_0^T\mathbf{c}_0 + \mathbf{c}_1^T\mathbf{c} \end{array}\right]$$

and

$$\mathbf{A}^T \underline{b} = \left[\begin{array}{c|c} \mathbf{I}_k & \mathbf{\Phi}^T \\ \hline \mathbf{c}_0^T & \mathbf{c}_1^T \end{array}\right] \left[\begin{array}{c} \underline{0} \\ \hline \underline{q} \end{array}\right] = \left[\begin{array}{c} \underline{0} \\ \hline \mathbf{c}_1^T \underline{q} \end{array}\right]$$

which yields the least-squares solution from

$$\left[\begin{array}{c|c} \mathbf{I}_k & \mathbf{c}_0 \\ \hline \mathbf{c}_0^T & \mathbf{c}_0^T\mathbf{c}_0 + \mathbf{c}_1^T\mathbf{c}_1 \end{array}\right] \left[\begin{array}{c} \underline{d} \\ \hline \underline{e} \end{array}\right] = \left[\begin{array}{c} \underline{0} \\ \hline \mathbf{c}_1^T \underline{q} \end{array}\right]$$

This is equivalent to solving the two blocks of equations

$$\underline{d} + \mathbf{c}_0 \underline{e} = \underline{0} \quad (7)$$

and