Role of hydrogen in the formation of metastable defects in hydrogenated amorphous silicon

W. B. Jackson, J. M. Marshall,* and M. D. Moyer
Xerox Corporation, Palo Alto Research Center, Palo Alto, California 94304
(Received 2 May 1988; revised manuscript received 31 August 1988)

This paper presents results of studies on carrier-induced metastable defect creation in hydrogenated amorphous silicon. The metastable defects were studied by measuring the threshold voltage shifts on thin-film amorphous silicon transistors and capacitors as a function of time, temperature, and bias. The kinetics (time, temperature, bias, and doping dependence) of these defects as well as most other metastable-defect processes are quantitatively explained by hydrogen diffusion and the creation of defects due to the presence of excess band-tail carriers.

I. INTRODUCTION

The origin of metastable-defect kinetics in hydrogenated amorphous silicon (*a*-Si:H) is an intriguing problem. The defects can be created by a variety of different means including illumination,[1] current,[2,3] electric fields,[4] doping,[5] and rapid thermal quenching,[6] and can be removed by annealing at temperatures above 150 °C. Despite considerable effort as well as great technological importance, the microscopic mechanisms responsible for these metastable changes are not well understood. In this paper, we concentrate on carrier-induced defect kinetics and show that the observed results for carrier-induced as well as many other metastable defects are quantitatively consistent with models involving hydrogen motion. Because the kinetics of carrier-induced defects are similar to those caused by other factors, the results support the unified picture that hydrogen is involved in the kinetics of all the observed metastable defects.[7]

Much of the previous work on metastable defects has focused on an understanding of light-induced defect formation known as the Staebler-Wronski effect, and several models have been developed to explain the defect-creation process. One of the most widely accepted classes of models involves the breaking of weak Si—Si bonds and subsequent stabilization of the newly created Si dangling-bond defects.[2,8–13] One significant uncertainty of the weak-bond models concerns the role of band-tail carriers. All the various forms of metastable and doping-induced defect formation involve the presence of band-tail carriers—either electrons or holes. Making the assumption that a single mechanism underlies all the various defect processes, we can conclude that the band-tail carriers must be involved in the destruction and formation of deep defects. The connection between band-tail carriers and reformation of bonds has been experimentally investigated[14] and has been attributed to changes in energetics of bonds due to the modified "$8-N$" rule.[15] Further elaboration of the model and the energetics has been discussed in Refs. 16–20. The band-tail carrier is hypothesized to destabilize weak silicon-silicon bonds with subsequent formation of dangling-bond defects. Studies of doped material have shown that in addition to changes in the number of defects, the number of active dopants also changes in response to excess carriers induced by light or rapid thermal quenching.[21,16] Consequently, the mechanisms for rearrangements are not necessarily unique to silicon and can involve dopant atoms as well.

A second uncertainty of the weak-bond model involves the mechanism by which these electronically induced defects and bond rearrangements are stabilized. The most widely proposed idea is that hydrogen motion is involved in the kinetics and stabilization of the metastable defects.[9–13,22] The basic model is that hydrogen, bonded to Si, is either released interstitially with subsequent trapping or switches positions to create and anneal dangling bonds and active dopants.

A number of observations suggest that hydrogen plays an important role in metastable changes in a *a*-Si:H. First, all the metastable changes anneal at temperatures between 150 °C and 230 °C. At these temperatures, hydrogen is known to diffuse quite readily through the network.[23,24] Second, annealing temperatures of greater than 420 °C are required to reduce the large spin densities found in unhydrogenated films and only recrystallization temperatures of over 600 °C reduce the defect densities to levels found in hydrogenated amorphous silicon.[25,26] If hydrogen did not play a role, unhydrogenated films should exhibit similar defect kinetics at low temperatures. Third, the motion of hydrogen atoms provides a natural mechanism for stabilizing metastable defects against recombination because the hydrogen is mobile and is not sterically constrained by more than one bond. On the basis of these considerations, a number of models have been proposed invoking hydrogen motion as an active participant in various individual metastable-defect phenomena.[7,9–13,21,24]

Additional support for the involvement of hydrogen has come from more detailed studies of the hydrogen diffusion in *a*-Si:H. Doping has been found to greatly affect the diffusion rate of hydrogen in both crystalline[27] and amorphous silicon.[24] Boron doping, in particular, increases the diffusion coefficient by nearly three orders of

 ©1989 The American Physical Society

magnitude while phosphorus doping results in nearly an order of magnitude increase. Metastable-defect kinetics indicate a similar dependence on doping, i.e., defects in boron-doped material anneal faster than those in phosphorus or undoped material.[28] Whether the increased hydrogen diffusion causes the metastable effects or the same mechanism results in enhanced hydrogen diffusion is unresolved. Furthermore, the hydrogen motion has been found to be dispersive exhibiting a time dependence of the diffusion coefficient similar to carrier transport in amorphous materials.[24] This dispersive hydrogen motion accounts for the stretched exponential decays of excess carriers in doped amorphous silicon,[22] the annealing of light-induced defects,[29] and the carrier-induced creation of defects.[30] Dispersive hydrogen motion also accounts for the observed Meyer-Neldel–type relation between the activation energy of annealing for metastable defects and the attempt-to-anneal factor.[31] An important testable consequence for the involvement of hydrogen in metastable processes is to compare the kinetics of defect formation under carefully controlled conditions with that expected from hydrogen diffusion.

In an independent development, significant progress has been made in understanding the origin of changes in the electrical characteristics of metal-insulator–amorphous-silicon (MIS) structures. Initial measurements attributed changes to trapped charges within the silicon-nitride layer used as the insulator in the MIS structure.[32] Time-of-flight measurements found evidence for both fast states (interface states) and bulk states on studies of nitride overlayers on silicon.[33,34] Because the deposition of nitride on silicon (top nitride) results in significantly worse devices than in the case of deposition of amorphous silicon nitride (bottom nitride), the causes of threshold shift for devices which use bottom nitride may be completely different. More recent photoconductive measurements on devices before and after stressing have been interpreted in terms of the creation of interface states.[35] The evidence relied on the understanding of the details of recombination under illumination in the device. Recent work on MIS structures[30,36,37] in which the insulator was replaced by a thermal gate oxide (an insulator known to have negligible trapping[38]) demonstrated that the defects are in fact created in the amorphous silicon within 10 nm of the interface. The metastable effects are, therefore, due to changes in amorphous silicon rather than in the insulator. The MIS structure is an ideal system to study metastable-defect formation because one can control the carrier density and the electrical measurements are extremely sensitive to small defect changes.

In this paper, the metastable-defect-formation kinetics for carrier-induced defects in MIS structures (capacitors and transistors) are studied in detail and compared to the results expected from hydrogen diffusion and to the kinetics of other metastable-defect processes. We find evidence consistent with the hypothesis that hydrogen diffusion is actively involved in metastable-defect formation in *a*-Si:H. The metastable-defect kinetics as a function of time, temperature, and carrier concentration are in quantitative agreement with models involving dispersive hydrogen motion. Because the kinetics of these defects are similar to those found for other metastable defects, the results provide evidence that hydrogen motion is involved in many of the metastable processes observed in *a*-Si:H.

The organization of the paper is as follows. Section II discusses the sample structures, the experimental procedure, and a brief review of the experimental techniques used to obtain the data. In Sec. III, results measured on these samples are presented. In Sec. IV, the defect kinetics expected from the stretched exponential equation involving dispersive hydrogen diffusion are derived. These relations are compared with the experimental time and temperature kinetics of carrier-induced defect creation of Sec. III. The stretched exponential equation also predicts a particular relation between hydrogen diffusion and the decay time of metastable changes which is verified by comparison to experimental data from a wide variety of metastable changes. Finally, possible physical origins of the stretched exponential equation are discussed.

II. EXPERIMENT

The generation, measurement, and annealing of carrier-induced metastable defects was studied on a variety of different MIS structures subjected to various temperatures, biases, and doping conditions. The defects were created by the application of a field, and the changes in the trapped charge density in these new states were determined by monitoring the shift in the threshold of either a capacitance-voltage (C-V) curve for a capacitor or of the source-drain current (I_{sd}) versus gate voltage (I-V) curve for a transistor. In this section, the principles behind the C-V and I-V measurements are briefly reviewed followed by a discussion of the specific sample structures used in this study, the measurement procedure used to determine the threshold-voltage shifts, and finally, the stressing sequences use in this study.

A. Sample structures

The results were obtained from a number of different devices. Many of the measurements were made on capacitors in which the gate electrode consisted of degenerately P-doped crystalline silicon. The insulator was either 30 or 75 nm thermal oxide followed by 500 nm of either undoped or P-doped *a*-Si:H deposited by glow discharge at a substrate temperature of 230 °C. The top contact was a 50–100-nm n^+-type layer topped with a 1-mm-diam aluminum pad defining the capacitor area (structures similar to those used in Refs. 30 and 36). In addition, capacitors and transistors were fabricated using silicon nitride as the dielectric in structures shown in Figs. 1 and 2. Initially, Cr was patterned for the bottom gate electrode. Plasma-deposited silicon nitride (300-nm thickness) followed by *a*-Si:H (100–500 nm) was deposited without interruption of the plasma (for further details see Ref. 39). In the case of the capacitors, a 60–150-nm n^+-type layer was also deposited followed by either a 100-nm-thick Al or an indium–tin-oxide semitransparent top contact. The layers were patterned into a structure as indicated in Fig. 1. In the case of the transistor, a pas-

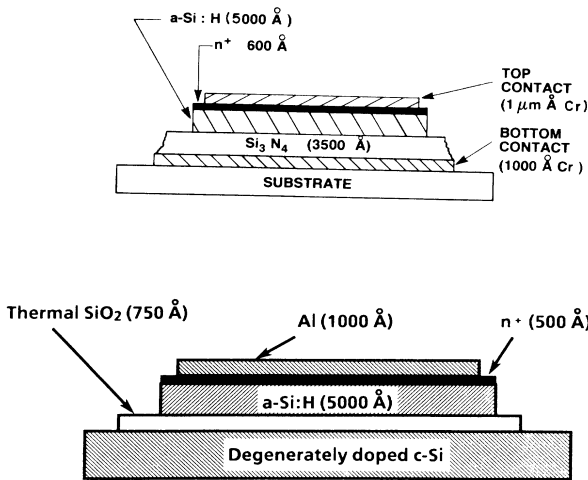

FIG. 1. The structure for capacitors used in this study. (a) nitride; (b) thermal oxide.

sivating layer of silicon nitride was deposited over the undoped *a*-Si:H. Holes were opened up in the nitride for the source and drain contacts, and then a 120-nm layer of n^+-type *a*-Si:H and Al layer were patterned to complete the fabrication. Areas of the capacitors ranged from 10^{-2} to 10^{-4} cm^2. In some cases transistors and capacitors were fabricated from the same silicon-nitride–amorphous-silicon deposition to permit comparison of the various devices.

