Measurement of the density of gap states in hydrogenated amorphous silicon by space charge spectroscopy

David V. Lang, J. David Cohen, and James P. Harbison
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 30 June 1981)

We report a comprehensive study of the dynamic response of junction space-charge layers in undoped and PH_3-doped *a*-Si:H films grown by the rf-glow-discharge technique. By using the numerical analysis methods discussed in the adjoining theory paper, we are able consistently to interpret a variety of transient response and ac admittance measurements in terms of a bulk density of gap states $g(E)$ which is characteristic of each sample. While the general shape of $g(E)$ seems to be a characteristic property of *a*-Si:H, the overall concentration of gap states depends on growth conditions and doping. The density of states at approximately midgap is observed to vary between values as low as about 2×10^{15} $cm^{-3}eV^{-1}$ in undoped films and as high as 1×10^{18} $cm^{-3}eV^{-1}$ in some PH_3-doped films. The general shape of our $g(E)$ is dominated by a deep minimum ($<10^{16}$ $cm^{-3}eV^{-1}$) between 0.3 and 0.6 eV from the conduction band and a broad shoulder of states extending from the valence band up to midgap. The significant difference between this type of bulk $g(E)$ and previous models for the density of states in *a*-Si:H may be explained by the effects of states at or near the surface of the films which strongly influence the previous types of measurements. We discuss recent transport and optical measurements and show that they provide strong support for our density of states as opposed to previous models for $g(E)$.

I. INTRODUCTION

Recently there has been considerable interest in measuring the density and energy distribution of gap states in hydrogenated amorphous silicon (*a*-Si:H).[1] Such states are obviously of central importance to the understanding of the transport and optical properties of *a*-Si:H. Indeed, many of these properties are controlled by states in the gap rather than by the much larger number of states in the valence and conduction bands. One of the interesting features of amorphous materials in general is that there is not a clear-cut distinction between "band" and "gap" states except insofar as the density of states at a particular energy leads to delocalization or localization of the electronic wave functions. The demarcation line between these two regions is called the mobility edge.[1] Since hopping conduction and optical absorption can both occur to varying degrees in the localized states, the "band gap" is somewhat ill-defined and may depend on the type of experimental technique used to measure it.

In spite of these difficulties, however, there has emerged over the past few years a widely accepted model[1,2] for the density of gap states in *a*-Si:H. This well-known density distribution, which we shall call $g_{SL}(E)$, was originally based on field-effect measurements of Spear, LeComber, and coworkers but has subsequently been shown to be resonably consistent with a variety of other measurements.[2] The overall consistency of the $g_{SL}(E)$ model is based on certain reasonable assumptions which, however, have not been independently verified.

It therefore seemed of utmost interest to determine $g(E)$ utilizing new techniques which do not rely on the same assumptions as these previous techniques. To this end we have undertaken an extensive series of measurements on *a*-Si:H Schottky barriers in an effort to extract the density of gap states $g(E)$ from the dynamic electric response of these structures. Preliminary accounts of this work have been published elsewhere.[3–5] Our $g(E)$ results differ substantially from $g_{SL}(E)$. While it might be argued that such differences are sample dependent, the type of optical and electric behavior upon which $g_{SL}(E)$ is based is quite generally observed by many different workers. Our samples in particular have the same basic transport and ESR properties that are typical of high-quality *a*-Si:H.

To set our experimental methods in perspective,

let us first briefly review the strengths and weaknesses of the various techniques which have been used in the past to obtain information about $g(E)$. The *field-effect* method gave historically the first evidence that the density of gap states in *a*-Si:H was much lower than in *a*-Si.[6] However, this method is subject to many potential problems.[4,7,8] The major difficulty is that the transconductance of a field-effect transistor is highly sensitivity to states at or near the oxide-semiconductor interface and that these measurements cannot inherently distinguish the relative contributions from these interface or "surface" states as opposed to true bulk states. Such difficulties have been recognized by the proponents of this method and it has been argued that the general consistency of $g_{SL}(E)$ with a variety of other experimental data justifies the neglect of interface states in the field-effect analysis.[2] However, there is now considerable evidence that interface states are an important factor in *a*-Si:H.[9] In some cases interface effects dominate the apparent bulk properties of thin insulating *a*-Si:H films.[10–15] Interface states have in fact been observed to influence field-effect measurements in a study which explicitly considered the dynamics of interfacial charge screening.[10] One would also expect lateral inhomogeneities in the interfacial charge distribution to give a spread of flat-band voltages and hence an apparent but spurious band tail in $g(E)$.[16] Such lateral inhomogeneities may be inferred from the work of Solomon.[17]

A second method closely related to the field effect is to obtain $g(E)$ from the *steady-state capacitance-voltage* (*C-V*) characteristics of an *a*-Si:H metal-oxide–semiconductor (MOS) structure,[19] Schottky barrier,[19–22] or *pn* junction.[2] In the limit where the frequency is low enough to allow midgap states to follow the oscillating capacitance measurement voltage, it is straightforward to obtain $g(E)$ from the *C-V* data *provided* interface states can be neglected. Unfortunately, this is precisely the set of conditions (low frequency and/or high temperature) where the anomalous effects due to interface states have been reported.[20,23]

Recently, a number of studies have inferred $g(E)$ at midgap from the *frequency and/or temperature dependence of the capacitance* of an *a*-Si:H Schottky barrier.[20,21,23–25] These sorts of measurements have an advantage in that they may be made in a regime of temperature and frequency which is less sensitive to interface states than are *C-V* and field-effect measurements.[20–23] Several important results have been obtained by such measurements. Viktorovitch[23] has shown that the bulk $g(E)$ at midgap depends strongly on sample-preparation conditions and that the interface state density may be very large in MOS structures. Such techniques have an inherent shortcoming, however, in that they are able to measure $g(E)$ only over a relatively small portion of the gap near the Fermi level. Furthermore, the identification of bulk and surface contributions to the signal is not always straightforward.

A closely related method is the study by Balberg[26–28] of the *admittance versus bias voltage and frequency of metal-insulator-semiconductor (MIS) tunnel diodes.* These results were interpreted as being due to interface states with a peak at $E_c - 0.5$ eV rather similar to $g_{SL}(E)$. It was claimed from dc *I-V* measurements that the same distribution of states extended into the bulk.[27] The major problem with this type of tunneling measurement is the lack of a detailed theoretical understanding of the *barrier,* which in turn precludes obtaining quantitative results for $g(E)$ or in separating surface from bulk effects.[29]

All of the above techniques (field-effect, *C-V*, *C-ω*, *C-T*, and tunneling) are *steady state* in the sense that the independent variable (V, ω, or T) is changed slowly so that the corresponding change in the dependent variable (I, C, or G) is complete for each measurement step. It is also possible to explicity study the *transient* response of the dependent variable following a step function change in the independent variable, usually V. Such transient studies are widely used to study deep levels in the gap of crystalline semiconductors.[30,31] Transient junction measurements have only recently been applied to *a*-Si:H[3–5,32–34] but show great promise as a means of unambiguously measuring $g(E)$. Such measurements probe the same physical parameters as does the steady-state frequency response. In principle the two are related by Laplace transform.[35] Thus, the choice between the steady-state and transient-measurement mode is ultimately related to the ease of data analysis or insensitivity to extraneous effects such as interface states.

Other more traditional bulk measurements such as transport, spin resonance, and optical effects have also been used to study the gap states.[1] The $g(E)$ results obtained from these methods usually ignore the effects of band bending near the interfaces and the effects of interface states. In spite of the fact that such "bulk" measurements may be seriously in error in *thin* films ($\lesssim$few micrometers) due to the interface effects,[9–15] such results are

often viewed as being more respresentative of the bulk than are those obtained from the various depletion layer methods mentioned above. Since these methods have been reviewed elsewhere,[1,2] we will only discuss a few specific techniques as they pertain to our experimental results later in this paper.

The methods which we have used to obtain $g(E)$ include the steady-state junction capacitance and conductance methods mentioned above as well as various transient measurements. The major innovation has been to apply the technique of deep-level transient spectroscopy[30,31] (DLTS) to the case of amorphous semiconductors. Our first quantitative results[4,5] for $g(E)$ were based on an extension of the standard DLTS analysis of discrete deep levels in crystalline semiconductors. One expects this extension to be valid for the limit of small $g(E)$. However, our results showed that the large concentration of deep gap states did not justify the small signal approximation. We therefore undertook a detailed theoretical study of the DLTS method as it applied to an arbitrary density of states.[35] In so doing, we first found it necessary to solve the amorphous diode problem in a more rigorous way than in previous treatments.[2,19–25] In particular, we used a defintion of the diode capacitance which related directly to the physics of thermal capture and emission at localized states rather than an *ad hoc* truncation of the depletion region or a lumped-element RC equivalent circuit as in previous calculations.[2,19–25] The results allow us to calculate the full dynamic response of the junction [C and G vs V,ω,T; thermally stimulated capacitance (TSCAP), thermally stimulated current (TSC), DLTS] for *any* arbitrary $g(E)$. This more exact analysis of the DLTS data gives slightly larger $g(E)$ concentrations (2 to 4 times greater) than our original more approximate analysis.[4,5]

From the unified theoretical treatment in Ref. 35, it is quite clear that *all* steady-state and transient admittance measurements relate to the same physical phenomenon, namely, the thermal capture and emission of carriers at gap states within the depletion region. One might ask, therefore, whether the DLTS technique provides sufficient advantages to justify its added complexity relative to the more straightforward steady-state measurements. We believe that capacitance transient DLTS is clearly superior to the other methods of analyzing deep levels. This should become clear as one reads this and the following paper. To summarize its major advantages: (a) the raw DLTS data in doped samples is nearly proportional to $g(E)$ over $\frac{2}{3}$ of the gap and the calculated DLTS line shape is very sensitive to small changes in the shape of $g(E)$, (b) capacitance DLTS is the least sensitive to interface states of all junction admittance measurements, (c) spatial variations in $g(E)$ can be directly measured in order to verify bulk behavior, (d) minority-carrier trap distributions can be measured, and (e) recombination and trapping rates can be measured directly.

We do not intend to imply, however, that DLTS should *replace* the simpler junctions measurements, but rather should *supplement* them. In fact, one only has a complete and consistent picture of $g(E)$ when *all* aspects of the junction dynamics can be consistently explained by a particular model. Each type of measurement provides a different viewpoint on the basic physics of the junction dynamics. In this paper we will organize our presentation by first discussing the results of the simplest measurement techniques and gradually work up to a discussion of the DLTS data. However, in spite of the fact that DLTS is perhaps more complex to explain and to implement than are the other techniques, it is actually the *easiest* to analyze in terms of $g(E)$. Thus our theoretical discussion is made in *reverse* order. Namely, we shall first obtain $g(E)$ from the DLTS spectra and then discuss the relationship of this $g(E)$ to the steady-state C-V and C-T results.

The paper is divided into eight sections, of which this introduction is Sec. I. In Sec. II we outline our standard rf glow-discharge growth technique and sample preparation procedures. Typical data are presented in Secs. III–V in connection with the discussions of the relevant experimental techniques: Sec. III covers steady-state methods, Sec. IV discusses transient techniques, and Sec. V discusses the DLTS data. The most important parts of the paper are Secs. V and VI. There we discuss the procedure for interpreting the DLTS data in terms of the density-of-states distribution function and show $g(E)$ for several representative types of samples. This is followed by a discussion in Sec. VII of the relationship of our $g(E)$ results to other results in the literature. Finally in Sec. VIII we summarize the main results of the paper and the conclusions which may be drawn from them.

II. GROWTH AND SAMPLE PREPARATION

All a-Si:H samples were prepared in a conventional capacitively coupled rf glow-discharge

growth system similar to that previously described by Knights,[36] with some minor exceptions. The walls of the growth chamber are insulating (fused silica), resulting in the grounded and ungrounded electrodes being roughly comparable in size, with the substrates always mounted on the grounded electrode ("anodic"[36] or "unbiased"[37] films). The entire dimensions of the growth chamber are relatively small, with a plate diameter of 2 cm and a plate spacing of 1.5 cm, all enclosed in a spherical chamber of 5-cm diameter. The system is vacuum pumped with a standard mechanical roughing pump used during the high gas load growth process itself, but for some of the later samples considered in the study a diffusion pump equipped with a liquid-nitrogen cold trap was added to give a starting base pressure in the mid-10^{-6}-Torr range before introducing the growth gases. As an additional precaution to keep unwanted background impurity levels at a minimum, after venting the system in an argon atmosphere for substrate introduction, the chamber was alternately vacuum pumped and argon backfilled at least three times before each growth run.

