Electronic traps and P_b centers at the Si/SiO_2 interface: Band-gap energy distribution

E. H. Poindexter, G. J. Gerardi, M.-E. Rueckel, and P. J. Caplan
U. S. Army Electronics Technology and Devices Laboratory (ERADCOM), Fort Monmouth, New Jersey 07703

N. M. Johnson and D. K. Biegelsen
Xerox Palo Alto Research Center, Palo Alto, California 94304

(Received 22 May 1984; accepted for publication 3 July 1984)

Energy distribution of P_b centers ($\cdot Si{\equiv}Si_3$) and electronic traps (D_{it}) at the Si/SiO_2 interface in metal-oxide-silicon (MOS) structures was examined by electric-field-controlled electron paramagnetic resonance (EPR) and capacitance-voltage (C-V) analysis on the same samples. Chips of (111)-oriented silicon were dry-oxidized for maximum P_b and trap density, and metallized with a large MOS capacitor for EPR and adjacent small dots for C-V measurements. Analysis of C-V data shows two D_{it} peaks of amplitude $2\times10^{13}\,eV^{-1}\,cm^{-2}$ at $E_v + 0.26$ eV and $E_v + 0.84$ eV. The EPR spin density reflects addition or subtraction of an electron from the singly occupied paramagnetic state and shows transitions of amplitude $1.5\times10^{13}\,eV^{-1}\,cm^{-2}$ at $E_v + 0.31$ eV and $E_v + 0.80$ eV. This correlation of electrical and EPR responses and their identical chemical and physical behavior are strong evidence that $\cdot Si{\equiv}Si_3$ is a major source of interface electronic traps in the 0.15–0.95 eV region of the Si band gap in unpassivated material.

I. INTRODUCTION

Electron paramagnetic resonance (EPR) has revealed a simple structural picture of the P_b center, which is a major characteristic defect at the Si/SiO_2 interface of metal-oxide-silicon (MOS) devices.[1,2] On (111)-oriented silicon, the P_b center has been identified as $\cdot Si{\equiv}Si_3$, that is, a triply coordinated silicon atom, with nonbonded or dangling orbital normal to the interface plane and aimed into the oxide. It has long been suspected or assumed that $\cdot Si{\equiv}Si_3$ is a source of interface electrical traps, the band-gap density of which is designated D_{it}. A complete definition of the structural, physical, chemical, and electrical nature of this center is important in further development of ultrasmall integrated circuit elements, especially in consideration of the defects arising from high internal operating fields or stressful external environments.

In this paper, the band-gap energy distribution of P_b levels on (111)-oriented silicon wafers is treated. Both EPR and electrical methods are discussed and compared. A straightforward method was used, specifically, application of an electric field to MOS capacitors and observation of interface trap occupancy by EPR and capacitance-voltage (C-V) analysis. The first qualitative measurements[3] of this type on an MOS structure showed an amphoteric trap which could hold 0, 1, or 2 electrons, and indicated the approximate positions of the ($0\leftrightarrow1$) and ($1\leftrightarrow2$) electron transitions in the lower and upper halves of the band gap, respectively. Subsequent observations made with a remotely applied field from corona discharge, which avoids practical problems of large-area MOS capacitors, supported and extended the amphoteric model.[4,5] Deep-level transient spectroscopy (DLTS), together with C-V and improved EPR studies of MOS structures, led to partially quantitative correlation of the P_b transitions with D_{it} peaks observed at 0.3 and 0.8 eV above the valence band.[6] A similar P_b level splitting was observed by the corona-discharge method.[7]

In this paper we present new results and extended discussion on the quantitative determination of the P_b level distribution, correlated with D_{it} on the same sample chips. The MOS method was chosen for best controllability and most realistic device impact. Data were gathered over a sufficient range for good definition of the features of interest; the results quantify and substantiate the amphoteric model of P_b centers and interface traps.