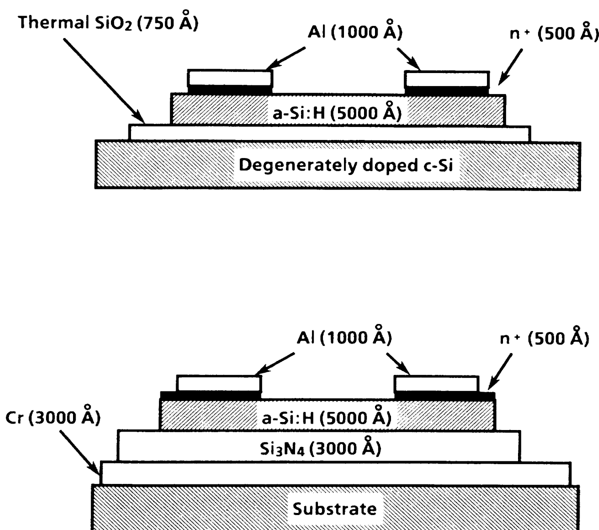

FIG. 2. The structure for transistors used in this study. (a) nitride; (b) thermal oxide.

B. Capacitance- and current-voltage measurements

C-V and *I-V* measurements are useful techniques for determining trapped charge and defect densities in MIS structures. In particular, information about the density and approximate location within the band gap can be determined. If the capacitance is measured as a function of voltage on MIS structure with n^+-type contacts shown in Fig. 1, the capacitance ranges from the geometrical capacitance (capacitance of the whole structure) to a maximum value given by the insulator capacitance.[40] Changing the applied voltage from negative to positive voltages applied to the gate sweeps the Fermi level in the bandgap at the interface from lower to higher energies (towards the conduction band). The band bending in the semiconductor changes from electron (hole) depletion (accumulation) to accumulation (depletion).[30] If the contact can inject holes and a hole-accumulation layer forms, the capacitance rises to the insulator capacitance for increasingly negative biases applied to the gate as well as for positive bias. Information about the lower half of the band gap can therefore be obtained. In the usual case when the contacts block holes, e.g., an n^+-type contact, the holes cannot respond to the oscillating field of the capacitance measurement so the capacitance remains equal to the geometrical capacitance for negative biases (Fig. 3). Newly created negatively charged defects in the lower half of the band gap or charge trapped in the insulator result in a parallel shift of the *C-V* curve to more positive

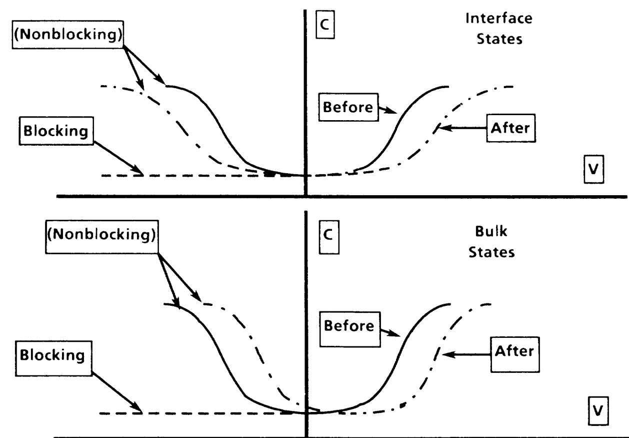

FIG 3. Effects of interface-state creation and bulk trapping on the *C-V* curves. The top figure shows the effect of additional interface states which cause a stretchout of the *C-V* curve. If the contact blocks holes, then the negative-voltage portion is not observed. Hence, if interface states are produced in the lower half of the gap, the positive-voltage portion of the *C-V* curve exhibits a parallel shift as observed in Fig. 5. The bottom figure shows the effects of bulk trapping within the nitride. There is a parallel shift of both the positive and negative portions of the *C-V* curve. Results of Ref. 30 have shown that positive stressing results in the situation represented in the upper curve.

gate voltages (Fig. 3). The amount of trapped charge per unit area, ΔQ, is given by the relation $\Delta Q = C_i \, \Delta V$, where C_i is the insulator capacitance per unit area and ΔV is the threshold-voltage shift. The areal density of defects created is therefore directly proportional to the threshold-voltage shift for parallel shifts. If states are created in the upper half of the gap, their occupancy changes during the voltage sweep causing the *C-V* curve to stretch out along the voltage axis. The rate at which the capacitance increases as a function of voltage is therefore a measure of defect density through which the Fermi level E_F sweeps during the measurement. The value of capacitance also indicates the approximate position of E_F in the gap. Changes of the *C-V* curves near the geometrical (insulator) capacitance are indicative of deep (shallow) state changes. In summary, voltage stretchout indicates interfacial defects in the upper part of the gap while parallel shifts indicate charged defects in the lower portion of the gap or charge trapping in the insulator.

The above discussion described the results for *C-V* curves obtained on capacitors, but the same information and considerations apply to transistors and *I-V* curves taken with small source-drain voltages to avoid pinch-off. The transistor has the additional property that the Fermi-level position at the interface can be determined by measuring the activation energy of the source-drain current. Electrical studies of both transistors and capacitors provide important information on the creation of interface states and charge trapping.

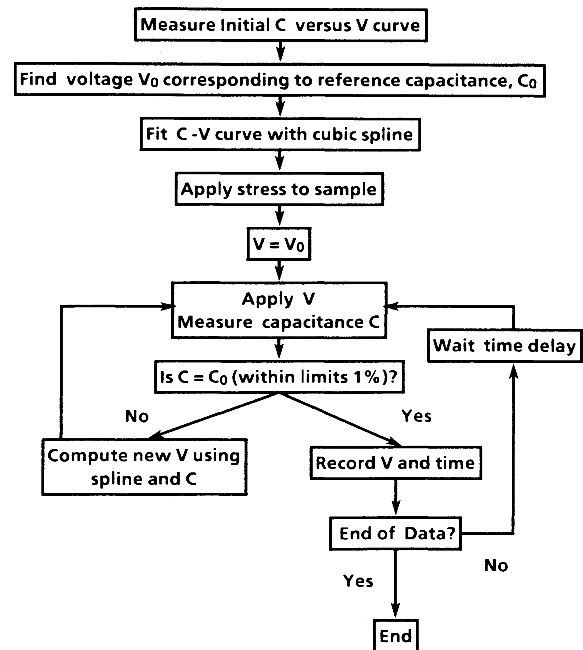

FIG. 4. Flow chart for the procedure to measure rapidly the threshold voltage shift as a function of time. At the end, a final *C-V* curve is measured and compared to the initial curve to ensure that a parallel shift is observed.

C. Measurement procedures

The experimental procedure was (1) to measure either the *I-V* or *C-V* characteristics, (2) apply a stressing sequence to the gate, (3) periodically interrupt the stressing to determine the threshold shift, (4) apply a relaxation sequence to the gate, (5) periodically interrupt the relaxation to determine the threshold, and (6) measure a final *I-V* or *C-V* curve. The stressing sequence used to generate the defects consisted of dc voltages in the range of -25 to $+25$ V depending on the insulator thickness. The sample temperature could be altered so the effects of temperature could be determined. The transistors and capacitors could also be illuminated using a tungsten lamp. The *C-V* measurements could be taken at any frequency from 5 Hz to 10 MHz although most results were taken at 100 kHz. The rate at which the *C-V* and *I-V* curves were obtained could be varied from 100 ms per point to a much longer time period. All results in this work were obtained using roughly 4 sec per point during which an average of five readings were computed. The experiment was computer controlled to permit the required long observation times (additional details may be found in Ref. 30).

A crucial requirement of the experiment is that the time for the threshold shift measurement be as short as possible to minimize the disruption of the stressing and relaxation cycles. A flow chart describing the measurement procedure is presented in Fig. 4. After an initial *C-V* or *I-V* curve is determined, a reference capacitance (C_0) or current (I_0) is selected. The electrical characteristic is fitted with cubic spline regression, and the voltage corresponding to the reference value is obtained (V_0). The sample is subsequently stressed for a specified time. The previous threshold voltage is applied to the sample and the capacitance or current is measured. If $C = C_0$ or $I = I_0$ within a preset limit, the current calculated voltage is the position of the threshold voltage. Otherwise a new voltage is calculated and applied to the sample and the cycle repeats itself. The stress (or relaxation) is reapplied to the sample until the next point is measured. Typically, the cycle converges within one cycle and takes less than 0.3 sec. This procedure minimizes the effects of the threshold measurement on the charging-discharging process and permits extremely small threshold shifts (<5 meV) at short times to be determined. The procedure relies on the assumption that the *C-V* or *I-V* curve exhibit only a parallel shift. A final *C-V* or *I-V* curve is compared to the initial one to ensure that this assumption is valid.

III. RESULTS

A. *C-V* and *I-V* measurements

Typical *I-V* measurements on a nitride transistor are depicted in Fig. 5 for various measurement temperatures. The rise in current for V_g near 2 V indicates that the Fermi level within the channel moves close to the conduction band for these gate biases. The increase in current for negative voltages is due to the increased carrier concen-

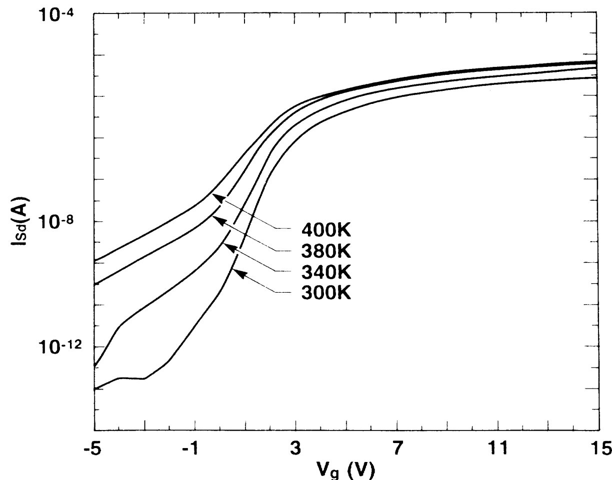

FIG. 5. The source-drain current for a nitride transistor vs gate voltage for different temperatures.

tration caused by the higher temperatures. Since the carrier density is known to be important for the defect creation, these measurements provide important information about the Fermi-level position as a function of gate bias. The Fermi-level position is determined by plotting the current at a given V_g versus $1/kT$. The activation energy is a rough measurement of the position of the Fermi level relative to the conduction band for various gate biases. Such activation-energy plots reveal that the Fermi level within the channel moves to within 0.09 eV of the conduction-band edge (Fig. 6). Hence, the carrier densities for 10–15-V gate biases correspond to heavy phosphorus doping of 10^{-2} in the gas phases while gate biases around 3 V correspond to lighter phosphorus-doping levels of 10^{-5}–10^{-4}. This correspondence between the various doping levels and V_g will be used in Sec. IV to compare the kinetics at different gate biases with those for doping levels giving comparable carrier densities (as determined by the Fermi-level position).