Gas flows during growth consisted of 50 cm^3/min of mixtures of silane (semiconductor purity,[38] 100 Ω cm minimum resistivity, *n* type) diluted in varying amounts of argon [ultrahigh purity (UHP), 99.999%]. For *n*-type doping, electronic-grade phosphine, PH_3, was introduced into the gas stream either in the form of a calibrated dilution mixture of PH_3 in UHP argon, or directly, through a precalibrated variable-leak valve. Table I summarizes the gas compositions together with the relevant accompanying growth conditions for each of the films discussed in detail in later sections of this paper.

TABLE I. Growth parameters for samples shown in figures.

Sample number	Silane in Ar (mol %)	Phosphine in silane (vppm)	rf power density (mW/cm^2)	Film thickness (μm)
1 (106)	2.7	undoped[a]	36	2.8
2 (123)	100	undoped	42	1.3
3 (138)	10	300	23	0.7
4 (139)	45	60	30	1.7
5 (152)	10	300	12	1.1
6 (228)	100	undoped	30	2.3

[a]See text.

The range of growth parameters for the entire series of samples included in this study were as follows: The substrate temperature and deposition pressure were held fixed at 250 °C and 0.36 Torr, respectively. The rf excitation at 13.56 MHz was at a power level ranging from 2 to 8 W which, when taking into account the actual fraction of rf power being coupled into the plasma, is 12 to 100 mW/cm^2 of electrode. This power density is obtained from the ratio of the calculated capacitance of the electrodes (assuming the dielectric constant of vacuum) to the measured stray capacitance of the entire system. Resulting growth rates were in the 0.3 – 3.0 μm/h range with 1.0 μm/h typical. The percentage of silane in argon ranged from 2.7 to 100 mol %, with vitually all of the extensively studied samples in the unresolvable columnar morphology growth regime as characterized earlier by Knights and Lujan.[39] In fact, we were unable to observe columnar fracture morphology even in scanning electron micrographs of 1 – 2 μm films grown at 2.7 mol % SiH_4 with maximum rf power, conditions most likely to give visibly columnar growth.

Films were doped with PH_3 in the range of 0 – 3000 volume parts per million (vppm) referenced to the silane gas flow. In order to make the films conductive enough to make the necessary capacitance measurements described in the rest of this paper, films had to be doped greater than about 30 vppm. The two exceptions to this rule were a set of undoped films which were used to measure the density of states in undoped material at the Fermi level (see Sec. VI) and an anomalous sample, 1 (106), in Table I, which, though it was nominally undoped, had a shallow Fermi level and hence acceptably high conductivity due to some as yet unexplained impurity. The origin of such an impurity has been difficult to identify since it was detected in only the first dozen or so samples grown in the new system. It is possible, however, that is is related to the incorporation of traces of fluorine, known to influence dramatically the properties of this material,[40] remaining from a standard CF_4-O_2 plasma etch of the growth system performed between each run in that batch of early samples. Subsequently, this etch procedure was discontinued and the growth system rebuilt. Several attempts to duplicate such conditions, even with CF_4-O_2 plasma pre-etches, have been unsuccessful. This anomalous sample does provide us with an excellent vehicle to study the density of

states of a phosphine-free film, however, as will be discussed in later sections.

Samples were deposited on a variety of substrate materials, often simultaneously, depending on the particular measurements to be done. The most common substrate material was heavily doped (10^{19} cm^{-3}), n- or p-type polished crystalline silicon (denoted n^+-Si or p^+-Si). Low-doped ($< 10^{15}$ cm^{-3}) silicon wafers polished on both sides were employed for infrared-absorption measurements. Before being introduced into the growth system the pieces of silicon were given an extensive surface cleaning and etching procedure terminating with a hot $HCl:H_2O_2$ etch,[41] resulting in a controlled 10–20 Å surface oxide, in order to assure a clean and repeatable surface. When compared with various simpler silicon etch pretreatments (e.g., HF; or 8 mol % HCl:20 mol % H_2O at 80 °C, or degreasing only without etching) such a treatment was found to form a superior junction with the a-Si:H film for the various measurements explained later in this paper. Other substrate materials used at one time or another in the study include polished fused silica (etched in 33 mol % HNO_3, 4 mol % HF, balance H_2O with detergent), evaporated chromium (300–1000 Å thick) on similarly prepared fused silica, and polished tantalum (either degreased only or degreased and etched in $5:HNO_3 + 1:HF$).

Problems associated with cracking of the film material were predominantly, though not exclusively, confined to the crystalline silicon substrates. This effect was presumably due to cumulative strain resulting from the difference in the thermal expansion coefficient of the film and the substrate during cooling from the growth temperature. However, in samples where this proved to be a problem, a combination of slow, controlled cooling (~5 °C/min) together with a minimization of the exposure time to air between the a-Si:H deposition and the top-electrode evaporation was usually sufficient to avoid the problem.

In addition to the measurements discussed in subsequent sections of this paper, samples were characterized in a number of more standard ways to insure that they were typical of the a-Si:H material reported elsewhere in the literature. Both ac dark conductivity and the position of the Fermi level as a function of doping confirmed that this material behaved in the same way as the a-Si:H material originally reported by LeComber and Spear.[2] The results of hydrogen effusion measurements on these samples, together with ESR measurements, both dark and light induced, luminescence, and infrared absorption, all indicate that the material being considered in this study is quite representative of the optimized glow-discharge a-Si:H reported in the literature.

The post-growth sample preparation and mounting necessary for the series of measurements to be discussed in the rest of this paper deserves elaboration. As mentioned earlier, the most common substrate material for these measurements was either n^+-Si or p^+-Si. Semitransparent top electrodes were formed over the a-Si:H film by evaporation of ~150-Å Cr followed by ~250-Å Au through a molybdenum shadow mask to form an array of 0.5-mm diameter dots which served as Schottky-barrier contacts. Alternately, to determine the degree of current spreading beyond the top electrode-defined area, a few samples were covered uniformly with Cr-Au, masked with a series of 0.5-mm steel balls and sand blasted from above through to the substrate, leaving an area physically defined by the remaining mesa. For a few of the measurements on the n-type films, the back contact with the substrate was made ohmic by using an n^+-Si substrate and growing the first ~1000 Å of a-Si:H heavily doped (10 000 vppm PH_3). For the most part, however, it was found to be an advantage to have a barrier at the back contact as well as the front Schottky by growing the n- or p-type a-Si:H film directly on a p^+-Si or n^+-Si substrate, respectively. Then, by varying the bias on the sandwich, now consisting of two diodes of reverse polarity, one could deplete either the front or back junction, and hence look at two different spatial regions of the sample with two completely different interfaces present in order to help further in discrimination of bulk from surface-related effects. Rear Schottky contacts were also tried for certain specialized measurements by using a tantalum or Cr-on-fused-silica substrate, but the buried p^+-n junction at the substrate interface proved to be superior in general because it had the lowest leakage in reverse bias. Ohmic contact was easily made to the back of the heavily doped silicon wafer by rubbing in an In-Ga eutectic mixture. Subsequently the sample was cleaved into 2-mm square pieces and glued to a TO-18 transistor header by thermally setting silver epoxy, thus making the header case one contact of the composite diode. The other contact was made by thermocompression bonding a 25-μm diameter gold wire between one of the 0.5-mm top electrodes and one of the header pins.

Samples thus mounted were in a convenient

form allowing rapid interchange on the probe for measurements involving a temperture sweep with the top transparent electrode facing outward for possible laser illumination. Before each of the measurements, the samples were heated in the dark to 200 °C at zero bias and held there for at least 5 min to place them into the annealed high-dark-conductivity state, labeled state *A* by Staebler and Wronski.[42] Samples which had been converted into state *B* by sufficient exposure to 6000-Å laser illumination showed significantly different properties which will be discussed elsewhere.[43] For the purpose of this paper, all measurements were made with the samples in state *A*. In particular, we took great pains to be sure that the laser excitation necessary for obtaining the DLTS or TSCAP hole-trap spectrum did not also convert the sample into state *B*. It was also found that the properties of the samples were altered by heating above 100 °C in the dark with applied bias. This effect could be reversed by the state-*A* anneal sequence and will be discussed elsewhere.[44] The net effect for the purpose of this paper is that all TSCAP and DLTS thermal scans were limited to a maximum temperature of 100 °C.

The sample numbers used in this paper refer to the diodes made from a specific growth run, e.g., sample 5 (152), is characteristic of diodes from run number 152. Since the details of the density of states depend specifically on the growth run, the sample number is important in labeling the results for this and subsequent publications.

III. STEADY-STATE CAPACITANCE MEASUREMENTS

A. Capacitance versus bias voltage (*C-V*)

One of the most straightforward measurements which can be made on our sandwich-type sample geometry is a standard *C-V* plot. Typical data are shown in Fig. 1 for a 300-ppm $-PH_3$-doped sample measured at 100 kHz in the dark at room temperature. The most striking feature of this *C-V* curve is the sharp peak in capacitance at approximately zero bias. This is typical of nearly all of our samples and indicates barrierlike behavior on both sides of the *a*-Si:H film. As discussed in Sec. II, however, this is useful in assessing the homogeneity of the bulk $g(E)$ and greatly aids in separating bulk from interface effects.

In our sign convention a negative-bias voltage

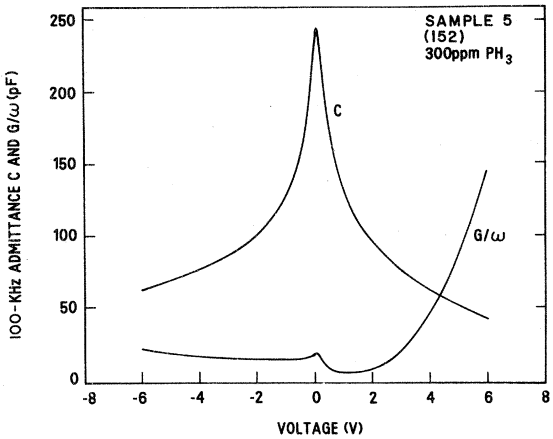

FIG. 1. Real (G/ω) and imaginary (C) parts of the steady-state complex admittance ($Y=G+i\omega C$) vs applied voltage for a typical PH_3-doped *a*-Si:H sample. The measurements were made at 300 K and 100 kHz.

corresponds to a reverse bias on the rear (substrate—*a*-Si:H) interface with a forward bias on the front (Schottky-barrier) interface. Conversely, the positive voltage scale in Fig. 1 corresponds to a reverse bias on the Schottky barrier with a forward bias on the substrate interface. Thus, for all biases farther than a few tenths of a volt from zero the sample behaves as if it were a small capacitor (the reverse biased interface) in series with much larger capacitor (the forward biased interface). Under such circumstances one may safely ignore the contribution to the overall capacitance from the forward biased side of the sample. We have verified the validity of this approach by measuring several n^+Si$-n^+$ *a*-Si:H$-n$ *a*-Si:H samples which have a reasonably good ohmic contact at the substrate interface. For such single-diode structures the capacitance does not peak at $V\cong 0$ as in Fig. 1 but increases rapidly when forward bias is applied to the Schottky barrier, as is expected for a single-barrier structure.

Another parameter which we routinely measure along with the capacitance C is the ac conductance G of the sample. In general we measure the complex admittance ($Y=G+i\omega C$) of the sample with a computer-controlled HP4274A *LCR* meter in the parallel *RC* equivalent-circuit mode using various frequencies and amplitudes of the measurement voltage. The curve labeled G/ω in Fig. 1 is related in part to the shunt leakage current thermally activated over (or tunneling through) the barrier in phase with the bias oscillation. By plotting G/ω and C, both the real and imaginary components of the current are displayed on the same scale in

capacitance units. As we will discuss later, the conductance may also include a trapping-detrapping component which is not necessarily related to the shunt leakage path. In general the leakage current as a given bias increases with increased PH_3 doping which effectively limits the maximum usable doping range for most of our measurements to less than 3000 ppm PH_3.