II. EXPERIMENTAL PRINCIPLES AND TECHNIQUES

The general theories and methods of EPR have been well presented in text.[8] Special considerations for study of the Si/SiO_2 interface have been summarized in a recent review article.[9] The principles of C-V analysis are exhaustively discussed in text,[10] and DLTS techniques in a previous paper.[11]

A. Sample preparation

Samples for correlated EPR and C-V measurements were fabricated on the same Czochralski-grown n- and p-type silicon chips of (111) orientation, with resistivity greater than 30 Ω cm. They were oxidized in dry O_2 at 1000 °C to 190-nm thickness and rapidly cooled in oxygen to produce maximum P_b concentration. The oxide was removed from the back of the sample to eliminate uncontrolled traps. A high dose of boron was implanted into the backside contact area of the wafers and activated during the oxidation to produce a p^+ layer for ohmic contact. The samples were metallized on the oxidized face, as shown in Fig. 1. The two small pads were provided for C-V measurements, in the range of capacitance most suitable for the apparatus used. A small area under the end with the C-V pads was metallized for good substrate contact. This was adequate for a satisfactorily low series resistance for the large pad, and it did not load the microwave cavity excessively. (The EPR cavity can ac-

 0021-8979/84/222844-06$02.40

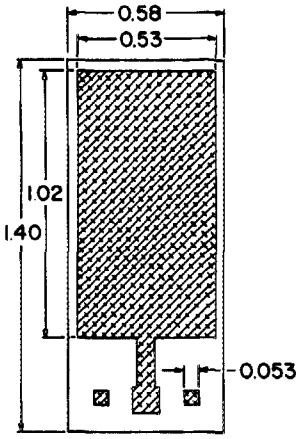

FIG. 1. Test chip with large-area MOS capacitor and small MOS pads for correlated EPR and C-V measurements. Crosshatched area is aluminum, 50 nm thick. Part of the rear surface (not shown) is boron-implanted and aluminized for substrate contact, as discussed in text. Dimensions are in centimeters.

commodate a single 1-cm^2 layer of Al$<$100 nm thick; but two layers, even $<$50 nm, separated by a wafer thickness, make the sample unusable.)

B. Electrical measurements

The C-V analysis of the sample properties serves two important functions: to obtain the band-gap distribution of interface trap levels D_{it} and to obtain the surface potential or band-bending as a function of gate bias for analysis of the EPR data.

Certain special and extreme sample properties determine the C-V approaches and assumptions to be used here. The high areal concentration of interface traps ($\sim 5\times 10^{12}$ cm^{-2}) dominates the interface capacitance over most of the band gap because of the relatively low doping ($>30\ \Omega$ cm); this obviates the need for dopant profile determination, and reduces surface carrier capacitance to a minor correction over most of the gap. On the other hand, the doping is not so low that series resistance distorts the potential applied to the large-area capacitor or complicates the analysis. A value of 30–150 Ω cm fulfills both requirements. Third, (and this a fortuitous circumstance, not an adjustable parameter) the trap response time is fast enough to allow low frequency (lf), rather than quasi-static, C-V determination of band-bending and trap density. This is a great advantage. At the trap densities observed, the C-V ramp must span at least a $\pm$ 100 V range. This corresponds to electric fields of 5×10^6 V/cm in the oxide, which is a severe stress, and often leads to breakdown spikes. The lf method recovers from these with little difficulty, while the quasi-static C-V method, which is sensitive to leakage currents even in the pA range, was found to be unreliable.

The low-frequency C-V curves were taken at 10 Hz, after it was determined that there was no detectable change in shape below 20 Hz. Sweep rate was generally 100 mV/s, although there was negligible change below 500 mV/s. The high-frequency (hf) C-V method was not found to be of great value in studying these samples, as it is incapable of wide-range surface potential determination[12]; and it provides no additional information on D_{it}. Its main uses were to check for series resistance, and for rapid qualitative comparisons and selection of best samples.