We consider next C-V measurements and the effect of various levels of dc gate-bias stressing. A typical C-V curve before and after a 20-min bias at 20 V is presented in Fig. 7 for a nitride capacitor. The change is primarily a parallel shift with a very slight change in shape indicating that the dominant process is the addition of charge in the dielectric or at the interface. The sign of the charge trapped is negative since the curve shifts to the right. Such a charge could lie either within the bulk of the dielectric or in the a-Si:H near the interface in states whose occupancy does not change during the measurement. These states are located at midgap or in the lower portion of the gap for the following reason. If the states were near the conduction band, their occupancy would change during the C-V measurement which moves the Fermi level through the upper part of the gap at the interface. The change in occupancy would result in a stretchout of the C-V curve. When a negative stress is applied to the gate, the C-V curve shifts to the left, indicating that additional positive charge or holes are now trapped within the nitride or interface.

B. Interface versus bulk

It is important at this point to establish that for positive gate biases, the defects are created within the amorphous silicon near the interface. There is a number of different experiments in this study, as well as in previous work, which supports this suggestion. In Ref. 35, it was

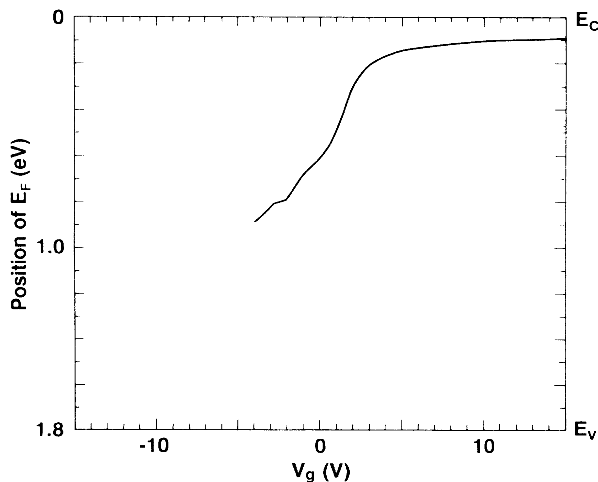

FIG. 6. The activation energy obtained from the data in Fig. 7 vs gate voltage. The conduction-band edge (E_c) is at 0 V while the valence-band edge (E_v) is roughly at 1.8 eV.

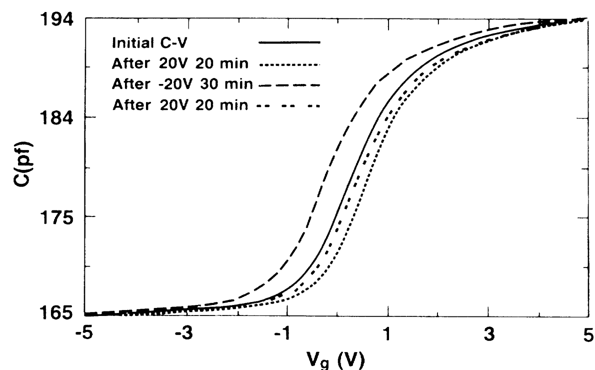

FIG. 7. The effect of various stressing conditions on the dependence of the 100-kHz capacitance vs bias applied to the gate (positive applied to the bottom metal in Fig. 1). The C-V curve initially shifts to the right, indicating the trapping of negative charge. The negative bias moves the curve to the right, indicating a net trapping of positive charge. Finally, the curve is returned roughly to its original state. The parallel shift indicates that no interface states in the upper part of the band gap are created.

shown that the induced charge could be neutralized by photodischarge in the silicon quite rapidly, indicating that the charge is in close proximity with the *a*-Si:H. Recent work on ambipolar transistors in which the channel current is carried by electrons and holes came to the same conclusion that positive gate biases generated interface states while negative biases resulted in charge exchange with the nitride.[30,36,37,41] Changes in the illuminated *I-V* characteristics have shown that for electron accumulation in nitride samples, states are created at or near the interface. In Fig. 8, dark and illuminated *I-V* characteristics before and after stressing are shown. The rise of the curve for negative voltages under illumination arises from conduction by holes and depends on the position of the Fermi level in the lower part of the gap. The stretchout of the illuminated-voltage characteristics demonstrates that the charge interchanges with the silicon during the *I-V* measurement. The same studies also indicate that hole accumulation, on the other hand, results in holes trapped within the nitride[30] due to the different offsets for valence and conduction bands.[42]

Additional corroborating evidence for the importance of the interface comes from studies of the effect of deposition on the rate of charge generation. The rate of charging is improved up to a certain level by changes in the nitride-deposition conditions. Beyond this level, changes in the nitride deposition, which significantly change its structure such as the Si—H bonding, do not result in reduced charging; the charging rate only increases.[43] Other indications that the interface is important come from the observation that the rate of the threshold shift is greatly increased by alteration of the growth conditions at the interface such as interrupting the plasma or striking a hydrogen plasma before silicon deposition.[43] Since the nitride is deposited first, the bulk nitride properties can not depend on plasma conditions during deposition of the interface. Consequently, the large increase in charging must reflect the changes at the interface. The rate of threshold shift also decreases if the silicon-deposition temperature is increased while keeping the nitride-deposition conditions fixed.[43,44] Hence, there is significant evidence from alteration of deposition conditions indicating that for the best nitride, the interface defect creation dominates the threshold-voltage shift.

Next, we present data from thermal oxide MIS structures indicating that the defects induced by bias stressing actually occur in the amorphous silicon rather than within the dielectric near the interface. Thermally grown oxide on *c*-Si has an extremely low trapping rate due to the large band gap and small numbers of traps.[38] Applications of 15 V stress to similar crystalline-silicon-thermal-oxide structures at elevated temperatures for durations of several days exhibited threshold shifts of less than 50 meV. Hence, charge trapping within the SiO_2 is negligible. The time dependence of the threshold-voltage shifts of an amorphous-silicon–oxide capacitor is virtually identical to the time dependence observed for an amorphous-silicon–nitride device for positive gate biases (Fig. 9). This is evidence that the states are not within the dielectric since both the number and trapping characteristics of states of a plasma-deposited nitride and a thermal oxide would be expected to be different. Furthermore, as will be discussed in Sec. IV, the time, temperature, and bias dependencies of the threshold-voltage shifts are entirely consistent with equivalent metastable effects which occur in the bulk amorphous silicon. The

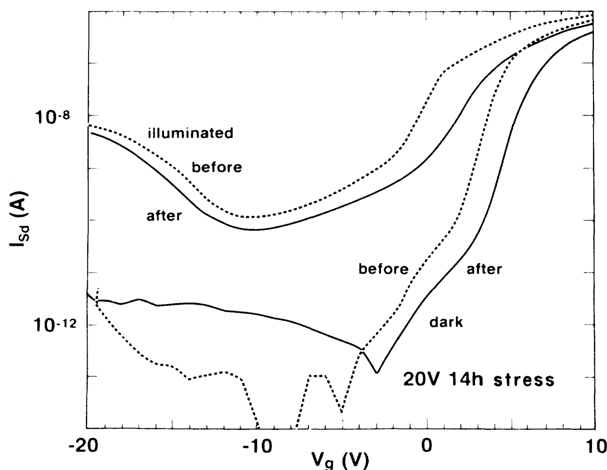

FIG. 8. The source-drain current vs gate voltage before (dashed) and following (solid) a 20 V, 14-h stressing time. The measurements were made under illumination and in the dark. The increase in current under negative gate biases for the illuminated case is due to hole conduction. The stretching of the illuminated curve along the voltage axis is indicative of interface-state creation.

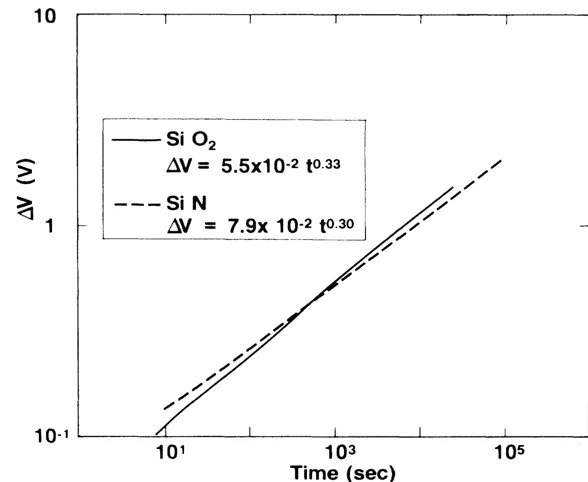

FIG. 9. A comparison of the threshold-voltage shift ΔV vs time for a silicon-nitride capacitor and a thermal-oxide capacitor. The shift is virtually identical, indicating that the threshold shifts are due to changes occurring within the *a*-Si:H rather than charge trapping within the nitride, or are characteristic of a special interface region between the silicon nitride and the *a*-Si:H.

defects are created within 100 nm of the interface because the accumulation layer in which the carrier density is high is less than 100 nm wide. Thus, positive biases which are less than 30 V primarily create interface states, and the density is proportional to the threshold shift. We can therefore determine defect kinetics in amorphous silicon through observation of the threshold shifts.

Consistent with above results, negative biases applied to oxide MIS structures are significantly different from those for silicon-nitride MIS structures. A negative bias applied to an oxide capacitor results in a *positive* threshold shift rather than the negative shift observed for nitride samples. This result indicates that states are created near the oxide interface which become negatively charged during the *C*-*V* measurement and cause a positive threshold shift. There is no significant bulk trapping in thermal SiO_2. For nitride films, on the other hand, negative biases cause significant densities of holes to become trapped within the bulk nitride. Hence, the threshold shifts to negative voltages confirming previous results.[30,37,41]

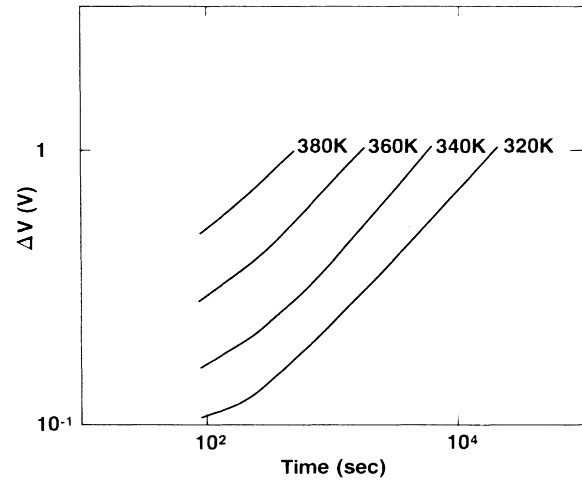

FIG. 11. The threshold-voltage shift transistors subjected to 20 V stress vs time for various temperatures. Each curve was obtained on a different device.