For the case of crystalline semiconductors where the dopant introduces a well-defined, shallow discrete state in the gap, the C-V plot is routinely used to evaluate the net dopant concentration by using the relationship

$$N_{CV}=-\frac{C^3}{\epsilon q A^2}\left[\frac{dC}{dV}\right]^{-1}, \tag{1}$$

where ϵ and A are the dielectric constant and area of the sample, respectively, and q is the electron charge. The quantity N_{CV} is well-defined for Schottky barriers or asymmetric (p^+n or n^+p) junctions in crystals with no deep levels. In such cases we have

$$N_{CV}=\begin{cases}N_D-N_A\ , & (n\text{ type})\\ N_A-N_D\ , & (p\text{ type})\ ,\end{cases} \tag{2}$$

where N_D (N_A) is the donor (acceptor) concentration. When deep levels are present, N_{CV} will include them in addition to the shallow levels if the voltage sweep used to measure C-V is slow enough so that the deep-level occupation can respond on the same timescale. Thus in the presence of significant concentrations of deep levels the meaning of N_{CV} depends on the manner in which the C-V data are taken.

The problem of interpreting C-V data is especially acute for a-Si:H since deep states often dominate the capacitance response. In this case the meaning of the C-V data and of N_{CV} is affected not only by the rate of the voltage sweep but also by the frequency and temperature of the measurement. Thus to properly analyze capacitance data we must also consider its temperature and frequency dependence. This we will do in Secs. II B–II D.

B. Capacitance versus temperature (C-T)

Perhaps one of the simplest ways to gain insight into the dynamic response of an a-Si:H barrier structure is to measure the capacitance at a fixed bias over a wide temperature range. This is also a necessary preliminary step to subsequent DLTS measurements. The typical data for three samples in Fig. 2 illustrate the types of changes in C-T plots which occur as a result of PH_3 doping.

A dominant feature of the C-T curves in Fig. 2 is the distinct step in capacitance for the two PH_3-doped films. The temperature of this step is the demarcation line between an active as opposed to a passive role for the a-Si:H film in a sandwich-structure capacitor. For temperatures below the step the a-Si:H film behaves like a passive dielectric medium which is not polarizable on the timescale of the measurement frequency (10 kHz in Fig. 2). Thus the measured capacitance of the structure is simply that of a parallel-plate capacitor with insulating a-Si:H as the dielectric. The low-tempertaure capacitance C_{LT} is proportioned to the inverse of the film thickness l, i.e.,

$$C_{LT}=\epsilon A/l\ . \tag{3}$$

As the temperature increases through the region of the step, the conductivity of the film becomes large enough so that current can pass through the film on the timescale of the measurement frequency and charge the depletion layer capacitors at each interface, i.e, the RC time constant of the structure is fast enough so that it can respond to the measurement voltage. One could also describe this capacitance "turn-on" by saying that the dielectric response time of the film is short enough so that it can behave as a polarizable dielectric on the timescale of the measurement frequency. This capacitance step is also sometimes referred to as the capacitance "freeze-out" when observed during a scan from higher to lower temperature. The zero-bias capacitance for temperatures just above

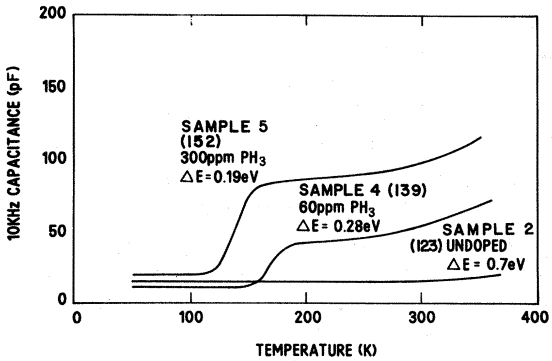

FIG. 2. 10-kHz steady-state capacitance vs temperature for three samples of different doping showing the turn on of the dielectric response at increasingly lower temperatures for more heavily doped films. The quanitity ΔE is the conductivity activation energy. The data were recored in all three cases in the dark with 2-V reverse bias on the front (Schottky) barrier.

the step is

$$C=\epsilon A/2W\ , \tag{4}$$

where W is the width of the depletion layers on each side of the film. The resistance of the bulk undepleted material in the interior of the film is then

$$R=\rho(l-2W)/A\ , \tag{5}$$

where ρ is the bulk resistivity.

The turn-on criterion is $\omega^{-1}=RC$, where ω is the angular frequency of the capacitance measurement. It is difficult, however, to exactly solve the problem at the turn-on temperature since one must explicitly consider both the transport equation and Poisson's equation simultaneously. One may approximately evaluate the turn-on criterion, however, for two limiting cases. On the low-temperature side of turn-on, where bulk transport parameters dominate the film's response, we can estimate $R\simeq\rho l/A$ and $C\simeq C_{\rm LT}=\epsilon A/l$, thus

$$\omega^{-1}\simeq\rho\epsilon\ . \tag{6}$$

The expression $\rho\epsilon$ is often referred to as the dielectric relaxation time of the material. For temperatures just above the turn-on point we may use Eqs. (4) and (5) to obtain

$$\omega^{-1}=\rho\epsilon\left[\frac{C}{C_{\rm LT}}-1\right]\ . \tag{7}$$

In cases such as Fig. 2 where there is a rather large and well-defined capacitance turn-on step we may have a capacitance ratio $C/C_{\rm LT}$ as large as 10, but typical values are in the range of 3 to 5.

For the typical measurement frequency of 10 kHz ($\omega=6.28\times10^4\ {\rm s}^{-1}$) in Fig. 2, we may estimate the resistivity at turn-on to be between $\rho=1.5\times10^7$ Ω cm on the low-temperature side of the step [Eq. (6)] and $\rho=3\times10^6$ Ω cm on the high-temperature side of the step [Eq. (7) with $C/C_{\rm LT}=5$]. Thus, the two PH_3-doped samples are conductive enough so that $\rho\sim10^6$ Ω cm at approximately 150 K while the undoped film is much more resistive and is just reaching $\rho\sim10^7$ Ω cm at 350 K. Because of the difficulties of analysis[10] we have not made planar dc resistivity measurements on our samples, but the values of ρ which can be inferred from the turn-on criterion are typical of high-quality glow-discharge films.[2,10] We may measure the *activation energy* of the conductivity quite accuratey, however, by considering the frequency dependence of the C-T curve, as will be discussed in the next section.

All of our data used for the determination of $g(E)$ are taken at temperatures above the turn-on point. Thus it is essential to have C-T measurements such as in Fig. 2 in order to assess the range of validity of our analysis. Although it is possible to draw some conclusions about $g(E)$ directly from the C-T measurement alone,[20,21,35] we find such analysis most useful and reliable as a corroboration for the $g(E)$ determined from DLTS.

C. Capacitance versus frequency (C-ω)

Measurements of capacitance versus frequency[20,21,23–25] in a-Si:H give in principle the same information as C-T measurements. Since C-ω measurements can be made at room temperature without a variable-temperature cryostat, they are also considerably easier to do. The relationship between C-T and C-ω depends on the basic phenomenon controlling the dynamic electrical response of the sample. When the dynamics are controlled by a thermally activated variable such as the resistivity or thermal detrapping rate, then the temperature scale is isomorphic to a logarithmic frequency scale. In such a case there is a clear advantage to C-T measurements because of the range of dynamic effects which may be probed. As a practical matter, C-ω measurements of a-Si:H samples are usually limited to the frequency range of 10^{-3} to 10^5 Hz.[20,21,23–25] Temperature variation, on the other hand, can probe many more decades of effective frequency for thermally activated processes. For example, a readily obtainable temperature scan of 100 to 400 K covers *25 decades* of the activated variable for an activation energy of 0.65 eV. The activation energy probed in a temperature scan at fixed frequency is proportional to the temperature, whereas for frequency measurements at room temperature each decade change in frequency is equivalent to a $2.3kT=57$ meV change in activation energy.

D. Frequency dependence of capacitance versus temperature (C-T-ω)

A method which is much more useful than either C-T or C-ω alone is to measure C vs T *and* ω simultaneously. The data are most conveniently presented as a family of C-T plots for various values of ω as in Fig. 3. This figure shows the capacitance and conductance versus temperature at three frequencies for the front and rear depletion layers of a 300-ppm PH_3-doped sample. Let us

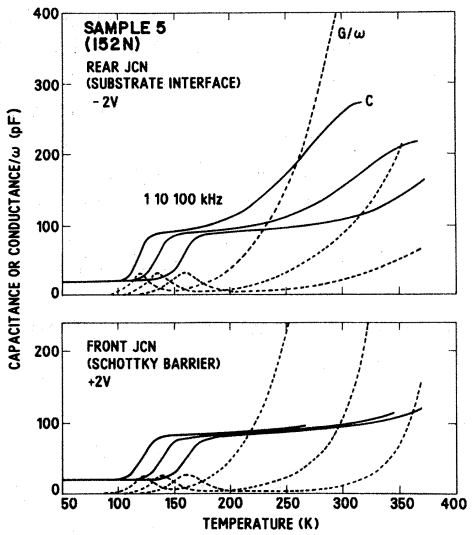

FIG. 3. Steady-state capacitance (solid lines) and conductance divided by angular frequency (dashed lines) vs temperature for three measurement frequencies. Note the difference between the behavior of the space-charge layers at the front and rear junctions.

first discuss the turn-on portion of the C-T curves in the temperature range 100—200 K. Note that as the frequency is changed the turn-on temperature shifts. Since the turn-on behavior is controlled primarily by the bulk transport according to Eqs. (6) and (7), this shift in the turn-on temperature versus frequency may be used to obtain the activation energy of the bulk conductivity of the sample. This is equal to the slope of a plot of the logarithm of the measurement frequency versus the inverse of the turn-on temperature. For the data in Fig. 3 the result is $\Delta E = 0.19$ eV. The activation energies in Fig. 2 were obtained in a similar manner. Note that this method of measuring ΔE is advantageous in that it is not subject to the surface band-bending problems which strongly influence dc resistivity measurements.[10] Indeed, the depletion layers play an integral role in our measurement. The values of ΔE obtained in this way agree very well with the conductivity activation energies versus PH_3-SiH_4 ratio as measured by Spear and co-workers.[2]

The conductance curves in Fig. 3 show both the smoothly varying shunt leakage as well as a peak at the turn-on temperature. The conductance turn-on peak is intimately related to the capacitance step according to linear response theory.[45,46] It is often called the dielectric loss peak and in G/ω units is always $\frac{1}{2}$ of the capacitance step.[45] Such an effect is purely an ac phenomenon and has no relationship to the dc leakage current. On the other hand, the fact that G/ω does not return to zero above the turn-on peak in Fig. 3 is due primarily to true leakage conductance.

The high-temperature portion ($T > 200$ K) of the admittance versus temperature and frequency curves in Fig. 3 shows several important effects which are quite typical of our samples. First, we note that the conductance increases monotonically with temperature until the dissipation factor $D = G/\omega C$ is greater than 10 so that C is no longer measurable by our LCR meter. Note also that while G/ω is quite different for the three frequencies, this is primarily an artifact of the division by ω. The value of G alone is rather independent of frequency in this temperature range, which, indeed, may be taken as the signature of a true dc leakage path as opposed to dielectric loss due to trapping and emission of carriers at bulk or interface states. The second notable feature of the temperature regime above turn-on in Fig. 3 is the significant difference between the capacitance response of the front and rear junctions. This is quite typical of most of our samples and may be due to several different effects; e.g., current spreading, bulk gap states, or interface states.

Let us first consider the current spreading effect. This term refers to the fact that the effective area of the substrate junction is limited only by the resistivity of the film (and consequent current spreading) from the well-defined area of the front Schottky barrier. The front Schottky barrier is 500 μm in diameter so that for a 1—2-μm film the effective area of the rear junction will begin to increase noticeably when the bulk conductivity of the film is roughly 3—4 orders of magnitude above its value at the turn-on temperature. Since the measured capacitance is proportional to the junction area, we see a monotonic capacitance increase at this point. We see no such current spreading effects in the few diodes which were fabricated in the more complex mesa structure. Current spreading can be recognized by the fact that the activation energy of the onset of the spreading effect is the same as that of the turn-on step.