C. EPR measurements

The major difficulty in reproducible EPR measurements under applied gate voltage rests in poor signal-to-noise (S/N) ratio at low values of microwave power. A S/N ratio of 40 : 1 can be obtained in conventional EPR at 300 K on a typical stack of five double-sided samples, but multiple double-sided MOS wafers are unfeasible for EPR due to cavity loading. Low temperature, which can help the S/N ratio, is unacceptable for C-V electrical measurements, and therefore was not considered for EPR since it would confuse correlations between techniques. Another problem is EPR relaxation. For simply identifying the P_b signal, or measuring spin concentration for correlation with oxidation, annealing, orientation, etc., a microwave power of up to 10 mW may be used without great error. But, for precise determination of signal height as a function of small voltage differences—essential to determine a well-defined band-gap distribution—much lower power must be used to avoid saturation effects. The latter can lead to erroneous distributions ranging from wrong energies, at best, to spurious P_b levels, at worst. For these reasons, a power not greater than 0.2 mW was used. This reduced the EPR sensitivity by a factor of about five.

The low S/N ratio requires the use of a signal averager and a hundred or more repeat traces of the EPR signal at any one voltage for the desired precision. Dwelling at a given gate voltage for long times (i.e., hours) has its disadvantages: any slow and systematic drifts are emphasized, and the large MOS sample is more prone to breakdown if left for long periods at an extreme gate voltage. To avoid these problems, a continuously swept voltage was applied to the specimen; the spectrometer was set to sample the signal voltage at the peak or valley of the EPR line, with the magnetic field held constant. Gate sweep rate was lowered until no further change or hysteresis in the EPR sweep was seen; this was $<$500 mV/s, consistent with corresponding C-V tests on the adjacent pads. The peak and valley curves were then subtracted algebraically to obtain the EPR signal amplitude, while at the same time reducing possible drift or wide background signals.

III. OBSERVATIONS AND ANALYSIS

A. Electrical measurements of trap density

Observed lf and hf C-V curves for a p-type MOS sample are shown in Fig. 2(a). The n-type curves are only slightly different due to low dopant concentrations used. On first inspection, the low-frequency capacitance C_{lf} appears to be without structure. However, a 20$\times$ expansion of the plot reveals much of interest, Fig. 2(b). There is clearly much "area" in the wings of the curve—comparable to the central peak. Integration of C_{lf} over the entire range by the method of Berglund[12] yields an area of 1.1 eV, as it should. Then, plotting this integral (oxide capacitance $= C_{ox}$)

$$\psi_s = \psi_{so} + \int_{V_{Go}}^{V_G} \left[1 - \frac{C_{lf}(V_G)}{C_{ox}}\right] dV_G \qquad (1)$$

point-by-point from the extreme negative gate voltage V_{Go} yields the band-bending at any value of V_G. This latter curve is plotted in Fig. 3, with ψ_{so} set equal to 0 at the valence band

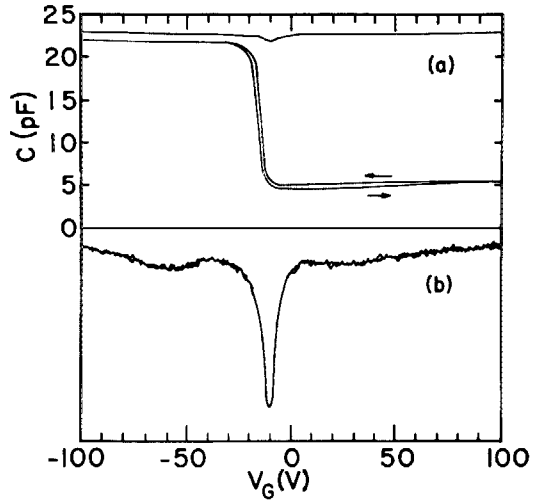

FIG. 2. (a) Low-frequency (10 Hz) and high-frequency (1 MHz) C-V curves for p-type MOS sample chip; (b) low-frequency curve expanded 20×.