C. Threshold voltage versus time

A typical threshold-voltage shift versus time is depicted in Fig. 10. The nitride was stressed with 20 V dc bias during the first part of the cycle. Subsequently, the gate voltage was set to zero and defect states induced by the gate voltage began to decay. Note that the creation rate of metastable defects is much more rapid than the annealing rate. Consequently, even if the dielectric remains unstressed for a long period of time, the charge will not be removed from the dielectric. In addition, metastable-defect formation and annealing is not exponential with time. Subtracting the initial voltage yields the change in the threshold, ΔV, as a function of time. Plotting $\log_{10}\Delta V$ versus $\log_{10}t$ yields a straight line for most of the samples irrespective of the temperature or type of device (transistor or capacitor). In other words, the charging time dependence is given by a power law of the form

$$\Delta V = \Delta V_0 t^{\beta} , \quad (1)$$

where ΔV_0 is the shift after 1 sec and β is an exponent.

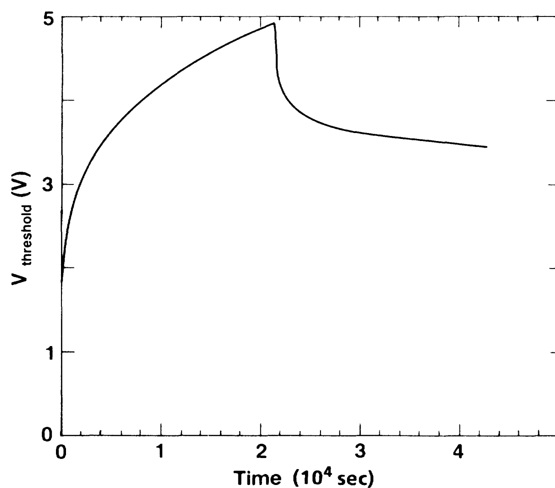

FIG. 10. The threshold of a capacitor vs stressing time. During the period from 0 to roughly 2×10^4 sec, 20 V dc was applied to the gate. After this period, the voltage was reduced to 0 V and the discharge was determined. The curve consists of hundreds of points demonstrating the high signal to noise ratio.

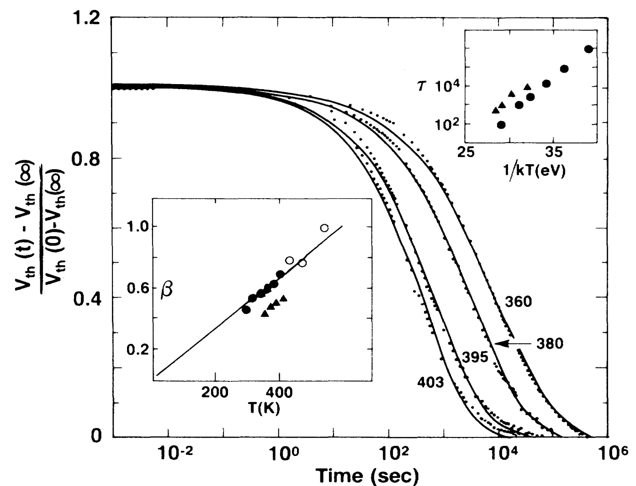

FIG. 12. The normalized threshold-voltage shift vs time for various indicated temperatures (K). The data are represented by points and the solid lines are least-squares fits to Eq. (10). The fitting parameters τ and β are represented in both insets by the solid triangles. The solid circles represent data obtained from fits to decays of excess band-tail carriers in doped *a*-Si:H from Ref. 22. The open circles represent β values obtained from hydrogen diffusion from Refs. 22 and 24.

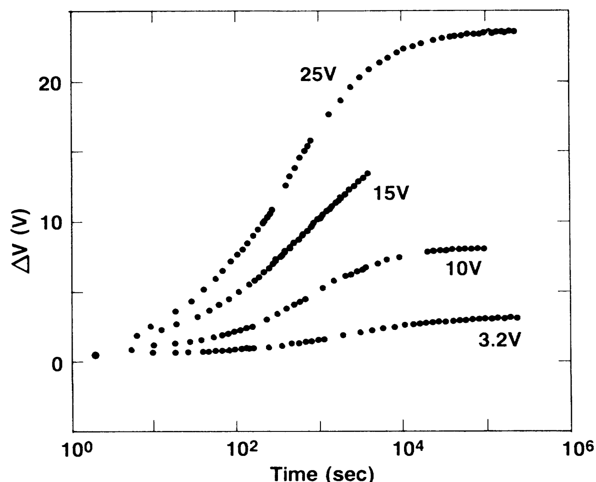

FIG. 13. The threshold-voltage shift for nitride transistors subjected to the various stressing voltages indicated vs time. Each curve was obtained on a different device.

An example of this behavior is plotted in Fig. 11 for different temperatures. Note the linear shape and equal spacing in log time for decreasing temperatures, indicating that β does not change greatly with temperature, while ΔV_0 has a significant temperature dependence. The values of β obtained from least-squares fits are roughly 0.5 ± 0.1 with some small increase with temperature. Equivalently, for these samples the threshold-voltage shift varies approximately as the square root of the charging time. The parameter ΔV_0, on the other hand, exhibits an activated behavior with an activation energy E_a of 0.4 ± 0.1 eV. A more detailed discussion of both β and ΔV_0 and their temperature dependencies is given in Sec. IV.

The metastable-defect-creation rate for longer periods of time for various temperatures is represented by the points in Fig. 12. The defects were created by the application of 20 V stress to a nitride capacitor. The threshold-voltage shift $\Delta V(t)=V_{th}(t)-V_{th}(\infty)$ (which is less than zero) versus $\log_{10}t$ has been normalized with respect to the total shift for 20-V stressing of a nitride capacitor. In Fig. 13, the threshold shift as a function of time is presented for various biases. Note that the net long-term change in threshold, i.e., $\Delta V(t)=V_{th}(t)-V_{th}(\infty)$ is approximately equal to the applied bias. Regardless of the conditions of stressing, a similar shape is observed for the creation of interface states.

IV. DISCUSSION

In the preceding section, a significant number of independent results demonstrated that the accumulation of charges in a MIS structure with *a*-Si:H results in the creation of states in the middle to lower part of the band gap. Moreover, the density and number of these states is directly proportional to the threshold-voltage shift of the electrical characteristics. Consequently, the time and temperature dependence of the voltage shifts must be related to the defect-creation kinetics. In order to explain the experimental changes of threshold with temperature, time, bias, and doping, we must discuss possible instability mechanisms that could account for these results. The creation of carrier-induced interface states is similar to other metastable effects such as the formation of light-induced defects,[29] and the decay of excess carriers due to rapid thermal quenching.[6] In all cases, deep defects are created within *a*-Si:H in response to the presence of excess carriers—a similarity pointing to a common origin.[7] In addition, the time and temperature of the kinetics of these effects are consistent with diffusive hydrogen motion.

In this section, we show how the extended Glarum model[45,46] applied to hydrogen diffusion not only explains the decay of excess carriers[22] but also explains numerous other metastable phenomena including the metastable-defect-formation results of Sec. III. The equation for excess carrier relaxation from Refs. 45 and 22 is used to derive equations for the time and temperature dependence of the defect-creation process (Sec. IV A). In Sec. IV B, the predictions of this equation are then compared with the experimental threshold-voltage results presented in Sec. III. In Sec. IV C, a relation between the hydrogen-diffusion coefficient and the characteristic time of decay of various metastable processes is derived from the hydrogen diffusion and is verified by experiment. Finally, in Sec. IV D we discuss a possible physical model which could give rise to the hydrogen diffusion equation—an equation which appears to be successful in quantitatively explaining a wide variety of metastable phenomena in *a*-Si:H.

A. Defect kinetics equation

Since the hydrogen motion provides a mechanism for creation and annealing of metastable defects, the formation rate of defects may be governed by hydrogen diffusion. According to the extended Glarum model in Refs. 45 and 46, the rate of defect formation is proportional to the rate at which hydrogen diffuses towards defect formation sites.[22] Consequently, the change of the defect density from equilibrium, ΔN_s, is given by

$$d\,\Delta N_s/dt=-AD(t)\Delta N_s\ , \tag{2}$$

where A is a constant of proportionality, and $D(t)$ is the time-dependent hydrogen diffusion coefficient.[22,45,46] The time dependence is due to the trapping and detrapping of the hydrogen with an anomalously wide range of dwell times as it diffuses through the network. Further discussion of the implications and possible origins of Eq. (2) as well as a discussion of the applicability of the kinetic equation is presented below in Sec. IV D. In the theory of multiple trapping for trap-limited band transport in approximately exponential distribution of trap energies, the dispersive diffusion coefficient is given by[47,48]

$$D(t)=D_{00}(\omega t)^{-\alpha}\ , \tag{3}$$

where D_{00} is a microscopic diffusion, ω is a hydrogen attempt frequency, and α is the temperature-dependent

dispersion parameter which is given by $\alpha=1-\beta=1-T/T_0$. The quantity kT_0 is the characteristic energy of the exponential distribution of trapping sites, and T is the measurement temperature. Solving Eq. (2) using Eq. (3) yields the characteristic stretched exponential[22,45]

$$\Delta N_s(t)=\Delta N_s(0)\exp\left[-\left[\frac{t}{\tau}\right]^{\beta}\right], \tag{4}$$

where $\beta=T/T_0$ as before, and

$$\tau=\tau_0\exp(E_\tau/kT), \tag{5}$$

where

$$\tau_0=1/\omega \tag{6}$$

and E_τ is given by

$$E_\tau=kT_0\ln\left[\frac{\beta}{A\tau_0 D_{00}}\right]. \tag{7}$$

Equations (4)–(7) relate the time dependence of defect densities to hydrogen diffusion and have a number of importance consequences which can be tested and verified by experiment. In the following subsections, the time, temperature, doping dependence, and consequences of these equations for the various metastable phenomena observed in a-Si:H are verified.