For many samples, we observe above turn-on a well-defined step in capacitance (and peak in conductance) which does not have the same activation energy as the bulk conductivity. This is therefore not bulk current spreading but must be due to deep states. A particularly extreme example of this phenomenon is shown in Fig. 4 were we see *three* capacitance steps (and three corresponding conduc-

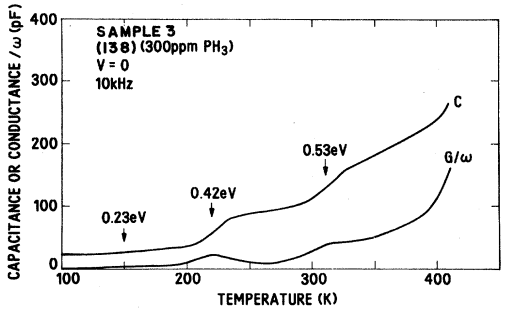

FIG. 4. Steady-state zero-bias capacitance (C) and conductance divided by angular frequency (G/ω) vs temperature for a sample exhibiting three well-defined capacitance steps (and corresponding conductance peaks).

tance peaks). The turn-on step (bulk conductivity) has an activation energy of $\Delta E = 0.23$ eV, while the two higher-temperature steps have $\Delta E = 0.42$ and 0.53 eV, respectively. The two steps above turn-on are caused by the onset of an additional polarizability of the depletion layers. This additional polarizability is due to the trapping and emission of carriers at deep gap states either in the bulk of the depletion layer or at the interface. Indeed, the analysis of the conductance due to dielectric losses is a well-established method for studying oxide-semiconductor interface states in MOS capacitors[26–29,47] and bulk deep levels in crystalline semiconductors.[45,46]

It is difficult to determine from a single set of fixed-bias C-T-ω data alone whether the capacitance increases above turn-on are due to bulk or interface states. However, one can use the front-versus-rear asymmetry as a clue that interface states may be responsible, since the front and back interfaces are usually quite different. In addition one can use the theory of Ref. 35 to calculate the C-T-ω response corresponding to the bulk $g(E)$ determined from TSCAP and DLTS experiments. Any significant deviations from the calculated bulk C-T-ω curves are then most likely due to interface states. This follows since one can show in general that *steady-state* C-T-ω measurements have a maximum sensitivity to the emission of trapped charge at the interface while capacitance *transient* measurements (TSCAP or DLTS) have *zero* sensitivity at the interface and very low sensitivity to the bulk near the interface.[30,35] Thus, comparisons of steady-state and transient capacitance data is a way to distinguish between bulk and interface states. We should note here, that our use of the term "interface states," while it is consistent with the a-Si:H literature, is not meant to be restricted only to the true two-dimensional interface between a-Si:H and the metal barrier (front junction) or the thin (10–20 Å) oxide on the silicon substrate (rear junction). In terms of the measurements we are discussing it is difficult to distinguish between such *true* interface states and bulk states that are located *near* the interface, i.e., within several hundreds of angstroms. At the risk of some potential confusion, but in the interests of brevity, we shall refer throughout this and the next paper to both situations as simply "interface" states.

IV. CAPACITANCE TRANSIENT MEASUREMENTS

A. General aspects of junction transient measurements

The basic physics of transient junction measurements is shown in Fig. 5. Here we show the potential-energy diagram of the band bending of a typical n-type Schottky barrier at two bias voltages: zero bias and a reverse bias, V_R, equal to roughly twice the band gap. The portions of the gap state distribution which are filled with electrons in equilibrium are roughly indicated by diagonal shading in both cases. The *difference* between the equilibrium gap state occupations for the two bias voltages is shown by dotted shading.[35] These difference states are full in equilibrium at zero bias but empty under steady-state reverse bias conditions. Thus, immediately after the bias is switched from zero to V_R, the dotted states in Fig. 5 are metastable and will begin to empty. The emission of these metastable trapped electrons may be observed directly by measuring either the current flow in the external circuit or the change in the junction capacitance. This is the essential physics of transient measurements.

Both current and capacitance transient measurements are in principle related to the same basic physical phenomenon. However, there are subtle practical differences between these different measurement modes which influence their relative signal-to-noise ratios and sensitivites to extraneous effects. Current measurements, for example, are subject to leakage current, thermoelectric currents, and secondary photocurrent. A distinct advantage of capacitance measurements is that the *sign* of ΔC depends on the sign of the trapped charge whereas current transients are all one sign.

In addition, there is an important difference be-

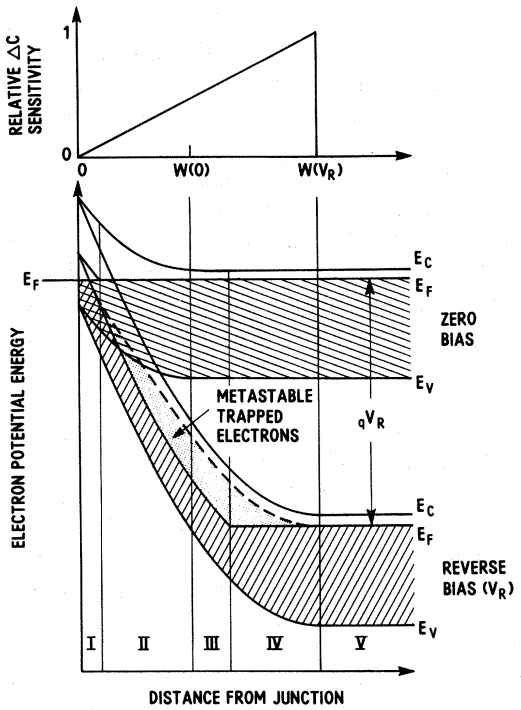

FIG. 5. Potential-energy (band-bending) diagram of an idealized barrier at both zero bias and an applied reverse bias of V_R. Regions of gap states which are filled under steady-state conditions are indicated for both biases by diagonal shading. The energy and spatial region where initially filled gap states are metastable immediately after the bias is switched from zero to V_R is shown by dotted shading. The spatially varying sensitivity of capacitance changes to the emission of a metastable trapped unit charge is shown in the upper part of the figure.

tween capacitance and current transients in their relative sensitivity to interface states. We show at the top of Fig. 5 the relative variation in spatial sensitivity for capacitance transients. The sensitivity at position x is defined as the capacitance change ΔC which would be induced by a unit charge at x being removed from the depletion region. The displayed linear spatial variation follows from the requirement that the dipole moment of the total junction space charge must remain constant for a constant applied bias.[30] This means, in particular, that charges removed at the thin (10–20 Å) surface-oxide interface of a Schottky barrier ($x \cong 0$) will produce a vanishingly small capacitance transient since such charges have essentially zero dipole moment with respect to the metal contact.

The relative spatial sensitivity of a *current* transient, on the other hand, is *maximum* at the surface ($x=0$) and falls linearly to zero at $x = W(V_R)$.[30] This is because only those emitted charges which are not part of the displacement current will be measured in the external circuit. It follows that this surface-weighted sensitivity profile for current transients is also appropriate to steady-state capacitance measurements since a capacitance meter is actually a phase-sensitive ac *current* measuring instrument. Thus features in steady-state capacitance, TSC, current transient, or current DLTS which do not appear in analogous TSCAP, ΔC transient, or capacitance DLTS are most likely due to states at the surface.

The interpretation of a capacitance or current transient depends not only on spatial variations but also on the initial and final occupation statistics of $g(E)$ in the depletion layer space charge. One type of initial condition has already been mentioned in connection with Fig. 5, namely, filling all states with electrons up to the zero-bias Fermi level. This initial condition can be modified somewhat to give information on the spatial variation of trap density by using an intermediate value of bias ($0 < V < V_R$) during the trap-filling phase.[30] However, using a reduction in bias voltage to fill gap states can only produce transients corresponding to majority-carrier traps. For n-type material this corresponds to electron emission from states in the upper half of the gap. In order to create a metastable initial condition for minority-carrier traps (hole traps in n-type material), one must either introduce mobile minority carriers to be captured or else directly excite the gap states by subgap light. In this paper we use 6000-Å laser excitation to create e-h pairs and thus to perturb the gap state occupation by carrier capture.

Since one measures the distribution of *occupied* metastable states in a junction transient experiment, the ultimate interpretation of the results in terms of $g(E)$ requires that one know the metastable occupation statistics of the initial state. This is often difficult to determine *a priori* for laser pulse excitation, but as we will discuss later, the data indicate that the laser-induced occupation is reasonably uniform in energy. The situation is easier for voltage pulse excitation since one knows the equilibrium occupation statistics at zero bias. As shown in Fig. 5, the zero-bias Fermi level is flat throughout the depletion layer and controlled by its values at the interface and in the bulk.[47] The *final*-state-occupation statistics can also be analyzed in a straightforward manner. In this case, however, one must know the spatial depen-

dence of the quasi-Fermi level under reverse bias.[47] The electron quasi-Fermi level for $V=V_R$ in Fig. 5 is shown as the solid line at approximately midgap in the spatial regions marked I, II, and III. In regions IV and V the true bulk Fermi level controls the electron occupation. In Ref. 35 we show that the electron quasi-Fermi level E_F^* is given for regions II and III of a reverse-biased diffusion-limited Schottky barrier[47] by

$$E_c - E_F^* = \frac{E_g}{2} + \frac{kT}{2} \ln\left[\frac{\nu_n}{\nu_p}\right], \tag{8}$$

for the limit of negligible reverse saturation ("leakage") current or for $\Phi_B > E_g/2$ and

$$E_c - E_F^*(x) = \Phi_B - kT \ln[\beta \mathscr{E}_0 / \mathscr{E}(x)], \tag{9}$$

for the limit of large reverse saturation current if $\Phi_B < E_g/2$. In these expressions, E_F^* is the electron quasi-Fermi-level, E_g is the band gap, ν_n and ν_p are the electron and hole emission rate prefactors for states near midgap, Φ_B is the Schottky-barrier height, β is the transmission coefficient ($0 \leq \beta \leq 1$) of the thin oxide layer between the Schottky-barrier metal and the semiconductor, $\mathscr{E}_0$ is the electric field at the Schottky-barrier interface ($x=0$), and $\mathscr{E}(x)$ is the electric field at point x within the depletion layer. Thus we see from Eqs. (8) and (9) that the deepest states which may be seen by an electron emission transient are near midgap in the nonleaky or large barrier case and near a depth equal to the barrier height in the case of a leaky barrier with $\Phi_B < E_g/2$. The quantitative criterion which controls the choice of limits is discussed in detail in Ref. 35. The barrier heights for undoped and lightly PH_3-doped a-Si:H are between 0.8 and 1.1 eV,[9,20] thus either limit would predict E_F^* at roughly midgap as shown in Fig. 5.

The general considerations of this subsection apply equally to the many forms of capacitance or current transient experiments. In Secs. IV B and IV C we will discuss in more detail the two forms which we have found most useful in our studies, namely TSCAP and capacitance DLTS.

B. Thermally stimulated capacitance

The techniques of thermally stimulated capacitance (TSCAP) and thermally stimulated current (TSC) are the simplest methods for surveying the distribution of gap states in semiconductors and have been widely used over the past 15 years in crystals.[30,45] The TSC method has recently been applied to a-Si:H.[33,34] We have used both TSC and TSCAP in our studies of a-Si:H and have developed a definite preference for TSCAP over TSC. One of the major difficulties which we have encountered with TSC is common to most current measurements and was mentioned earlier, namely, the sensitivity to extraneous currents, thermal emission from interface states, leakage currents, thermoelectric currents, and secondary photocurrents. As discussed eariler, capacitance transient measurements such as TSCAP or capacitance DLTS are not strongly affected by these phenomena and hence are much more straightforward to interpret.