edge, since we are using the plot only to establish the position of D_{it} and P_b features in the gap. The adjusted ψ_s then represents the band-gap position of the Fermi level at the interface, $E_F - E_{v,surf}$. Interface trap density is derived from the low-frequency curve by the relations[12]

$$C_{it}(V_G) = \left(\frac{1}{C_{lf}} - \frac{1}{C_{ox}}\right)^{-1} - C_s \tag{2}$$

and

$$D_{it}(V_G) = \frac{C_{it}(V_G)}{q_e}, \tag{3}$$

where C_{it} is capacitance due to interface traps, C_s is silicon surface capacitance from free carriers, q_e is unit electron charge. C_s was calculated approximately from published curves.[13] The derived trap density for a typical sample is plotted in Fig. 4. It shows two clearly defined peaks at energies $E_v + 0.26$ eV and $E_v + 0.84$ eV, with density maxima 2.2×10^{13} $eV^{-1}cm^{-2}$ and 2.0×10^{13} $eV^{-1}cm^{-2}$, respectively. The accuracy of the distribution declines very rapidly within 0.15 eV of the band edges, where the roughly determined C_s becomes greater than C_{it}. The total integrated trap concentration between 0.15 and 0.95 eV is 8.4×10^{12} cm^{-2}.

The C-V measurements were confirmed by comparison with previous DLTS measurements[6] on separate but identically oxidized n- and p-type wafers with properties optimized for DLTS. The D_{it} peak positions were at $E_v + 0.27$ eV and $E_v + 0.87$ eV with maxima $D_{it} = 1.4\times10^{13}$ $eV^{-1}cm^{-2}$ and 1.45×10^{13} $eV^{-1}cm^{-2}$, respectively, and a total integrated trap concentration of 8.0×10^{12} cm^{-2}. Thus, D_{it} distributions obtained by both C-V and DLTS on identically oxidized (111)-oriented Si wafers are in good agreement.

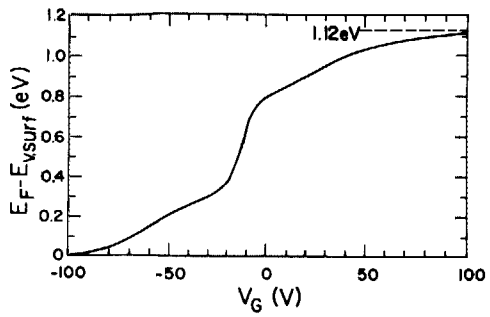

FIG. 3. Surface Fermi-level position E_F–$E_{v,surf}$ vs gate voltage V_G for p-type MOS test chip used for EPR study.

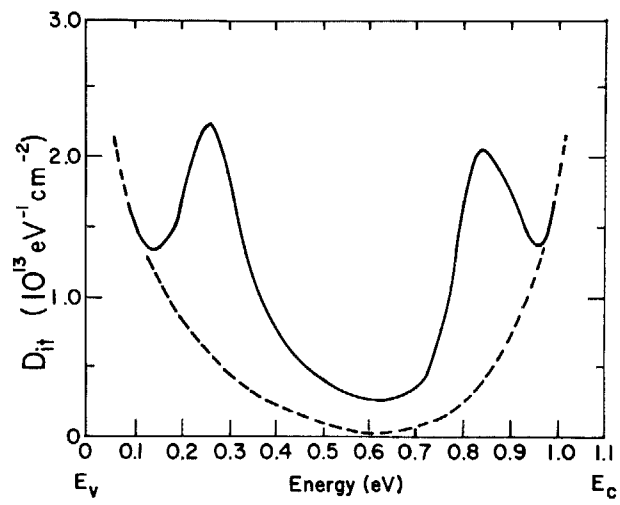

FIG. 4. Band-gap distribution of interface trap density D_{it} in as-oxidized p-type MOS sample measured by the low-frequency C-V technique. The solid line is the observed total D_{it}; the dashed line is a postulated continuum component of D_{it}.