B. Carrier-induced defect creation and threshold voltage shifts

One of the most straightforward means for testing Eqs. (4)–(7) is to compare these equations to the time dependence of the carrier-induced defect density presented in Sec. III. In order to compare the above results to data in Sec. III, a connection between the carrier density and the threshold voltage must be established. In the MIS structure, the carrier density and hence the induced defect density depends on the distance x from the insulator. Inserting the x dependence explicitly, Eq. (2) becomes

$$d\,\Delta N_s(x,t)/dt=-AD(t)\Delta N_s(x,t). \tag{8}$$

Equation (8) can be integrated over x through the thickness of the film. Defining ΔN_s according to the relation

$$\Delta N_s=\Delta N_s(t)\equiv\int_0^\infty \Delta N_s(x,t)dx, \tag{9}$$

it becomes an areal charge density. In the rest of the paper, N_s, ΔN_s, and n_{BT} (the density of band-tail carriers) refer to areal charge densities defined in a manner similar to Eq. (9). Equation (2) and all other equations in the paper hold for these areal charge densities. Accordingly, the time dependence of areal defect densities in amorphous silicon should follow a stretched exponential behavior indicated in Eq. (4). Since the threshold-voltage shift, $\Delta V_{\mathrm{th}}(t)=V_{\mathrm{th}}(t)-V_{\mathrm{th}}(\infty)=\Delta N_s/C_i$, where C_i is the insulator capacitance per unit area, the time dependence of the threshold-voltage shift is given by

$$\frac{V_{\mathrm{th}}(t)-V_{\mathrm{th}}(\infty)}{V_{\mathrm{th}}(0)-V_{\mathrm{th}}(\infty)}=\frac{\Delta N_s(t)}{\Delta N_s(0)}=\exp\left[-\left[\frac{t}{\tau}\right]^{\beta}\right]. \tag{10}$$

According to Eq. (10), the threshold-voltage shift should exhibit stretched exponential behavior with τ given by Eq. (5) and $\beta=T/T_0$. In Fig. 12, the normalized threshold shift measured on nitride transistor structures is plotted for different temperatures. The corresponding least-squares fit of the stretched exponential is also depicted (solid lines); the fit is quite reasonable for the various temperatures. The insets show the τ and β values from these fits (triangles) along with the corresponding values for the decay of excess carriers in P-doped amorphous silicon (solid circles) and from hydrogen diffusion (open circles).[24] The expected activated form of τ is observed with $E_\tau=0.95$ eV and $\tau_0=6\times10^{-10}$ sec. The activation energy is virtually identical to that observed for the decay of excess band-tail carriers.[16] This result is expected since in both cases the carrier-density reduction proceeds by the creation of defects via hydrogen diffusion. The prefactor τ_0 is slightly different due to the fact that the transistor channel in which the defects are formed is undoped, which lowers the hydrogen-diffusion rate (see discussion in Sec. IV C). The small difference in T_0 is within the experimental uncertainty since β is rather sensitive to the initial and equilibrium values of the threshold—values which are difficult to obtain.

Using data from hydrogen diffusion,[24] Eqs. (4)–(7) are in quantitative agreement with observations as well. The value of T_0 found from the hydrogen diffusion[24,22] shown in the lower inset of Fig. 12 gives a value of $T_0=650$ K compared with a value of approximately 800 K for the transistor data (lower inset of Fig. 12). From Ref. 24, the microscopic diffusion constant D_{00} was found to be 5×10^{-14} cm^2/sec for a 10^{-2} P-doped sample. Using a value of $A\simeq2.5\times10^{14}$ cm^{-2}, a value which will be shown to be physically reasonable in Sec. IV D, E_τ obtained from Eq. (7) is 0.95 eV, in good agreement with the observed activation energy. This same value of A was used in Eqs. (2)–(7) to explain the exponential "Meyer-Neldel" relation between the attempt-to-anneal factor and the annealing activation energy found for light-induced defects (see Ref. 31 for further details). This value is rather constrained by experiment.

For short times and small threshold shifts, the quantity in the square brackets in Eq. (10) is much less than unity and Eq. (10) becomes

$$\Delta V(t)=V_{\mathrm{th}}(t)-V_{\mathrm{th}}(0)=[V_{\mathrm{th}}(\infty)-V_{\mathrm{th}}(0)]\left[\frac{t}{\tau}\right]^{\beta}. \tag{11}$$

If we define

$$\Delta V_0=[V_{\mathrm{th}}(\infty)-V_{\mathrm{th}}(0)]\tau^{-\beta} \tag{12}$$

we recover the observed power-law form in Eq. (1). The power-law exponent is the same as the exponent in the stretched exponential. Hence, the β values given in Sec. III also apply to the stretched exponential. The form of charging presented in Eq. (10) provides the transition from short times to very long times and represents a generalization of Eq. (1). Because the threshold shift cannot exceed the applied bias for long times, Eq. (1) must break

down and is replaced by Eq. (10).

The values of β obtained by a fit to Eq. (10) are similar to those found in a fit to Eq. (1), indicating the consistency of the two equations. The β values from the slope of the curves in Fig. 10 are 0.56, 0.63, 0.62, and 0.67 for 320–380 K. These points are in close agreement with the β values in the lower inset of Fig. 12. The quantity T_0 has the value of 600 K in reasonable agreement with the values found for the stretched exponential fit. Since τ is thermally activated according to the relation $\tau=\tau_0\exp(E_\tau/kT)$, then Eq. (12) predicts that ΔV_0 will be given by

$$\Delta V_0=[V_{\rm th}(\infty)-V_{\rm th}(0)]\tau_0^{-\beta}\exp(-\beta E_\tau/kT) \tag{13}$$

so that ΔV_0 is activated and has an activation energy given by βE_τ or $0.5\times0.95=0.5$ eV. The observed value for activation energy of ΔV_0 from Fig. 11 is approximately 0.5 eV, in good agreement with Eq. (11). Furthermore, the value of τ_0 found from ΔV_0 is also within a factor of 10 of that found for decay of excess carriers due to rapid thermal quenching and hydrogen diffusion. The quantity ΔV_0 is related directly to microscopic parameters of hydrogen motion. Substituting Eqs. (5)–(7) into Eq. (12) yields

$$\Delta V_0=[V_{\rm th}(\infty)-V_{\rm th}(0)]A(D_{00}\omega^{-\alpha}) . \tag{14}$$

The factor in parentheses is directly proportional to the diffusion coefficient through Eq. (3) and $V_{\rm th}(\infty)-V_{\rm th}(0)\simeq V_g$, the applied bias. The threshold-shifting parameters can therefore be related directly to physical parameters.

Summarizing the results of this subsection, we observe that carrier-induced defect creation kinetics as a function of time and temperature are in quite good agreement with the hydrogen-diffusion equation [Eq. (2)]—the same equation which accounts for the decay of excess carriers in doped material and the annealing of light-induced defects.[22,30]

C. General relation between τ and hydrogen diffusion

One of the most significant consequences of Eq. (2) is that it predicts a close connection between the stretched exponential parameter τ and hydrogen diffusion. This relation applies for all metastable effects governed by Eq. (2) and, hence, should be observed for the decay of excess carriers, annealing of light-induced dangling bonds, as well as the carrier-induced creation of defects. Since it is known that the hydrogen diffusion is greatly affected by temperature and doping, this is a strong test of Eqs. (2)–(7). In order to make this connection, we must first correct the hydrogen-diffusion measurements of Ref. 24 for the effects of dispersion. We demonstrate that the usual means of determining the diffusion constant apply for dispersive transport. Once intricacies of dispersion are dealt with, the relation between the characteristic time τ and the hydrogen diffusion follows immediately.

1. Effects of dispersion on hydrogen diffusion

Because the hydrogen diffusion is dispersive, the diffusion coefficient will exhibit an anomalous dependence on the diffusion distance and time as in the case of dispersive carrier transport.[49,50] In Ref. 24, the hydrogen-diffusion coefficient was determined by fitting the hydrogen profile achieved after annealing at a known temperature and time t to an error-function profile. The diffusion coefficient D was determined by the relation

$$D=L^2/4t , \tag{15}$$

where L is the distance at which the hydrogen error-function profile falls to e^{-1} of its peak value. The question immediately arises whether this procedure is appropriate if D is time dependent. The second question is how to deal with the dependence of the diffusion constant on L.

Beginning with the first question, we show that Eq. (15) is indeed valid for dispersive diffusion. According to the one-dimensional dispersive diffusion theory in Ref. 51, the normalized diffusion profile depends on the distance x and time t only in the combination

$$s=\frac{x}{(D_{00})^{1/2}\omega^{\beta/2-1}t^{\beta/2}} . \tag{16}$$

Hence, the normalized hydrogen profile remains invariant if plotted versus s. This equation indicates how a dispersive diffusion profile progresses in time and distance, and one can follow the diffusion of a given concentration level versus time setting s equal to a constant, e.g., $s=2$. Then Eq. (16) gives the following relation between the distance L that the hydrogen has diffused and the time to diffuse this distance, t_L, namely

$$s=2=\frac{L}{(D_{00})^{1/2}\omega^{\beta/2-1}t_L^{\beta/2}} . \tag{17}$$

This relation may be rearranged to solve for the diffusion at a time t_L, yielding

$$D(t_L)=D_{00}(wt_L)^{-\alpha}=L^2/4t . \tag{18}$$

This equation is identical to Eq. (15). Thus, the time-dependent diffusion coefficient may be determined by measuring the distance L that a given concentration level has traveled in a time t_L. The most accurate way of determining the distance that a specific concentration level has traveled is to fit the profile by a functional form and then find the appropriate distance. Such a procedure utilizes concentration information over a number of points. In Ref. 24, an error-function fit was used to determine the distance to use in Eq. (18).

We next address the question of how the L dependence of the diffusion measurements in Ref. 24 should be handled. The basic approach is to correct the diffusion measurement to a constant diffusion distance. Setting Eq. (15) equal to Eq. (3), solving for the diffusion time t_L to diffuse a distance L, and then substituting into Eq. (15) yields

$$D(t_L)=\left[\frac{4D_{00}}{\omega^{\alpha}L^2}\right]^{1/\beta}\frac{L^2}{4} . \quad (19)$$

Equivalently, to correct a diffusion constant evaluated from diffusing a distance L_2 to the value for diffusing a distance L_1 at a temperature T the relation

$$D(t_{L_1})=\left[\frac{L_1}{L_2}\right]^{2-2/\beta}D(t_{L_2}) \quad (20)$$

is used. Equation (20) was used to correct the data in Ref. 24 to a constant diffusion distance of 100 nm using $T_0=650$ K for all temperatures and doping levels. This correction is actually fairly small for the data in Ref. 24 since the data correspond to a situation approaching constant-distance diffusion. The correction does result in a slight increase of the activation energy to approximately 1.5–1.4 eV from 1.25–1.35 eV, yielding better agreement with other measurements of the hydrogen-diffusion activation energy.

2. *Relation between hydrogen diffusion and* τ

We are now in a position to derive a general relation between the diffusion coefficient and the stretched exponential parameter τ. Combining Eqs. (3), (5)–(7), and (19), the relation between τ and the constant distance diffusion coefficient D is found to be

$$\tau=\left[\frac{\beta 4^{\alpha}}{AL^{2\alpha}}\right]^{1/\beta}\frac{1}{D} . \quad (21)$$

This important relation predicts that the relaxation time is inversely related to the hydrogen diffusion corrected to constant distance. Since the diffusion constant depends strongly on the temperature, doping concentration, and type of dopant, this is an important relation which can be tested using the diffusion data from Ref. 24, excess carrier relaxation data, and the carrier-induced defect kinetics.