The TSCAP technique is essentially the observation of a slow capacitance transient during a temperature scan. As with an isothermal transient, the gap states in the depletion layer are first prepared with some initial metastable occupation which is then "frozen in" at the lowest temperature of the scan. As shown in Fig. 6, the capacitance is recorded as a function of temperature during the warm-up scan in the same manner as the C-T data of Sec. III. The signal corresponding to the metastable initial state slowly changes with increased temperature due to thermal emission of trapped charge until it coincides with the steady-state C-T curve (labeled "baseline" in Fig. 6). The typical TSCAP experiment in Fig. 6 shows three C-T curves recorded at the same bias (-2 V) for the two different metastable initial conditions as well as the steady-state baseline curve for this bias. The C-T curve labeled "zero bias fill" corresponds to the initial condition of Fig. 5, namely, all gap states are filled with electrons up to the zero-bias Fermi level. The curve labeled "laser fill" corresponds to the initial condition of saturation with 6000-Å laser illumination at low temperature ($\lesssim 100$ K).

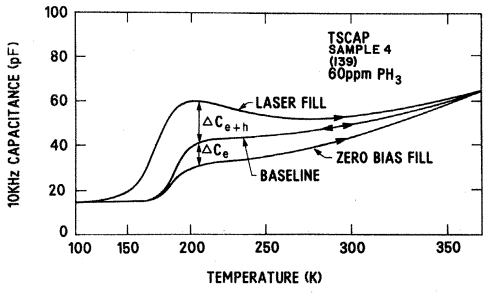

FIG. 6. Typical set of the thermally stimulated capacitance (TSCAP) temperature scans for three different initial conditions as described in the text.

The negative difference ΔC_e between the zero-bias-fill and baseline C-T curves just above the dielectric response turn-on temperature is related to the integral of $g(E)$ between about midgap and some energy in the upper half of the gap which depends logarithmically on the thermal scan rate and linearly on the turn-on temperature. For the case of Fig. 6 the upper energy limit of this integral is about $E_c - 0.55$ eV. For the caes of the laser-fill C-T curve in Fig. 6, the positive difference ΔC_{e+h} is related to the integral of trapped holes between about $E_V + 0.55$ eV and midgap *minus* the same trapped electron integral as in ΔC_e. We relate these metastable capacitance differences to the $g(E)$ integrals by numerical solutions of Poisson's equation and our analysis of the diode admittance as discussed in Ref. 35. The specific relations for our typical experimental conditions are discussed in Sec. VI C.

One could in principle obtain $g(E)$ from the derivatives of the differences between the metastable C-T curves and the baseline. For an ideal system this is formally the same as obtaining $g(E)$ from the DLTS data of the next section, although the range of energies covered is small with TSCAP. However, for a number of practical experimental reasons we believe that the DLTS spectra are more reliable for the qualitative *shape* of $g(E)$ while in most cases TSCAP is most reliable in setting the overall quantitative scale of $g(E)$.

The different and complementary aspects of TSCAP versus DLTS should be more clear after we have discussed DLTS. For now let us note that a major difference is that TSCAP is a single-shot transient experiment whereas DLTS is based on repetitive transients. Thus, an advantage of TSCAP is that it is possible to spend sufficient time to carefully prepare the initial occupation conditions. This makes it easier to prove that the metastable effects are fully saturated, and hence accurately reflect the true magnitude of $g(E)$. For example, the zero-bias-fill sequence can be initiated at room temperature or above and held there for hours if necessary before the cool down to guarantee the achievement of thermal equilibrium at zero bias before the switch to the reverse bias scan voltage. The initial conditions for DLTS must by necessity be repetitively set on a much shorter timescale, and thus are more subject to incomplete saturation, although even in the DLTS case we have experimentally demonstrated saturation in all samples studied in detail. On the other hand, the changes in the metastable occupation conditions must accurately reflect *only* thermal-emission processes if the data are to be interpreted as $g(E)$, i.e., the energy scale must be related to the dynamics of thermal emission. This condition is much more difficult to fulfill for TSCAP than for DLTS because there is a greater likelihood that extraneous effects such as capture of leakage current or deep-state-hopping conduction could affect the high-temperature portions of the very slow (~ 10 min) TSCAP metastable C-T scan as opposed to the relatively fast DLTS transient (20—300 msec). We can show, however, by observing low-temperature isothermal transients that the magnitude of the metastable trapped charge corresponding to ΔC_e and ΔC_{e+h} has most likely not been altered by such extraneous effects. Therefore, we can trust the *integrals* of $g(E)$ obtained from the low-temperature TSCAP results, but are concerned that the metastable capacitance differences remaining by the time the scan has reached higher temperatures may have been prematurely altered by slow effects other than the thermal emission of trapped carries to the bulk conduction path. Indeed, such differences between TSCAP and DLTS offer the future possibility of studying alternate conduction paths and bulk leakage effects.

C. Deep-level transient spectroscopy: Review of the discrete level case

The DLTS form of capacitance transient spectroscopy has been widely used since its introduction in 1974 to study deep levels in crystalline semiconductors. The basic concepts and details of implementation and analysis for the case of discrete defect states in the bulk of crystalline semiconductors have been adequately discussed in the original reference[31] and subsequent reviews.[30,45] However, for the sake of the reader who may be unfamiliar with the technique, we will outline the main concepts here as well.

The main idea of DLTS is essentially the synchronous detection of repetitive, thermally activated transients of a given time constant. It can be shown[35] that under such conditions the DLTS algorithm constitutes essentially a spectrum analyzer of the impulse response of the trap system, i.e., the ideal DLTS spectrum is directly related to the Laplace transform of the transient time dependence. This is the *definition* of the thermal emission density of states for the case of energy-independent matrix elements between the free and bound states.

The DLTS algorithm can be illustrated and explained most simply for the case of a discrete trap level where a double boxcar (dual-gated integrator) is used as the DLTS signal processor. This is shown schematically in Fig. 7 with the two gate times denoted as t_1 and t_2. The DLTS signal is defined in this case as the difference between the transient amplitude at time t_1 minus the amplitude at time t_2. For the case of a capacitance transient this difference is denoted $C(t_1)-C(t_2)$. As a function of temperature the $C(t_1)-C(t_2)$ difference signal for a discrete level with a well-defined activation energy goes through a maximum at the temperature where the transient time constant is on the order of the gate spacing. For this very simple case, we show in a straightforward manner that the time constant $\tau_{\max}$ corresponding to the maximum DLTS signal is

$$\tau_{\max}=\frac{t_2-t_1}{\ln(t_2/t_1)} . \tag{10}$$

The inverse of this time is called the DLTS "rate window." It is obvious, of course, that in practice the boxcar gates at t_1 and t_2 must have finite widths. In fact, for optimum signal-to-noise ratio the gate widths should be as large as possible.[30,45]

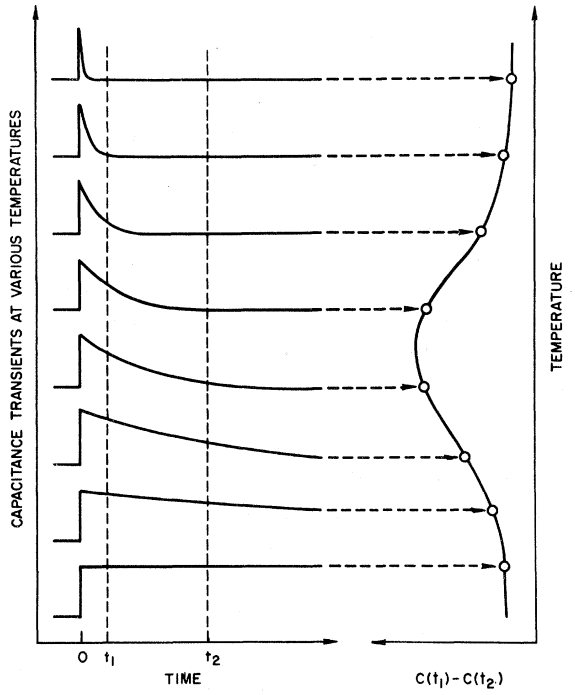

FIG. 7. Basic principle of the DLTS transient-analysis technique as illustrated for emission from a discrete trap level.

In this case the simple derivation of Eq. (10) is no longer valid. It has been shown,[48] however, that even in the wide-gate limit the expression in Eq. (10) is still reasonably close to the true solution provided t_1 and t_2 are taken as the *midpoints* of the gates

A DLTS *spectrum* is defined as the $C(t_2)-C(t_1)$ difference signal plotted as a function of temperature. For the case of several sharp levels this specturm consists of well-defined peaks of width ΔT, where ΔT for each peak is about 10% of the temperature of that peak. The energy scale of such a spectrum is related to the thermal activation energy of the detrapping process. The electron thermal-emission rate e_n for a particular trap level of depth E is given by

$$e_n=\nu_0\exp(-E/kT) , \tag{11}$$

where ν_0 may be interpreted as an attempt-to-escape frequency with the Boltzmann factor being the escape probability. In the small signal limit, the time constant of the exponential capacitance transient due to this trap is the inverse of the thermal emission rate given in Eq. (11). Thus an Arrhenius plot of $\log\tau_{\max}$ versus the inverse temperature of the DLTS peaks for several spectra recorded with different rate windows (i.e., different t_1 and t_2) may be analyzed to give both ν_0 and E for each trap level.

For most deep levels in crystalline semiconductors the exponential prefactor in Eq. (11) is in the range $10^{11}<\nu_0<10^{13}\ \mathrm{s}^{-1}$. As is often pointed out, the largest values in this range are roughly of the order of acoustic-phonon frequencies. However, in a more realistic treatment of the thermal emission process, the ν_0 factor is actually related to the product of the matrix element for the bound-to-free transition and the effective density of final states N_c within approximately kT of the bottom of the conduction band. A well-known expression derived from detailed balance arguments relates ν_0 to these more fundamental parameters, namely,

$$\nu_0=\sigma_n\langle v_n\rangle N_c , \tag{12}$$

where σ_n is the electron-capture cross section of the defect, $\langle v_n\rangle=(3kT/m)^{1/2}$ is the average thermal velocity of conduction electrons, m is the conduction-electron effective mass, and N_c is the weighted integral of the conduction-band density of states times a Boltzmann factor (this is usually called the effective density of states[49]). Since the terms in ν_0 may themselves be temperature dependent, one must exercise caution in interpreting ex-

perimental values of E and ν_0 derived from an Arrhenius plot. For this paper we have adopted the usual convention[30] of taking account of the T^2-temperature dependence of $\langle v_n \rangle N_c$ by subtracting $2kT_m$ from the measured value of E, where T_m is the average temperature of the Arrhenius plot for each peak in the spectrum. The temperature dependence of σ_n is a more serious problem and, without an independent measurement of the carrier capture process, cannot be unambiguously determined. Unless this temperature dependence is explicitly taken into account it will automatically be included in the measured activation energy E.[30] In this case the measured activation energy is the *sum* of both the potential barrier to capture and the equilibrium depth of the bound state, i.e., E is the energy difference between the ground state of the defect and the *highest* energy it must achieve to become free even if this maximum energy is greater than the final conduction path.

The *concentrations* of the various deep levels in the discrete defect case are in nearly all cases directly related to the respective magnitudes of the peaks in the DLTS spectrum. This is to be contrasted with optical emission or absorption spectra where quantitative relationships depend in detail on the magnitudes of various matrix elements which are nearly always vastly different for different defects. In the small-signal ($\Delta C/C << 1$), large bias ($V_R >> E_g/q$) limit, the relationship between the magnitude of the capacitance transient ΔC associated with a trap of concentration N_T is simply

$$\frac{\Delta C}{C} \cong 2\frac{N_T}{N_S} , \tag{13}$$

where C is the overall capacitance of the depletion layer and N_s is the net shallow-level concentration which determines the shape and size of the depletion layer.[30] It is immediately clear from the TSCAP data of Fig. 6 that we are not in the $\Delta C/C << 1$ limit for a-Si:H. Nevertheless, our detailed analysis in Ref. 35 shows that Eq. (13) is still useful in giving order-of-magnitude estimates of deep- versus shallow-level concentrations in our a-Si:H samples.