B. EPR observations

The EPR peak-to-peak signal amplitude as a function of gate voltage is shown in Fig. 5 for three values of microwave power: 10, 1, and 0.1 mW. The misleading character due to excessive microwave power is evident. Note that the two apparent peaks at 10 mW have nothing to do with interface trap peaks; they are artifacts the appearance of which is governed by electron relaxation time. The most nearly correct curve is that at 0.1 mW; this one was analyzed in detail. No further change in shape was noted at 0.2 mW or below.

The EPR amplitude A, or spin density, as a function of Fermi-level position in the band gap is shown in Fig. 6. This curve was derived from the average of the swept-voltage curve of Fig. 5 and two-wide range, manually run curves, together with the band-bending function of Fig. 3. Unlike the D_{it} plots, this curve is actually the total number of singly

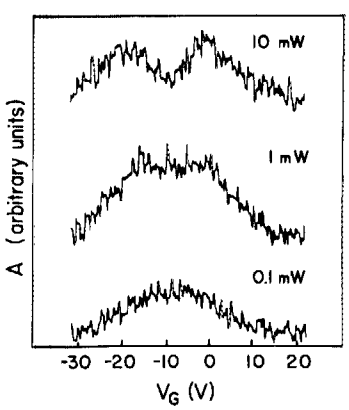

FIG. 5. EPR signal amplitude A vs gate voltage V_g in large-area p-type test chip for three different microwave powers. The spectra have been vertically displaced for clarity. The lower border of the drawing is the baseline for the 0.1-mW curve.

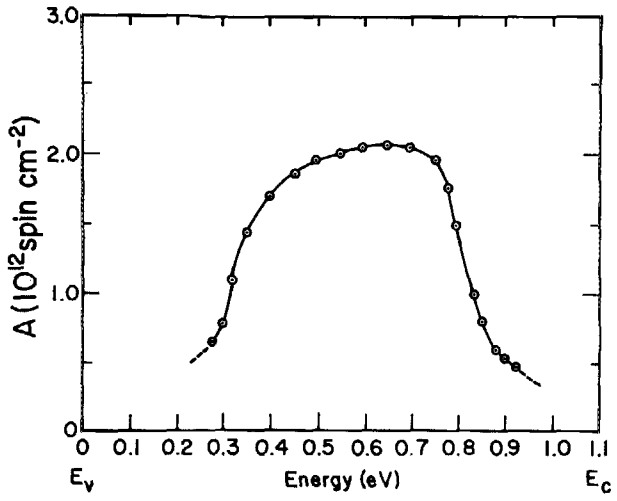

FIG. 6. Calibrated EPR signal amplitude A (spin density) vs Fermi-level position, derived from data in Figs. 3 and 5.

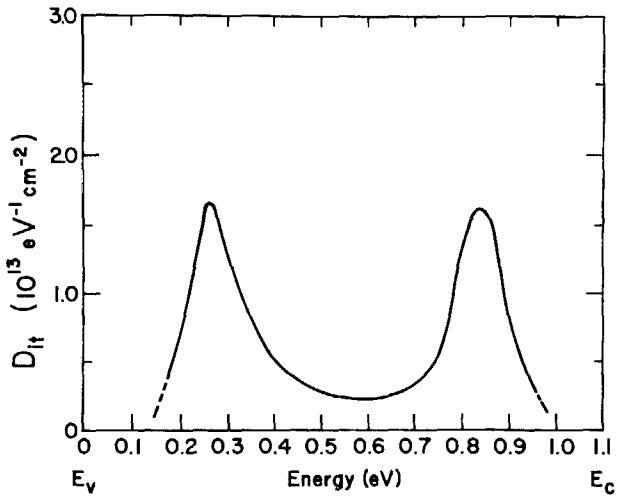

FIG. 8. D_{it} peaks deconvolved from observed total D_{it}, as discussed in text.

occupied P_b centers at any given band-bending, and thus is the integral of the distributed density of P_b levels up to the corresponding energy. The P_b *center* areal concentration is slightly more than 2×10^{12} cm^{-2}; the corresponding P_b *level* concentration is 4×10^{12} cm^{-2}, since each center can trap two electrons.