In order to test Eq. (21) the stretched exponential parameter τ and the diffusion coefficient must be determined at the same temperature and doping level. Unfortunately, since the hydrogen-diffusion measurements require macroscopic diffusion over 10–100 nm, they must be performed at temperatures of 200–300 °C. On the other hand, the relaxation times for the various metastable phenomena are extremely short at these temperatures. As a consequence, the defect-creation and relaxation measurements must be done at temperatures of 30–150 °C. It is necessary either to measure the temperature dependence of τ at the lower temperatures and extrapolate to the higher temperatures, or to measure the temperature dependence of D and extrapolate to the lower temperatures. In the following, both procedures are performed to verify Eq. (21).

We first extrapolate the τ values to higher temperature for a variety of different samples. From the inset in Fig. 12 and Refs. 22, 29, and 30 it is found that the activation energy of τ is 0.9–1.0 eV for the decay of excess carriers, annealing of light-induced metastable defects, and creation of field-induced dangling bonds. Using this activation energy, the τ values for a variety of samples were extrapolated to 513 K where the constant distance diffusion coefficient for a variety of samples has been determined. The results are depicted in Fig. 14. For the x axis, the right-hand side of Eq. (21) was evaluated using experimentally measured diffusion constants from Ref. 24 corrected to a constant diffusion distance of $L=100$ nm while A has the same value (2.5×10^{14} cm^{-2}) as used in the preceding subsection. The quantity $\beta=1-\alpha=T/T_0$ is evaluated with $T=513$ K and $T_0=650$ K from the hydrogen diffusion as before. Included in the data are τ's from (1) the annealing of light-induced dangling bonds in undoped a-Si:H at three different temperatures determined from electron-spin resonance (solid triangles); (2) the decay of excess carriers due to rapid thermal quenching in 10^{-2} B-doped samples (open triangle) (Ref. 16); (3) the decay of excess carriers from rapid thermal quenching in 10^{-5}, 10^{-4}, 10^{-3}, and 10^{-2} P-doped samples (solid circles) (Ref. 16); (4) the creation of metastable defects induced by field-induced hole accumulation in undoped amorphous silicon (open circle) (the diffusion coefficient was taken to be the same as 10^{-2} B-doped sample); and (5) the creation of metastable defects induced by electron accumulation in 10^{-5} P-doped material for two different biases corresponding to increased carrier densities roughly equivalent to 10^{-5} and 10^{-2} P-doped material. The solid line is the value predicted by Eq. (21).

The agreement is quite good for a variety of samples with different carrier densities, doping levels, and at different temperatures. Note that if excess holes are created either by field accumulation or by doping, the hydrogen-diffusion rate is rapid and the relaxation time is short. The relaxation times also become increasingly greater as the phosphorus doping is decreased. This decreases the carrier density and causes the hydrogen diffusion to decrease. Finally, the undoped material in the annealing of light-induced dangling bonds has the longest relaxation time and the lowest H diffusion. Figure 14 demonstrates that τ is inversely related to the hydrogen-diffusion coefficient as it is altered by changing the carrier density. The results presented in Fig. 14 are compelling evidence indicating that the defect kinetics for various metastable defects are due to hydrogen motion and verifying our assumption that H motion is involved in stabilizing the metastable defects.

We can also test Eq. (21) by varying the temperature in addition to the type of sample. This comparison is accomplished by measuring the temperature dependence of the constant-distance diffusion coefficient and extrapolating to the metastable-defect measurement temperatures (Fig. 15). Because the diffusion constant is difficult to measure accurately, the extrapolation is not nearly as accurate as the τ extrapolation used for Fig. 14. We have included the results of 10^{-2} P doped for various temperatures as well. Note that the relaxation times are much longer, as expected, and the light-induced defect points which overlapped in Fig. 14 are now spread out corre-

sponding to the different temperatures of annealing. The solid line indicates in Fig. 15 again the result predicted by Eq. (21). The relaxation rate is inversely proportional to the diffusion coefficient even if the diffusion coefficient is altered by a change of temperature. This result may seem somewhat surprising since the relaxation rate has an activation energy of 0.95 eV while the constant distance diffusion has an activation energy of 1.4–1.5 eV. The difference is due the temperature dependence of the quantity in the prefactor of Eq. (21). This factor is thermally activated. The hydrogen-diffusion model thus accounts for the different activation energies between defect creation and hydrogen diffusion.

Knowing that the relaxation time varies with the hydrogen-diffusion rate, we can use the relaxation time to study the hydrogen-diffusion rate under conditions which would be impossible using the typical hydrogen-diffusion

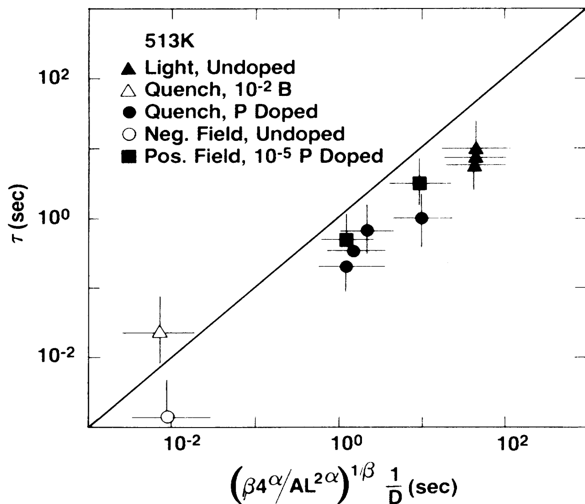

FIG. 14. A test of Eq. (21) demonstrating the inverse relation between the constant distance hydrogen-diffusion coefficient, D, and the relaxation time τ. The relaxation times were measured as a function of temperature and extrapolated to 513 K, the temperature at which the constant distance hydrogen-diffusion measurements [data from Ref. 22 and corrected using Eq. (20)] were made for a variety of metastable phenomena in samples with different doping levels. The annealing of light-induced defects in undoped silicon as determined from electron-spin-resonance measurements from different temperatures (Ref. 29) (solid triangles), the decay of rapid-thermal-quenching-induced carriers in 10^{-2} B-doped films measured by charge sweep-out (Ref. 16) (open triangle), the decay of rapid-thermal-quenching-induced carriers in P-doped samples with doping ranging from 10^{-5} to 10^{-2} (Ref. 16) (solid circles), creation defects in MIS structures under hole accumulation using C-V threshold shifts (this work) (open circle), and the creation of defects in MIS structures under electron accumulation for 3.2 and 20 V bias (from Fig. 13) (solid squares). The solid line represents the prediction of Eq. (21) with constant values obtained from other experiments.

measurements. First, we can estimate the hydrogen diffusion in undoped a-Si:H. Because the relaxation data are available over wider temperature ranges and at lower temperatures, extrapolation of the relaxation time is more reliable than the extrapolation of the diffusion constant. Using an activation energy of 0.95 eV obtained from the data (inset of Fig. 12), Eq. (21), and the relaxation time of 10^4 sec for annealing of light-induced defects at 403 K, we calculate a value of 10^{-27} cm^2/sec for the diffusion constant in undoped a-Si:H at 300 K. The relaxation time is roughly 3×10^7 sec or nearly 1 yr. One would therefore expect that light-induced defects may anneal away on the order of a 1-yr time scale.

Second, we can show that much of the variation of the hydrogen-diffusion rate observed for various doping levels is due to the presence of band-tail carriers rather than dopant atoms. By increasing the gate bias, the excess carrier density is correspondingly increased. Equation (10) predicts that if the change in threshold is normalized by the total threshold shift after long times, $V_{\rm th}(0)-V_{\rm th}(\infty)\simeq -V_g$, the result will reflect the changes in the diffusion coefficient due to the increased carrier concentration. This result is observed in Fig. 16 where the normalized threshold-voltage shift of a transistor with silicon-nitride gate dielectric is plotted for various gate biases along with the corresponding stretched exponential fits. The decrease of the relaxation time by approximately a factor of 8 indicates that the hydrogen-diffusion rate increases as a result of the increased carrier density. This result suggests that the increase in the hydrogen-diffusion rate for increased P doping observed in Ref. 24 is due to the increased carrier density rather than the dangling-bond density.

Further corroboration that the carrier density affects the hydrogen-diffusion rate comes from the observation of negative bias stressing (hole accumulation) on a capacitor with undoped a-Si:H and a thermal-oxide dielectric. The threshold shifts quite rapidly ($\tau=1000$ sec) even at room temperature. This result is expected since hydrogen diffuses 3–4 orders of magnitude more rapidly in 10^{-2} B-doped material. This point is represented by the open circle in Figs. 14 and 15. These results demonstrate that it is the carrier density which causes the increased hydrogen-diffusion rate in B- and P-doped material rather than either the doping-induced increase in dangling-bond density or doping-induced changes in film microstructure.

This close relation between the creation of defects due to rapid thermal quenching, hydrogen diffusion, and the threshold-voltage shift has important implications for the physical mechanism causing threshold shifts. The data indicate that the a-Si:H near the nitride creates defects in response to the increased density of carriers present at the interface. The process stops only after the Fermi level is restored to midgap. Assuming that the defects are generated within the accumulation region, which is roughly 10 nm wide, the generated defect density is approximately 5×10^{17} states cm^{-3}. The fact that the kinetics of the charging process match the defect creation of other metastable processes associated with bulk a-Si:H suggests that the defects are being formed within the a-

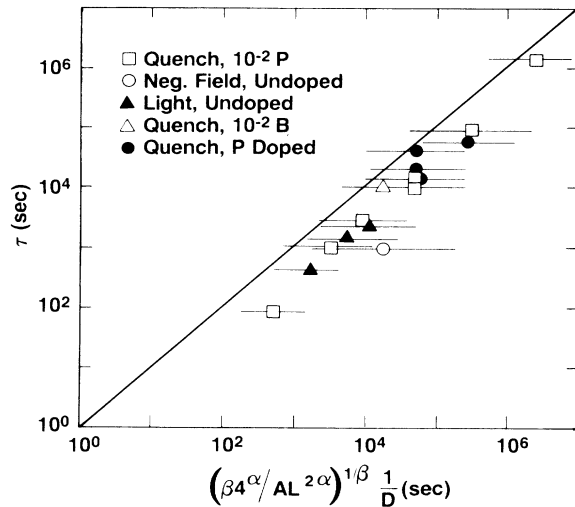

FIG. 15. A test of Eq. (21) demonstrating the inverse relation between the constant distance hydrogen-diffusing coefficient D and the relaxation time τ. The constant distance hydrogen-diffusion coefficient [using Eq. (20)] was extrapolated to the various temperatures at which τ was measured for a variety of different metastable effects. The symbols are the same as in Fig. 14, except that the open squares represent results from the decay of excess carriers from 10^{-2} P samples at various temperatures determined by voltage sweep-out (Ref. 16).