V. DLTS RESULTS FOR a-Si:H

The general concepts of DLTS were derived for the case of sharp levels, as discussed in the last section. However, even in the case of crystalline semiconductors there exist cases where the DLTS spectra are broad and relatively featureless as in a-Si:H. Two good examples are the broad defect bands introduced by ion implantation[50] or by plastic deformation.[51] A distribution of noninteracting defects with different energies will give rise to a spectrum of thermal emission rates. In the small signal case the transient response of the junction is the sum of a series of decaying exponentials with the rate constants determined by Eq. (11). The DLTS spectrum is thus simply an unresolved superposition of sharp DLTS lines with a weighting function which is proportional to the trap concentration distribution over the various possible energies, i.e., the density-of-state function $g(E)$. In other words, the DLTS spectrum is proportional in the small signal limit to $g(E)$ convoluted with the sharp DLTS line shape. Since $\Delta T \propto T$ for each line, this convolution is essentially the product of $g(E)$ times the temperature.

The activation energy of each component of a broad, unresolved DLTS signal cannot be measured directly by a set of Arrhenius plots as in the case of resolved lines. Thus, the energy scale can only be determined from the inverse of Eq. (11), namely

$$E = kT \ln(\nu_0 \tau_m) , \tag{14}$$

where τ_m is the DLTS rate window given by Eq. (10). If there is at least one distinct feature in the spectrum, the values of E and ν_0 may be independently determined for that feature by the usual Arrhenius plot. The extension of that ν_0 to the entire spectrum via Eq. (14) is not *a priori* justified. However, as we will discuss, it is possible to test the validity of a fixed ν_0 by making carrier-capture measurements.

The original extension of DLTS to the case of a-Si:H was done according to these simple concepts.[3–5] We found, however, that a detailed analysis of the data proved to much more difficult than in the case of crystals. The difficulties arose because of three major factors which did not affect most previous work with small concentrations of discrete defects in crystals. Namely, (1) the admittance of a-Si:H diodes was not well understood, (2) the deep-level concentration in a-Si:H is typically so large that is has a non-negligible effect on the shape of the band bending in the depletion region, i.e., we are not in the small signal ($\Delta C/C << 1$) limit, and (3) the DLTS spectra are broad with few well-defined distinguishable features. The theoretical work in the adjoining paper[35] deals with these problems in a way which now makes us as confi-

dent of our analysis of a-Si:H as we have been of sharp levels in crystals. Our analysis is rather complex and necessitates numerical calculations to obtain $g(E)$ from the DLTS spectra. Nevertheless, as we shall discuss in Sec. VI, the DLTS spectrum is extremely sensitive to changes in $g(E)$ and hence constitutes a powerful method for studying $g(E)$.

The two types of DLTS spectra which we routinely measure correspond to the two types of metastable initial conditions used for the TSCAP scans in Fig. 6, namely, (1) filling the electron trap states in the upper half of the gap by momentarily lowering the reverse bias (voltage pulse filling), or (2) saturating both electron and hole traps with 6000-Å laser excitation (laser pulse filling). Two cycles of the DLTS experimental sequences are shown in Fig. 8 for a DLTS rate window of 10 s^{-1} ($\tau_m = 100$ ms). A typical set of laser pulse and voltage pulse DLTS spectra which correspond to the respective sequences in Fig. 8 are shown in Fig. 9. The dashed line at 200 K is the 10-kHz dielectric response turn-on temperature for this sample. For temperatures below this point the spectra are not readily interpreted in terms of a density of states.

The DLTS spectrometer which was used is similar to ones which are commonly used for discrete levels in crystals.[30,45] The major difference for the application to a-Si:H is the introduction of a low-

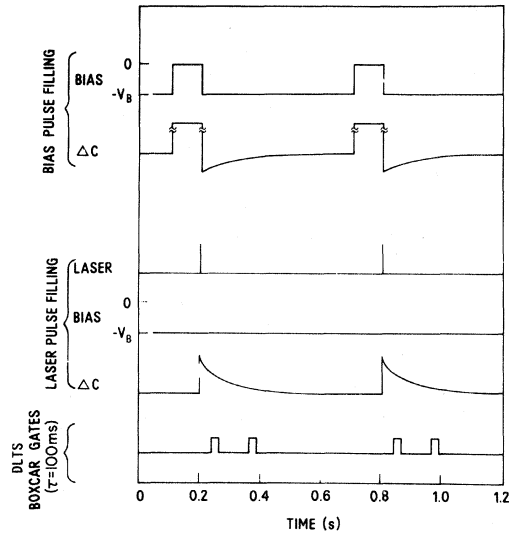

FIG. 8. Typical experimental timing sequence for the repetitive transients which are converted into DLTS spectra by the principle illustrated in Fig. 7. Unless noted, all DLTS spectra in this paper are recorded using the times shown in this figure, i.e., a 10-s^{-1} rate window.

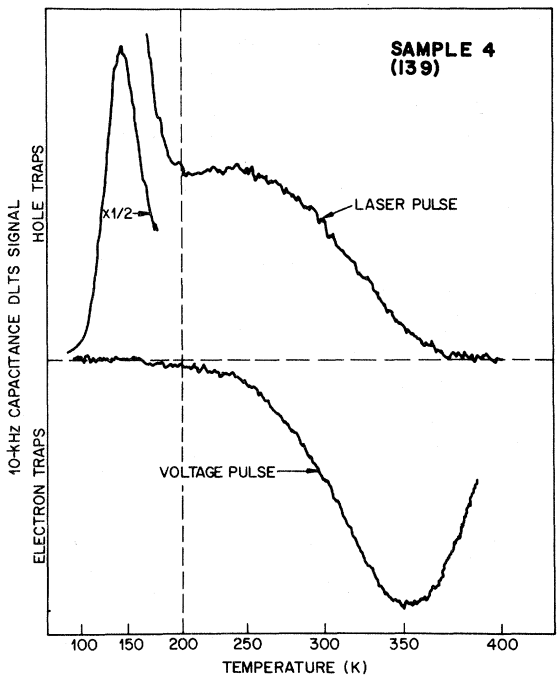

FIG. 9. DLTS hole- and electron-emission spectra for sample 4(139) doped with 60-vppm PH_3. (10-s^{-1} rate window, 5-V reverse bias on rear junction.)

frequency (1—100-kHz) capacitance meter (lock-in amplifier) in place of the 1—20-MHz capacitance meter typical of most DLTS systems. A low measurement frequency is necessary to make the dielectric response turn-on temperature low enough to observe $g(E)$ over an appreciable fraction of the gap. Unfortunately, the signal-to-noise (S/N) ratio is proportional to the frequency. Therefore, the typical value of 10 kHz was chosen as the best compromise between a wide DLTS spectrum and good S/N ratio. The choice of the DLTS rate window is a result of a similar trade-off between maximum spectral range (which favors a fast rate window) and the requirement that for reproducibility of the measurements the maximum temperature of the DLTS scan must be less than 100 °C (which favors a slow rate window). The typical value of $\tau_m = 100$ ms was chosen as the best compromise.

We will discuss the detailed analysis of DLTS spectra in Sec. VI. However, even the raw DLTS data in Fig. 9 may be roughly interpreted in terms of the gross features of $g(E)$ as follows. The negative voltage-pulse signal is due to electron emission and is roughly proportional to $g(E)$ measured *down* from the conduction-band edge to midgap with the energy scale proportional to the temperature scale according to Eq. (14). Thus the small

signal between 200 and 250 K reflects a corresponding minimum in $g(E)$ in the upper half of the gap. The apparent peak in the electron signal is due to the midgap cutoff imposed by the electron quasi-Fermi level. The positive laser-pulse spectrum is the *net* hole *minus* electron-emission signal. Therefore, the baseline for the hole states alone is not zero but is actually the voltage pulse signal. If this baseline is used instead of the true zero, then the laser-pulse signal may be roughly interpreted as $g(E)$ measured *up* from the valence-band edge to midgap with the same energy scale as for the electron-trap signal. Thus the shape of the laser pulse signal in Fig. 9 indicates a broad tail of hole states extending from the valence band to midgap.

As we mentioned earlier, a major question about DLTS as opposed to TSCAP is whether or not the saturated initial conditions have actually been achieved for each transient. The operational test for answering this question is to vary the length or intensity of the laser or voltage pulse until the DLTS signal no longer increases. The saturation of the voltage pulse signal is shown in Fig. 10 for a typical sample. This shows five successive DLTS scans with a different voltage pulse length τ_p for each scan. The times shown in the figure (1 μs to 10 ms) correspond to the duration of the respective zero-bias pulse before each transient.

Figure 10 can also be interpreted to directly give the capture rates for the trap-filling phase of the

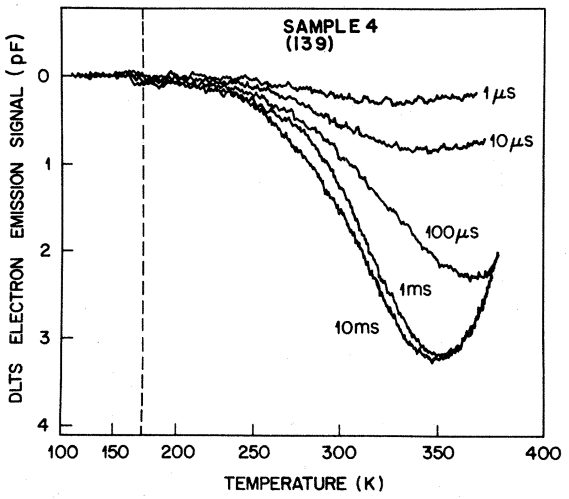

FIG. 10. Set of five DLTS electron-emission spectra for sample 4 (139) corresonding to different durations of the bias voltage pulse used to fill the gap states with electrons. (10-s^{-1} rate window, 5-V reverse bias on rear junction.)

experiment. Note that for most of the spectrum the characteristic time of electron capture is of the order of 100 μs. The states are completely filled by a 10-ms pulse, and thus the DLTS data can be interpreted as $g(E)$ for this pulse length. We can show the achievement of complete filling by two independent means. First, the apparent saturation at 10 ms is further supported by the fact that the signal produced by the 100-ms pulse is essentially identical to the 10-ms spectrum shown. By detailed balance we know that the thermal filling time constant for a given trap below the Fermi level in neutral material cannot be longer than the thermal emission time. Since the emission time constant (DLTS rate window) in Fig. 10 is 100 ms for *all* parts of the spectrum, we expect that a 100-ms zero-bias-filling pulse should be sufficient to fill the states, as is observed. A second way to verify that the DLTS initial condition corresponds to complete saturation is to compare the integral of the DLTS signal with the TSCAP signal (ΔC_e) in Fig. 6. Recall that for the TSCAP case the zero-bias-filling time may be made arbitrarily long (several hours, if necessary) so that there is absolutely no doubt that thermal equilibrium, and hence complete filling of states up to the zero-bias Fermi level, has beeen achieved. We find that the DLTS integral is in fact somewhat *larger* than the ΔC_e seen by TSCAP. This discrepancy is generally less than a factor of 2 and can often be explained by the high-temperature area-spreading effect discussed in Sec. III.

The saturation of the laser-induced DLTS spectrum is shown for a typical sample in Fig. 11. The tendency for the DLTS signal to saturate with a single pulse as the laser *intensity* is increased may be compared with the complementary saturation experiment possible with TSCAP. Namely, the saturation of ΔC_{e+h} in Fig. 6 as a function of the *number* of laser pulses of the *same* intensity. The TSCAP results typically correspond to of the order of 100 laser pulses. As with voltage pulses, the DLTS-TSCAP agreement is within a factor of 2 and indicates no measurable nonlinearities.