The absolute value of the derivative dA/dE yields the P_b level density. The derivative in the lower half of the gap follows the successive addition of one electron to each empty center, called the (0→1) electron transition. The centers go from positively charged to neutral. The derivative in the upper half of the gap follows the (1→2) electron transition; as a second electron is added to each center, the P_b centers become negatively charged. The second electron enters a spatial orbital already occupied by a previously trapped electron. Since two electrons in one orbital must have spins opposed, the EPR signal declines as more second electrons are added, hence the negative dA/dE in this region.

The P_b level distribution is plotted in Fig. 7. It shows maxima at $E_v + 0.31$ eV and $E_v + 0.80$ eV, with amplitudes 1.4×10^{13} eV^{-1} cm^{-2} and 1.7×10^{13} eV^{-1} cm^{-2}, respectively.

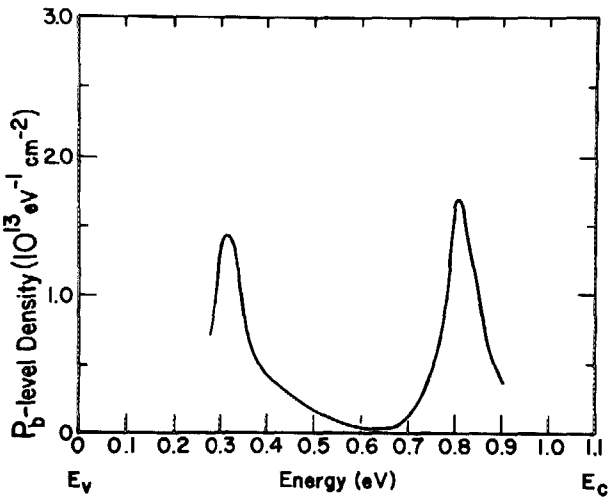

FIG. 7. P_b-level distribution in as-oxidized p-type MOS sample.

In previous studies, over a smaller range in the gap, the electron transition peaks have been assumed to be at the half-maximum points of the EPR amplitude plot. Thus, in Fig. 6 we find $E_v + 0.32$ eV and $E_v + 0.83$ eV, respectively, very nearly the same as the derivative peaks. When feasible, however, the derivative approach is preferred on principle. It is physically more meaningful; it yields numerical values for the peak amplitudes, and it obviates the need for baseline determination, which can be uncertain because of a spurious EPR signal from uncontrolled P_b centers.

The band gap positions of the P_b transitions correspond closely to the peaks of D_{it} observed by C-V and DLTS. Furthermore, the concentration of P_b levels (4×10^{12} cm^{-2}) may be reconciled with the concentration of electrical traps by consideration of other contributions to the observed total D_{it}.

C. Comparison of D_{it} and P_b peaks

The D_{it} distributon suggests the superposition of at least two sources of interface traps which, respectively, comprise the two peaks and an U-shaped continuum often ascribed to band-tail levels. This is in line with the annealing of integrated circuit MOS wafers by hydrogen, where it is observed that the peaks and a large part of the continuum gradually disappear; but ultimately, a low-density U-shaped background remains.[14,15] There is no established function to guide deconvolution, but a hypothetical continuum is shown in Fig. 4 by the dashed line. Subtraction of the continuum yields the net two-peaked D_{it} distribution in Fig. 8. The same deconvolution has been applied to the DLTS data previously reported[6] and yields a similar result.

No deconvolution is necessary for the EPR plot, since EPR is responding only to the paramagnetic singly occupied center associated with the two peaks; apparently the U-shaped continuum is not due to defects which are normally paramagnetic. The net C-V peaks and EPR peaks are similar. The pertinent parameters derived from the EPR and the deconvolved C-V and DLTS analyses are summarized in Table I. Although the deconvolution is mainly a working hypothesis at this time, the EPR and electrical results are seen to be in good quantitative agreement.