Si:H rather than the nitride. One might expect the defect creation rates to differ significantly between the two materials. Thus, the kinetics are consistent with the hypothesis of interface-state generation.

In conclusion, Eq. (21) explains quantitatively the relation between the hydrogen-diffusion coefficient and the time scale of a variety of metastable changes observed under various temperatures, carrier densities, and doping conditions. This result suggests that the hydrogen-diffusion equation, Eq. (2), and subsequent equations describe the metastable kinetics in a-Si:H.

D. Origin of stretched exponential equation

Having established that Eq. (2) involving hydrogen diffusion explains the defect kinetics for a variety of different metastable phenomena, we discuss possible origins and physical implications of this diffusion equation as well as showing that the value found for A used in previous calculations can be derived in terms of physical quantities such as the hydrogen density and trapping parameters.

1. Hydrogen motion—Cause or effect?

There are two possible explanations for the observation of such a close correspondence between the rate at which metastable changes occur and the hydrogen diffusion presented above. Either hydrogen motion is involved directly, i.e., causes the metastable changes, or alternatively, some other defect or lattice reconstruction such as a fivefold silicon atom causes both hydrogen motion and the metastable changes. The latter hypothesis is unlikely for the following reasons. First, it would be remarkable if the creation or annealing of metastable defects by this third defect was not altered by temperature, doping, carrier density, or carrier type. For example, one would expect that moving the Fermi level might increase or decrease the energy required to change the defect energy or that the defect-creation process has some additional activation energy so that the temperature dependence of the defect kinetics would be different from that observed for hydrogen diffusion. Holes might decrease the rate of metastable-defect formation but increase the rate of hydrogen diffusion. Second, such a defect involving only silicon would either have to be extremely mobile or be present in large numbers. If it is present in large numbers, then it should be readily observable by transport or optical techniques. If the defect is extremely mobile, then this defect and the corresponding changes should occur rapidly at lower temperatures in unhydrogenated silicon, and there should not be significant differences in the rate changes between a-Si and a-Si:H. These consequences are in disagreement with experiment.[25,26] Third, the time dependence of the defect formation should not necessarily depend on the hydrogen diffusion rather than some other aspect such as the hydrogen density. Finally, one can always add additional complexities to a theory to explain results. It is clear, however, that in the present case, hydrogen motion is sufficient to explain the kinetics of light-induced defects in a-Si:H without the need for the added complexity of additional mechanisms. Because the hydrogen diffusion in a-Si:H is quite similar to that observed in c-Si without significant numbers of defects, the hypothesis of additional defects is also unnecessary in that case.[27] While completely definitive experimental evidence is not yet available, the present experiments indicate that the involvement of hydrogen in metastable-defect kinetics is the most likely possibility.

2. Possible origin of the hydrogen-motion equation

The experimental data indicate that there is a complex interaction between carriers, the silicon, and hydrogen motion, with many of the details still unknown. It is difficult to determine the exact physical picture responsible for metastable-defect changes. The model presented in the following is one of the simpler possible ways which will given rise to Eq. (2) and the observed results. Undoubtedly, the real situation will turn out to be more complex.

Begin by considering ways in which carriers can induce bond rearrangements that produce metastable defects deep in the gap, approximately at the energy location of the silicon dangling bond. The carriers can alter either silicon-silicon bonds or silicon-hydrogen bonds in undoped material. In doped material, a third possibility is that the metastable changes could be due to changes in dopant bonding as well.[21,16] Any metastable change in bonding configuration caused by carriers must be stabilized against subsequent relaxation. Because of the vari-

ous arguments presented previously, we shall assume that the bonding changes are stabilized by hydrogen motion. Either (1) the excess carriers break weak silicon-silicon bonds[24] which are subsequently stabilized by the diffusion of hydrogen (Fig. 17), (2) the excess carriers can increase the release rate of hydrogen from the silicon (break Si—H bonds), or (3) a combination of both may occur. In either case the result is the creation of new dangling bonds, by a process involving both excess carriers and the diffusion of hydrogen. As discussed below in more detail, the excess carriers can also alter the rate at which hydrogen diffuses through the network.

Beginning with case (1) depicted in Fig. 17, the excess carriers may disrupt the weak silicon-silicon bond by occupying the antibonding state.[24] However, the defect is stable only if a hydrogen atom diffuses to the site creating metastable dangling bonds. This motion creates dangling bonds which trap the band-tail carriers and reduce the carrier density. Eventually, the carrier density decreases and further dangling-bond creation decreases. In the case of doped material, the hydrogen can also alter the doping efficiency by changing the coordination of the dopant atoms. The following discussion demonstrates how the mathematical description of the defect-formation process along with dispersive hydrogen diffusion gives rise to the observed field-induced defect formation as a function of time, temperature, doping, and field.

The rate of defect formation, dN_s/dt, is given by

$$dN_s(t)/dt = R(t) N_p P_{\rm occ} , \tag{22}$$

where N_p is the density of sites at which the defects may be created, $R(t)$ is the current of hydrogen flow towards these sites, and $P_{\rm occ}$ is the probability of the site being occupied by a band-tail carrier. Now $R(t)$ is given by

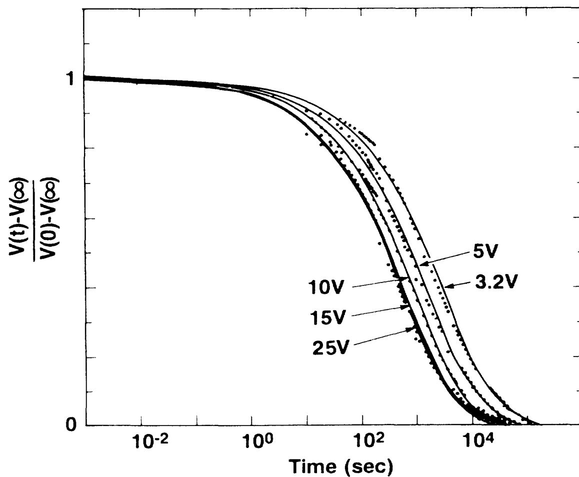

FIG. 16. The normalized threshold-voltage shift for nitride transistors at 403 K as a function of bias. The large biases shift faster due to the change in hydrogen-diffusion rate and the number of band-tail carriers.

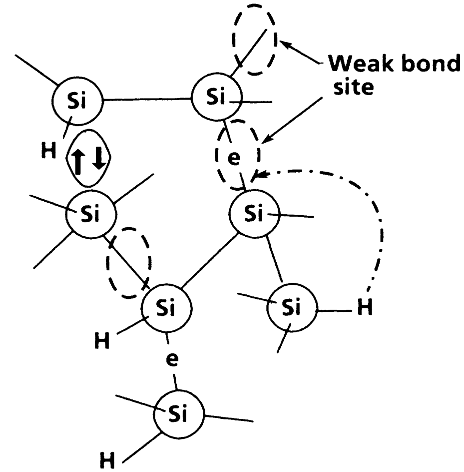

FIG. 17. The hydrogen-diffusion model for creation and annealing of defects. The hydrogen atom diffuses to a special site (weak Si—Si bond) where it creates a dangling bond and leaves behind another dangling bond. The breaking of the weak Si—Si bond, creation of the dangling bond, and trapping of the hydrogen at this site occurs only if a band-tail carrier is present. The process reduces the band-tail carrier density and brings the Fermi level back to midgap.

$$R(t) = 4\pi r_0 N_{\rm H} D(t) , \tag{23}$$

where r_0 is the capture radius of the site, $N_{\rm H}$ is the density of hydrogen, and $D(t)$ is the trap-dominated dispersive diffusion coefficient of hydrogen. The factor $P_{\rm occ}$ is given by

$$P_{\rm occ} = n_{\rm BT}/N_T , \tag{24}$$

where $n_{\rm BT}$ is the density of band-tail carriers and N_T is the number of tail states. In general, the density of defect-creation sites is on the order of the number of tail states or $N_T \simeq N_p$. Combining Eqs. (22)–(24) yields

$$dN_s(t)/dt = A D(t) n_{\rm BT}(t) , \tag{25}$$

where A is given by

$$A = 4\pi r_0 N_{\rm H} (N_p/N_T) . \tag{26}$$

This relation gives the parameter A in terms physical parameters of the hydrogen density and the capture radius. Since the hydrogen is captured by a silicon bond, the capture radius is approximately 0.2–0.3 nm, and $N_{\rm H}$ is approximately 10^{21} cm^{-3}. With these values, we find that $A \simeq 2.5 \times 10^{14}$ cm^{-2}, the value that was used throughout the previous sections to give good agreement with experiment.

Considering case (2), when the excess carriers create dangling bonds by increasing the rate of hydrogen release, one would expect that the release rate (dangling-bond creation rate) would be proportional to the density of band-tail carriers. Furthermore, since the rate-limiting step in H diffusion is likely to be the release of hydrogen from its silicon bond, the release rate should be propor-

tional to the hydrogen diffusion as well. The defect-creation rate would thus have a similar form to Eq. (25), although the constant would be related to the density of hydrogen and the release radius. Thus, regardless of whether the excess carriers increase the rate of stabilized destruction of weak Si—Si bonds or the rate of hydrogen release, one would expect the defect-creation rate to be given by Eq. (25). In both cases (1) and (2), it should be pointed out that the excess carriers affect the rate of hydrogen diffusion through the lattice once it is released—an effect known to occur in crystalline silicon. Therefore, the hydrogen diffusion $D(t)$ depends on the carrier density as well.

We now derive the kinetics which result from Eq. (25). The density of band-tail carriers is related to the number of defects by the relation[6]

$$n_{\rm BT}(t)=N_{\rm ex}(t)-N_s(t) \; , \tag{27}$$

where $N_{\rm ex}(t)$ is the density of band-tail carriers in the absence of defects. In the case of doped material, $N_{\rm ex}(t)$ is equal to the density of fourfold dopant atoms contributing carriers. Because the number of fourfold dopant atoms changes as a function of time, $N_{\rm ex}$ is in general time dependent. In the case of field-induced doping at the amorphous-silicon–dielectric interface, $N_{\rm ex}$ is independent of time and is equal to the density of carriers induced by the gate field in the semiconductor before any states near the interface are created. This density is roughly proportional to the gate voltage V_g for voltages above threshold. At equilibrium when $t=\infty$, $dN_s/dt=0$ and $n_{\rm BT}$ becomes very much smaller than N_s. Hence, we can make the approximation that $n_{\rm BT}(\infty)\simeq 0$. The equilibrium defect density $N_s(t=\infty)=N_{\rm ex}(t=\infty)$. Letting

$$\Delta N_s(t)=N_s(t)-N_s(\infty) \; , \tag{28}$$

Eq. (28) becomes

$$d\Delta N_s/dt=-AD(t)\Delta N_s \; , \tag{29}$$

which is identical to Eq. (2). Thus, we see that Eq. (2) describes the dispersive diffusion of hydrogen towards specific sites where defects are created or destroyed, which is similar to the Glarum model which is based on diffusion of defects towards dipoles. The rate of metastable changes can be altered by lowering the hydrogen-diffusion rate and eliminating the hydrogen.