The flash-lamp-pumped dye laser used in our experiments produced 6000-Å pulses of approximately 1 μs duration with a typical repetition rate of 15 Hz for setting the TSCAP initial conditon and a sequence of triggered single pulses as shown in Fig. 8 for the DLTS spectra. The arbitrary intensity scale of 100 in Fig. 11 corresponds to approximately 5×10^{16} photons/cm^2 for each pulse, or 5×10^{22} photons/cm^2 s during the pulse. For a typical *a*-

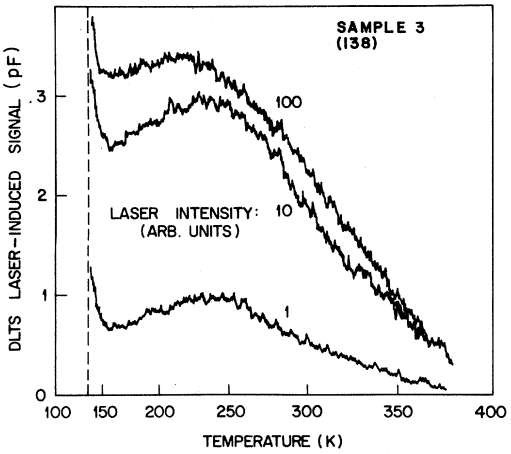

FIG. 11. Set of three DLTS hole-emission spectra for sample 3 (138) corresponding to different intensities of the laser pulse used to fill the gap states with holes and electrons. (10-s^{-1} rate window, 5-V reverse bias on rear junction.)

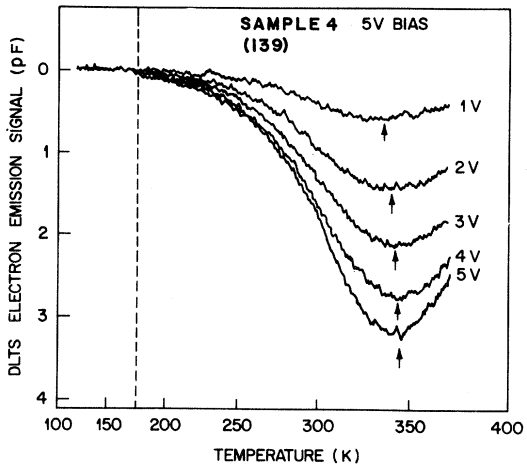

FIG. 12. Set of five DLTS electron-emission spectra for sample 4 (139) corresponding to different amplitudes of the 10-ms bias voltage pulse used to fill the gap states with electrons. The arrows indicate the DLTS peak positions calculated for pulse amplitudes less than 5-V using a spatially uniform $g(E)$ fit to the 5-V pulse data. (10-s^{-1} rate window, 5-V reverse bias on rear junction.)

Si:H absorption coefficient of $\alpha \sim 10^4$ cm^{-1} at 6000 Å,[52] this photon density would correspond to a total number of *e-h* pairs of 5×10^{20} cm^{-3} per pulse. The actual number of *e-h* pairs is 1—2 orders of magnitude less than this due to the 1—10% transmission of the thin Schottky-barrier metal film on the surface. The instantaneous electron and hole concentrations depend on the respective carrier lifetimes.

The *spatial variation* of the deep levels giving rise to the DLTS signal may be measured *directly* by using voltage pulses of various amplitudes.[30,45] Typical data of this sort are shown in Figs. 12 and 13. Because of the discrepancies among the various $g(E)$ models in the literature and the fact that interface effects have not been properly separated from bulk effects in many cases, it is crucially important to show directly that our DLTS spectra relate to the true *bulk g(E)*. As discussed earlier, we know that capacitance transient measurements are considerably less sensitive to interface states than are current transients or steady-state admittance measurements. However, the data in Figs. 12 and 13 allow us to demonstrate directly that our DLTS spectra are associated primarily with bulk states.

The DLTS spectra in Fig. 12 correspond to five scans with voltage pulses of 1 to 5 V amplitude, respectively. An *increased* amplitude of the pulses corresponds to a *smaller* reverse bias during the electron-capture phase, and hence a large filled proportion of the −5-V depletion region during the emission phase. In this notation a 5-V filling pulse amplitude for a −5-V steady-state bias may also be referred to as a zero-bias-filling pulse, since the diode is at zero bias during the pulse.

The amplitude of the maximum DLTS signal is plotted as a function of the filling-pulse amplitude in Fig. 13 for two values of the filling-pulse duration, corresponding to $\frac{1}{10}$ and $\frac{1}{2}$ of the rate window time constant τ_m. The longer of the two

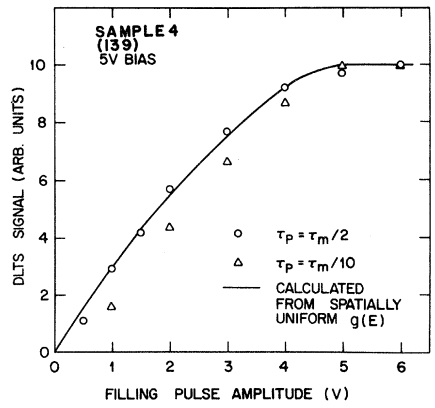

FIG. 13. Magnitude of DLTS electron-emission peak vs bias-pulse amplitude from data similar to that in Fig. 12. The two sets of data correspond to voltage-pulse durations of $\frac{1}{2}$ and $\frac{1}{10}$ of the DLTS emission time constant, respectively. Note the good agreement between theory and experiment for the longer-pulse data.

pulse durations corresponds to complete filling of the states up to the quasi-Fermi-level during each filling pulse. The solid curve in Fig. 13 is calculated using the methods of Ref. 35 for the complete filling case by assuming that the DLTS-determined $g(E)$ for this sample is spatially uniform in the bulk. Furthermore, the calculation correctly reproduces the small shift in the DLTS peak position with pulse amplitude shown by the arrows in Fig. 12. The ability of our numerical DLTS simulation program to accurately reproduce such fine details of the intermediate-pulse voltage data in Figs. 12 and 13 *with no adjustable parameters* is noteworthy when one considers that the $g(E)$ for this sample was determined only by the saturated zero-bias pulse DLTS and TSCAP data according to the method which we will discuss in the next section. We consider this sort of agreement to be very strong confirmation of our method of analysis and also of the bulk nature of the DLTS spectra. We therefore assert that the states seen by DLTS correspond to complete saturation of the true bulk $g(E)$.

Let us now discuss the measurements needed to characterize the energy scale of the DLTS spectra. Earlier we pointed out the difficulty in performing the straightforward sharp-line Arrhenius plot analysis for the case of broad DLTS spectra. The only feature of the *a*-Si:H DLTS spectra which can be readily identified is the peak in the electron emission (voltage pulse) spectrum at about 350 K for $\tau_m = 100$ ms. This peak does not necessarily correspond to an isolated peak in $g(E)$ as it might at first appear. In fact, our analysis shows that the sharp drop in electron signal above 350 K in Figs. 9, 10, and 12 is due to the quasi-Fermi-level cutoff in $g(E)$ imposed by the final-state occupation statistics discussed in connection with Fig. 5. This cut-off occurs because states below a point roughly at midgap [see Eqs. (8) and (9)] are *always* filled with electrons in steady state for all reverse bias voltages, and thus can never give rise to a metastable emission transient with only voltage pulse filling.

However, since the quasi-Fermi-level cutoff corresponds to a well-defined energy which depends only weakly on temperature, we may use this feature of the DLTS spectrum to obtain an activation energy. We measure the shift of the DLTS spectrum with rate window in the standard manner and use the positions of this broad electron "quasi"-peak to construct an Arrhenius plot. Such a plot is shown in Fig. 14 for a typical sample.

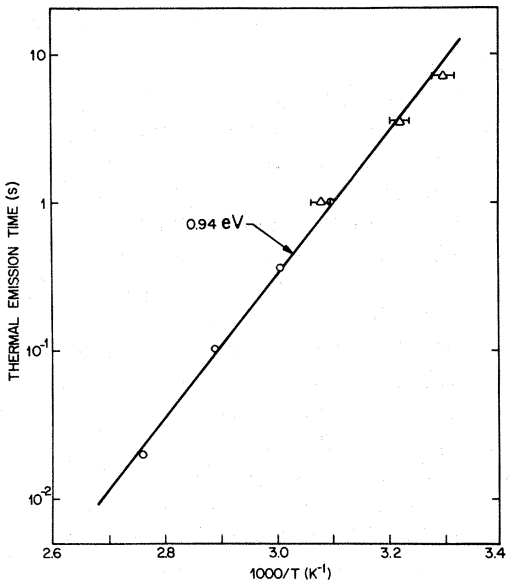

FIG. 14. DLTS emission time constant τ_m (inverse of rate window) vs the inverse temperature of the peak in the electron-emission spectrum at that τ_m for a typical sample. The activation energy of 0.94 eV is the slope of the raw data without the $2kT$ correction (see text). The circles are from DLTS spectra while the triangles are calculated from directly recorded capacitance transients.

The 0.94-eV slope of the raw data in Fig. 14 must be corrected for the weak temperature dependence of the prefactor in Eqs. (11) and (12). According to standard procedure,[30,45] this corresponds to subtracting $2kT = 0.06$ eV from the measured slope to obtain the corrected activation energy for this sample of $E = 0.88$ eV $\pm 10\%$. The conservative $\pm 10\%$ limits of error reflect both the scatter in the data as well as the uncertain validity of the $2kT$ correction procedure. The best value for the average corrected activation energy of several samples is $E = 0.85$ eV $\pm 10\%$. This implies a prefactor value of $\nu_0 \sim 10^{13}$ s^{-1} and sets the energy scale for our standard 100-ms rate window at $E = 28kT$ ($\pm 10\%$) [see Eqs. (1)–(14)].

The validity of extending this constant value of ν_0 to all parts of the spectrum can be tested in two ways. First, we can use the electron-capture data of Fig. 10 to estimate the variation in the capture rate over the electron-trap spectrum. The time constant of electron capture τ_e is related to the emission parameters of Eqs. (11) and (12) by the definition

$$\tau_e^{-1} = \sigma_n \langle v_n \rangle n \ , \tag{15}$$

if the free-electron concentration n is given by

$$n = N_c \exp[-(E_c - E_F)/kT] \,. \quad (16)$$

Thus the temperature dependence of τ_e involves *both* the temperature dependence of σ_n and n. If we follow the usual convention[2] and take the activation energy of the conductivity to be $E_c - E_F$, then for the sample shown in Fig. 10 we would expect n to have an activation energy of 0.29 eV. In fact, this is almost exactly what is observed for the capture rate. The capture time constant in Fig. 10 varies from 183 μs at 322 K to 46 μs at 368 K; this corresponds to an activation energy of 0.3 eV. Other samples with $E_c - E_F$ in the range of 0.2 eV show activated τ_e behavior close to 0.2 eV. These results are consistent with the interpretation that σ_n is relatively independent of temperature (from 320 to 370 K) *and* energy (from $E_c - 0.78$ eV to $E_c - 0.89$ eV) with the observed variation in τ_e being due to the temperature dependence of n.

Unfortunately, the exponential prefactors in the detailed balance relationships in Eqs. (11), (12) and (15), (16) are not internally consistent. The apparent capture time of ~ 100 μs is several orders of magnitude too slow to be consisent with both ν_0 and a Fermi-level position inferred from the activation energy of the conductivity. The origin of this discrepancy is not clear. There are two points which should be noted, however. First, detailed balance is usually accepted for the crystalline case[30,45,49] but it is by no means obvious for *a*-Si:H that detailed balance between capture and emission applies when emission occurs in an electric field of $\sim 10^5$ V/cm and capture is measured at zero field. Thus, for example, different conduction paths might be dominant in these two electric-field limits. Second, bulk transport limitations are *not* likely to be responsible for the slow trap-filling rate since a drift mobility of 10^{-5} cm^2/V s at 350 K would be we required to explain a 100-μs time constant for the flow of electrons from the reverse biased configuration to the narrower depletion-region characteristic of zero bias. Furthermore, an extrapolation of the *RC* time constant associated with the 10-kHz turn-on temperature for the sample in Fig. 10 implies $RC \lesssim 16$ ns at 350 K. It is entirely possible, however, that the relatively large current surge (~ 10 A/cm^2 for 1 ns) needed to fill the traps in a time predicted by detailed balance is too large for the type of non-ohmic contacts on our samples.