TABLE I. Measured parameters of electrical interface traps and P_b centers in oxidized (111) silicon wafers. The data refer to the two peaks in the band-gap distribution of the P_b signal observed by EPR, or deconvolved from the D_{it} distribution observed by C-V or DLTS, as discussed in text.

Parameter	Technique		
	C-V	DLTS	EPR
Defect concentration (cm^{-2})	2.7×10^{12}	2.4×10^{12}	2.1×10^{12}
Trap level concentration (cm^{-2})	5.3×10^{12}	4.7×10^{12}	4.1×10^{12}
(0↔1) peak position (eV)	$E_v+0.27$	$E_v+0.28$	$E_v+0.31$
(1↔2) peak position (eV)	$E_v+0.83$	$E_v+0.86$	$E_v+0.80$
(0↔1) max. density ($eV^{-1}\,cm^{-2}$)	1.6×10^{13}	1.1×10^{13}	1.4×10^{13}
(1↔2) max. density ($eV^{-1}\,cm^{-2}$)	1.6×10^{13}	1.1×10^{13}	1.7×10^{13}
Min. density position (eV)	$E_v+0.62$	$E_v+0.60$	$E_v+0.64$
Two-electron correlation energy, eV	0.56	0.58	0.49

IV. DISCUSSION AND CONCLUSIONS

A minor consideration (but suggestive for future studies) is the region just above midgap, where P_b, D_{it} distributions show a minimum. The D_{it} at this point is, typically, about 10% of the amplitude of the peaks. This midgap D_{it} value is used in monitoring of IC chips in manufacture. Furthermore, in some previous EPR studies,[1,2] midgap D_{it} has often been tacitly compared with EPR spin density, without specific justification. The midgap value appears to be proportional to the peak amplitudes of both EPR and D_{it} until the densities become very low after extended H_2 annealing. Whether this means that the dangling orbital P_b center distribution extends into the midgap region, or whether the midgap D_{it} is due to a portion of the U-shaped continuum, correlated with the peaks, cannot be decided at this time. It will be assumed here that the continuum states drop to near zero at midgap, and that there is a small overlap of (0↔1) and (1↔2) P_b levels, which gives rise to the midgap D_{it}. An overlap could be caused by fixed charge fluctuations, shifting P_b levels by varying amounts. This means that a minor fraction ($\lesssim5\%$) of the P_b centers are never seen by EPR, and this estimate was used in data analysis (although it is essentially negligible).

Although a small point in analysis of the P_b and D_{it} peaks, the origin of the midgap D_{it} is obviously an important consideration in seeking to establish the origin of the U-shaped continuum traps. The behavior of D_{it} peaks and continuum traps during H_2 anneal and postmetallization anneal implies that hydrogen can attack more than one source of D_{it}. Perhaps the most likely other defects to be attacked are distorted back-bonding orbitals associated with $\cdot Si{\equiv}Si_3$. Presumably, annealing of the dangling P_b orbital by H would relax the P_b center and its closest neighbors into a more normal bonding configuration, thus shifting their orbital energies toward the normal valence and conduction band values of the perfect lattice, and thereby reducing deep traps.

Dangling silicon orbital and band-tail states in amorphous silicon (α-Si:H) have been intensively studied by EPR and optical EPR, and reasonable models have been derived.[16,17] Preliminary optical[18] and optical EPR[19] studies of the Si/SiO_2 interface are promising, although the outcome is not predictable at this time. Note, however, that the results in the present paper set the two-electron correlation energy at slightly greater than $+0.5$ eV, comparable to $+0.4$ eV for α-Si:H.[17]