Recently, it has been pointed out that there is some theoretical question concerning the applicability of the kinetic-type equation, such as Eq. (29) for the case of dispersive transport.[52] Some previous work found that the stretched exponential of Eq. (4) describes the decay of defects due to dispersive recombination of defects.[45] Other theoretical work has found that the long time decays for dispersive diffusion to a defect where the first arriving hydrogen neutralizes or creates the defects results in a power-law decay at long times rather than the exponential decay predicted by the kinetic equations.[53–55] In the latter case, the time dependence of the defect-creation process should exhibit the form

$$\Delta N_s(t)=\frac{\Delta N_s(0)}{1+(t/\tau)^{\beta}} \tag{30}$$

rather than the stretched exponential form given by Eq. (4). The form given by Eq. (30) is very similar to Eq. (4) except at very long time scales. It fits the time-dependence data at least as well as Eq. (4) and therefore is also consistent with the data. The equations for τ, β, all of the results in Eqs. (2)–(7), the short-time limiting forms given by Eqs. (11)–(14), and Eq. (21), hold for Eq. (30) as well. It can be shown theoretically and by fitting the data to Eq. (30) that the τ's in the two cases differ by a small constant factor independent of doping, temperature, etc., and the resulting value of T_0 (which determines β) has similar values. Therefore, the results presented above apply regardless of which form correctly describes the experimental situation. The comparison of the data with Eq. (30) will be further discussed in a future paper.

V. CONCLUSIONS

The model of hydrogen-mediated defect creation of interface states quantitatively predicts the defect-creation rate as a function of time, temperature, doping, and bias from the hydrogen-diffusion data. Specifically, the following predictions have been verified. (1) The stretched exponential time dependence of threshold shift has been observed for the decay of excess carriers, the creation of states near the interface, and the annealing of light-induced defects. (2) The activation energy and prefactor of the characteristic time τ are predicted to be less than the observed activation energy of hydrogen due to dispersive transport of hydrogen. (3) A Meyer-Neldel relation between the attempt-to-anneal frequency and the activation energy was quantitatively observed. (4) The linear dependence of the stretched exponential factor β on measuring temperature was observed and is consistent with dispersive hydrogen diffusion. The characteristic temperatures (650 K) are in rough agreement as well. (5) The inverse relationship between the characteristic stretched exponential time τ, and the constant distance diffusion coefficient for different temperatures, times, doping levels, and biases has been confirmed. (6) The effect of bias on defect-creation rate was predicted. On the basis of the above predictions, hydrogen diffusion must be considered as a likely candidate for many if not all the metastable-defect phenomena observed in hydrogenated amorphous silicon and provides strong support for the results in Ref. 7. This model naturally explains the similarity between the kinetics of the various metastable processes arising from the common mechanism of hydrogen motion and provides a physical explanation of the various parameters in terms of hydrogen motion.

ACKNOWLEDGMENTS

We gratefully acknowledge many stimulating discussions with R. A. Street, J. Kakalios, K. Winer, Z. Smith, and H. Scher. This work was supported in part by the Solar Energy Research Institute (Golden, CO).

*Permanent address: Department of Materials Engineering, University College of Swansea, Singleton Park, Swansea SA2 8PP, United Kingdom.

[1]D. L. Staebler and C. R. Wronski, Appl. Phys. Lett. **31**, 292 (1977).

[2]D. L. Staebler and C. R. Wronski, J. Appl. Phys. **51**, 3262 (1980).

[3]H. Pfleiderer, W. Kusian, and W. Krühler, Solid State Commun. **49**, 493 (1984).

[4]D. V. Lang, J. D. Cohen, and J. P. Harbison, Phys. Rev. Lett. **48**, 421 (1982).

[5]M. Stutzmann, D. K. Biegelsen, and R. A. Street, Phys. Rev. B **35**, 5666 (1987), and references therein.

[6]R. A. Street, J. Kakalios, C. C. Tsai, and T. M. Hayes, Phys. Rev. B **35**, 1316 (1987).

[7]W. B. Jackson in *Proceedings of KOSEF/NSF Joint Seminar: The Physics of Semiconductor Materials and Applications,* edited by C. Lee and W. Paul (Korea Science and Engineering, Seoul, 1987), p. 47.

[8]K. Morigaki, I. Hirabayashi, M. Nakayama, S. Nitta, and K. Shimakawa, Solid State Commun. **33**, 851 (1980).

[9]J. I. Pankove and J. E. Berkeyheiser, Appl. Phys. Lett. **37**, 705 (1980).

[10]H. Dersch, J. Stuke, and J. Beichler, Appl. Phys. Lett. **38**, 456 (1980).

[11]D. E. Carlson, Appl. Phys. A **41**, 305 (1986).

[12]M. Stutzman, W. B. Jackson, and C. C. Tsai, Phys. Rev. B **32**, 23 (1985).

[13]G. Müller, S. Kalbitzer, and H. Mannsperger, Appl. Phys. A **39**, 243 (1986).

[14]R. A. Street, J. Kakalios, and T. M. Hayes, Phys. Rev. B **34**, 3030 (1986).

[15]R. A. Street, Phys. Rev. Lett. **49**, 1187 (1982).

[16]R. A. Street, M. Hack, and W. B. Jackson, Phys. Rev. B **37**, 4209 (1988).

[17]Z. E. Smith, S. Aljishi, D. Slobodin, V. Chu, S. Wagner, P. M. Lenahan, R. R. Arya, and M. S. Bennett, Phys. Rev. Lett. **57**, 2450 (1986).

[18]Z. E. Smith and S. Wagner, Phys. Rev. B **32**, 5510 (1985).

[19]M. Stutzmann, Philos. Mag. B **56**, 63 (1987).

[20]D. Redfield, Appl. Phys. Lett. **49**, 1517 (1986).

[21]W. B. Jackson, in *Stability of Amorphous Silicon Alloy Materials and Devices (Palo Alto, CA, 1987),* Proceedings of an International Conference on Stability of Amorphous Silicon Alloy Materials and Devices, AIP Conf. Proc. No. 157, edited by B. L. Stafford and E. Sabisky (AIP, New York, 1987), p. 17.

[22]J. Kakalios, R. A. Street, and W. B. Jacskon, Phys. Rev. Lett. **59**, 1037 (1987).

[23]D. E. Carlson and C. W. Magee, Appl. Phys. Lett. **33**, 81 (1978).

[24]R. A. Street, C. C. Tsai, J. Kakalios, and W. B. Jackson, Philos. Mag. **56**, 305 (1987).

[25]P. A. Thomas, M. H. Brodsky, D. Kaplan, and D. Lepine, Phys. Rev. B **18**, 3059 (1978).

[26]W. B. Jackson and C. C. Tsai (unpublished).

[27]N. M. Johnson, Phys. Rev. B **31**, 5525 (1985).

[28]J. Jang, T. M. Kim, J. K. Hyun, J. H. Yoon, and C. Lee, J. Non-Cryst. Solids **59&60**, 429, (1983); W. den Boer and S. Guha, J. Appl. Phys. **57**, 5539 (1985).

[29]W. B. Jackson and J. Kakalios, Phys. Rev. B **37**, 1020 (1988).

[30]W. B. Jackson and D. M. Moyer, Phys. Rev. B **36**, 6217 (1987).

[31]W. B. Jackson Phys. Rev. B **38**, 3595 (1988).

[32]M. J. Powell, J. Appl. Phys. Lett. **43**, 597 (1983).

[33]R. A. Street and C. C. Tsai, Mater. Res. Soc. Proc. **70**, 367 (1986).

[34]R. A. Street and C. C. Tsai, Appl. Phys. Lett. **48**, 1673 (1986).

[35]A. R. Hepburn, J. M. Marshall, C. Main, M. J. Powell, and C. van Berkel, Phys. Rev. Lett. **56**, 2215 (1986); A. R. Hepburn, C. Main, J. M. Marshall, J. C. van Berkel, and M. J. Powell, J. Non-Cryst. Solids **97&98**, 903 (1987).

[36]R. E. I. Schropp and J. F. Verwey, Appl. Phys. Lett. **50**, 185 (1987); R. E. I. Schropp, A. J. Boonstra, and T. M. Klapwijk, J. Non-Cryst. Solids **97&98**, 1339 (1987).

[37]W. B. Jackson, M. D. Moyer, C. C. Tsai, and J. Marshall, J. Non-Cryst. Solids **97&98**, 891 (1987).

[38]D. J. DiMaria, in *The Physics of* SiO_2 *and its Interfaces,* edited by S. T. Pantelides (Pergamon, New York, 1978), p. 160, and references therein.

[39]H. Tuan, Mater. Res. Soc. Symp. Proc. **33**, 247 (1984).

[40]S. M. Sze, *Physics of Semiconductor Devices,* 1st ed. (Wiley, New York, 1969), Chap. 9.

[41]M. J. Powell, C. van Berkel, I. D. French, and D. H. Nicholls, Appl. Phys. Lett. **51**, 1242 (1987).

[42]A. Iqbal, W. B. Jackson, C. C. Tsai, and J. W. Allen, J. Appl. Phys. **61**, 2947 (1987).

[43]D. Smith (private communication).

[44]Y. Kaneko, A. Sasano, T. Tsukada, R. Oritsuki, and K. Suzuki, *Extended Abstracts of the 18th International Conference on Solid State Devices and Materials, Tokyo, 1986,* edited by T. Sugamo (Komiyama, Tokyo, 1986), p. 699.

[45]M. F. Shlesinger and E. W. Montroll, Proc. Nat. Acad. Sci. USA **81**, 1280 (1984).

[46]S. H. Glarum, J. Chem. Phys. **33**, 1371 (1960).

[47]M. Grünewald, B. Pohlmann, B. Movaghar, and D. Würtz, Philos. Mag. B **49**, 341 (1984).

[48]H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).

[49]J. Orenstein and M. Kastner, Phys. Rev. Lett. **46**, 1421 (1981); T. Tiedje and A. Rose, Solid State Commun. **37**, 49 (1981).

[50]T. Tiedje, in *Semiconductors and Semimetals,* edited by J. I. Pankove (Academic, New York, 1984), Vol. 21C, p. 207, and references therein.

[51]P. N. Butcher and J. D. Clark, Philos. Mag. B **42**, 191 (1980).

[52]H. Scher (private communication).

[53]B. Movaghar, J. Phys. C **13**, 4915 (1980).

[54]J. Orenstein, M. A. Kastner, and V. Vaninov, Philos. Mag. B **46**, 23 (1982).

[55]H. Scher, J. Phys. (Paris) Colloq. **42**, C4-547 (1981).