In conclusion, P_b centers have now been shown to have a near-perfect 1 : 1 correspondence to interface traps in a variety of experiments. The gross quantitative correspondence has previously been observed for three main Si wafer orientations (111), (110), (100)[20]; over several different oxidation and thermal annealing procedures, including variable oxidation time[1,2]; in the Deal oxidation triangle[9]; in H_2 anneal and N_2 restoration of traps[1]; and during defect generation by radiation.[21] The declines of P_b density and either D_{it} peak were found to track closely in variable-temperature, postmetallization anneals.[6] In the present paper, it has been shown that the detailed band-gap energy spectra of P_b levels and interface traps correspond qualitatively and quantitatively. Thus, it has now been clearly demonstrated that P_b centers are the source of a large part of the interface traps in as-oxidized silicon wafers, and, specifically, are the source of the two main D_{it} peaks at $E_v+0.3$ eV and $E_v+0.8$ eV in (111) wafers.

ACKNOWLEDGMENTS

The authors are pleased to thank M. D. Moyer for technical support. The work at the Xerox Corporation was supported by the U.S. Army ERADCOM.

[1] P. J. Caplan, E. H. Poindexter, B. E. Deal, and R. R. Razouk, J. Appl. Phys. **50**, 5847 (1979).
[2] E. H. Poindexter, P. J. Caplan, B. E. Deal, and R. R. Razouk, J. Appl. Phys. **52**, 879 (1981).
[3] E. H. Poindexter, P. J. Caplan, J. J. Finnegan, N. M. Johnson, D. K. Biegelsen, and M. D. Moyer, in *The Physics of MOS Insulators*, edited by G. Lucovsky, S. T. Pantelides, and F. L. Galeener (Pergamon, New York, 1980), pp. 326–330.
[4] C. Brunstrom and C. Svensson, Solid State Commun. **37**, 399 (1981).
[5] P. M. Lenahan and P. V. Dressendorfer, Appl. Phys. Lett. **41**, 542 (1982).
[6] N. M. Johnson, D. K. Biegelsen, M. D. Moyer, S. T. Chang, E. H. Poindexter, and P. J. Caplan, Appl. Phys. Lett. **43**, 563 (1983).
[7] P. M. Lenahan and P. V. Dressendorfer, Appl. Phys. Lett. **44**, 96 (1984).
[8] J. E. Wertz and J. R. Bolton, *Electron Spin Resonance: Elementary Theory and Applications* (McGraw Hill, New York, 1972).
[9] E. H. Poindexter and P. J. Caplan, Prog. Surf. Sci. **14**, 201 (1983).
[10] E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982), Chaps. 2–4, 8, 10, 12.

[11]N. M. Johnson, J. Vac. Sci. Technol. **21**, 303 (1982).
[12]E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982), pp. 325–331, 489.
[13]E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982), p. 87.
[14]R. R. Razouk and B. E. Deal, J. Electrochem. Soc. **126**, 1573 (1979).
[15]N. M. Johnson, D. K. Biegelsen, and M. D. Moyer, in *The Physics of MOS Insulators*, edited by G. Lucovsky, S. T. Pantelides, and F. L. Galeener (Pergamon, New York, 1980), pp. 311–315.
[16]D. K. Biegelsen, in *Nuclear and Electron Spectroscopies Applied to Materials Science*, edited by E. N. Kaufmann and G. K. Shenoy (Elsevier-North Holland, Amsterdam, 1981), pp. 85–94.
[17]H. Dersch, J. Stuke, and J. Beichler, Phys. Status Solidi B **105**, 265 (1981).
[18]N. M. Johnson, W. B. Jackson, D. K. Biegelsen, and M. D. Moyer, Bull. Am. Phys. Soc. **29**, 339 (1984).
[19]B. Henderson, Appl. Phys. Lett. **44**, 228 (1984).
[20]E. H. Poindexter and P. J. Caplan, in *Insulating Films on Semiconductors*, edited by M. Schulz and G. Pensl (Springer, Berlin, 1981), pp. 150–160.
[21]P. M. Lenahan, K. L. Brower, P. V. Dressendorfer, and W. C. Johnson, IEEE Trans. Nucl. Sci **NS-28**, 4105 (1981).