A second means of justifying the approximate energy and temperature independence of ν_0 is to consider the extent to which our theoretical model fits the changes in the DLTS spectra with rate window. The most straightforward analysis assumes that ν_0 is independent of temperature and energy, i.e., that the energy scale of $g(E)$ may be given by Eq. (14). Since ν_0 enters only logarithmically into the energy scale, small deviations from this assumption will be relatively unimportant. To the extent that the variations in ν_0 are large enough to significantly distort the $g(E)$ obtained by our fitting procedure, we should expect to see significant deviations in the fit to the DLTS spectra for different rate windows. A comparison between the calculated DLTS spectra and the actual data for two different rate windows is shown for a typical sample in Fig. 15. The $g(E)$ for this sample was obtained by fitting the $\tau_m = 100$ ms DLTS data with $\nu_0 = 2 \times 10^8$ T^2 s^{-1} obtained from data similar to that in Fig. 14. The fact that the DLTS spectra for $\tau_m = 20$ ms can also be fitted quite well with the *same* ν_0 *and no adjustable parameters* is strong evidence that the variations in ν_0 with trap depth are small enough to be well within our estimated 10% absolute accuracy for the energy scale of $g(E)$.

Thus we believe that the energy scale obtained from DLTS measurements is trustworthy within the $\pm 10\%$ limits of error, in spite of the fact that such energies are more difficult to justify than for the discrete-level case. Concerns about the *absolute*

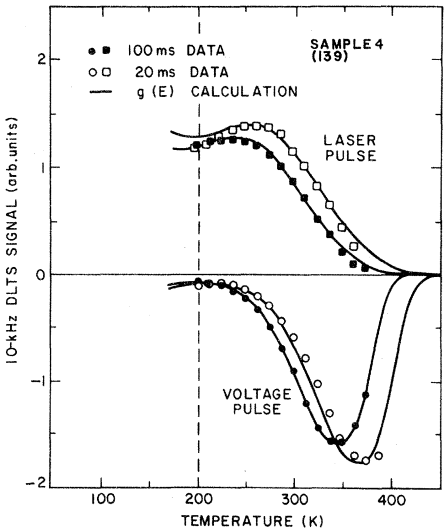

FIG. 15. DLTS hole- and electron-emission spectra for two different rate windows for sample 4 (139). The solid lines are calculated from the density of states in Fig. 23, which is derived solely from the 100-ms. DLTS data. (10- and 50-s^{-1} rate windows, 5-V reverse bias on rear junction.)

energy scale are, furthermore, somewhat secondary to the fact that the *relative* energy positions of certain specific features of the density of states are quite reproducible from sample to sample. Thus the possibility exists to unravel the various components of $g(E)$ in spite of the possible small systematic distortions of the energy scale inherent in the simple assumption of a constant ν_0. In the next section we shall show how variations in $g(E)$ in different samples can easily be resolved by our analysis.

VI. DENSITY OF GAP STATES

A. DLTS fitting procedure

As discussed in Sec. V, the general characteristics of $g(E)$ are apparent even from the raw DLTS data. This was the spirit of the analysis used to obtain $g(E)$ in our earlier work.[5] However, because our previous results differed substantially from the widely accepted $g_{SL}(E)$ obtained by field-effect measurements,[1,2] it was felt that a more rigorous analysis was necessary to justify the relationship between $g(E)$ and the DLTS spectra. The theoretical groundwork for this analysis is given in the following paper.[35] In this section we shall discuss in detail the procedure for obtaining $g(E)$ from the dynamic response data presented in this paper.

In the context of our model[35] for the dynamic response of a Schottky barrier with an arbitrary distribution of gap states, the relationship between the data and $g(E)$ is unique and rigorous. Thus the validity of $g(E)$ rests on the validity of the underlying assumptions of the model and not on the details of the analysis. These assumptions are discussed in Ref. 35. The main points relevant to the case of *a*-Si:H are the following: (1) the dynamic response above the dielectric turn-on temperature is considered to be limited by thermal emission from gap states and not by carrier transport, (2) the electric properties (E_g,ρ,E_F) of the material are considered to be laterally uniform on a scale greater than the Debye screening length $l_D=(kT\epsilon/q^2n)^{1/2}$, (3) the exponential prefactor ν_0 in Eqs. (11), (12), and (14) is taken as a constant times T^2, (4) the final-state quasi-Fermi-level in regions I—III of Fig. 5 is taken as $E_g/2$, (5) the initial state for voltage-pulse filling is considered to be in equilibrium with the bulk Fermi level during the filling pulse, and (6) the initial state for laser-pulse filling is assumed to be uniformly at half-occupancy for gap states in the range of observable DLTS energies.

We conservatively class these six conditions as assumptions since we feel they have not been rigorously verified. However, most of them seem to be nearly true on the basis of experiments which were designed to test their validity. In particular, we feel that conditions (3)—(5) are close enough to being true that any slight deviations would be within our stated limits of error. Condition (6) is difficult to verify. We know from the data such as in Fig. 11 that the laser-induced signal is *saturated* but we do not know *a priori* the occupation fraction. Thus for states in the lower half of the gap which capture electrons much more readily than they capture holes (i.e., $\sigma_n >> \sigma_p$), the concentrations would be grossly underestimated by this fitting procedure. Fortunately, however, this situation is not expected since most deep levels in crystals have the largest capture cross section associated with carriers in the nearest band so that we might actually expect $\sigma_p >> \sigma_n$ for states below midgap. In this case condition (6) would overestimate the trap concentration by a factor of 2. Note that such kinds of ambiguities do not exist for the voltage-pulse DLTS spectrum [condition (3)].

Assumptions (1) and (2) require more discussion since some results in the literature might seem to contradict them. In particular, Viktorovitch and Moddel[25] argue that the dynamic response of an *a*-Si:H Schottky barrier is controlled by carrier transport and not by capture and emission at states in the gap. Their studies of the frequency dependence of the barrier admittance of undoped *a*-Si:H films were made in the vicinity of the dielectric turn-on of the film. We would agree in this region of temperature and measurement frequency that transport effects limit the response (see Sec. III). However, our assumption (1) applies to temperatures well above the dielectric turn-on temperature where the discussion of Viktorovitch and Moddel does not apply.

We can give a convincing argument for the validity of condition (1) as follows. For DLTS this condition is valid if the thermal emission time constant [Eq. (11)] is much longer than the time required for the emitted charge to be swept out of the depletion region. Because of the high electric field in the depletion region ($\sim 10^5$ V/cm) the carrier sweep-out time is on the order of 1 ns for a mobility of 1 $cm^2/V\,s$. This is to be compared with a typical thermal emission time of 20—300 ms for DLTS experiments. Thus the drift mobility would have to be less than 10^{-7} $cm^2/V\,s$ for condition (1) to be violated in our DLTS measurement.

Furthermore, thermal emission from a level seen by DLTS must be the rate-limiting process because the only empty states available to retrap the emitted carrier as it leaves the depletion region are shallower than the emitting state. This follows because the deeper states are still full and because the reemission time from shallower states depends exponentially on the depth of the state. If retrapping *were* a serious problem it would manifest itself as a severe distortion of the DLTS spectrum for partially filled initial conditions. Such distortions are not seen (cf. Figs. 10 and 11). Furthermore, to extend condition (1) to *admittance measurements* we need only show that for small reductions in bias, the time necessary to fill $g(E)$ up to the Fermi level is less than or equal to the time necessary for these states to empy when the bias is increased. The verification of this behavior is implicit in Figs. 12 and 13. There we calculate the filling-pulse dependence of the DLTS signal *assuming* condition (1) to be valid. The resulting excellent fit to the experimental data clearly provides strong justification that carrier transport does not restrict the filling of gap states on time scales comparable to the thermal emission times.

Condition (2) is necessary to keep our analysis from becoming too cumbersome. However, there is now a large body of evidence from structural studies[39,53,54] that *a*-Si:H films are not laterally uniform but exhibit columnar growth morphology with hydrogen-rich polysilane material in the intercolumnar regions.[54] Whether such morphology will strongly affect the transport properties of *a*-Si:H films is an interesting question which has not been answered at the present time. Indeed, the extent to which our results may or may not be consistently analyzed with assumption (2) may shed some light on this problem.

In order to best illustrate the use of the numerical methods of Ref. 35 in obtaining $g(E)$ from the DLTS spectra let us consider the various steps of the fitting procedure for a typical sample [5(152)]. The saturated DLTS spectra for sample 5(152) are shown in Fig. 16 for voltage-pulse and laser-pulse filling. Recall that on the basis of the discrete level case, we expect in the small-signal limit ($\Delta C << C$) that the DLTS spectra are directly proportional to the product of $g(E)$ times T. This is the sort of simple analysis used in Ref. 5 to obtain an approximate $g(E)$. A detailed analysis shows, however, that the nonlinearities inherent in the $\Delta C \simeq C$ case tend to cancel the effect of the linear increase in DLTS resolution with temperature.

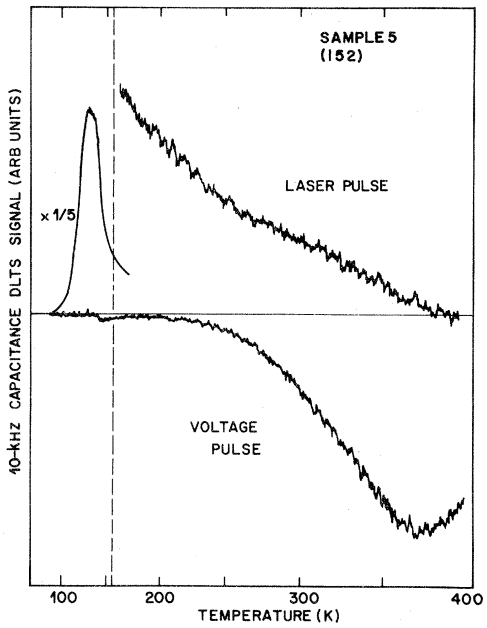

FIG. 16. DLTS hole- and electron-emission spectra for sample 5(152) doped with 300-ppm PH_3. (10-s^{-1} rate window, 5-V reverse bias on rear junction.)

Therefore, it turns out that the best initial guess for $g(E)$ is the actual DLTS spectra and not the spectra divided by temperature as was done in Ref. 5.

The first trial $g(E)$ for sample 5(152) is shown in Fig. 17. The shape of $g(E)$ in the upper half of the gap in Fig. 17 is proportional to the DLTS electron-emission signal. This electron emission signal also defines the thermal emission band gap when the high-temperature cutoff due to E_F^* is taken as $E_g/2$ [see Eqs. (8) and (9)]. In the lower half of the gap the first trial $g(E)$ is the *sum* of twice the laser-induced DLTS net hole emission signal *plus* the electron emission signal as indicated in Fig. 17. The factor of 2 follows from the assumed initial half-occupancy of gap states following laser excitation. The shape of the conduction-band tail in the vicinity of E_F is not given directly by the DLTS spectra. Rather it is adjusted to give the correct value for the steady-state capacitance at temperatures just above the dielectric turn-on (see Fig. 2). We discuss further how band-tail behavior may be inferred from our results in Sec. VI C.

The approximate $g(E)$ in Fig. 17 is used as the first trial function in the numerical DLTS simulation program discussed in Ref. 35. The DLTS spectra calculated from Fig. 17 are shown in Fig.

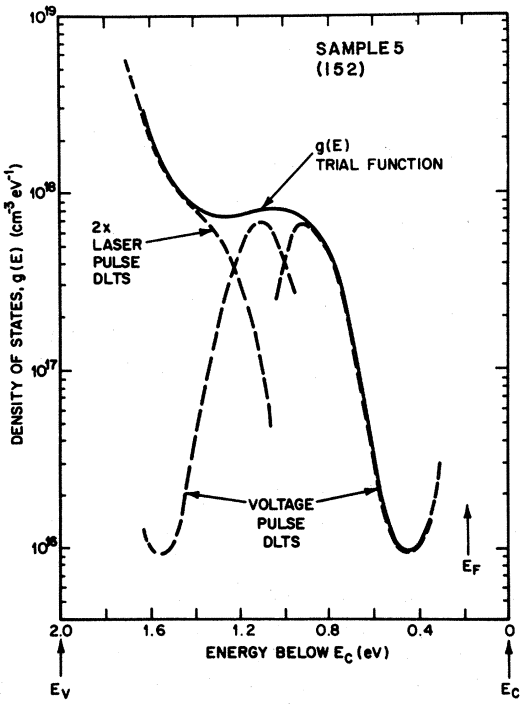

FIG. 17. First $g(E)$ trial function to be used in the iterative fitting procedure which ultimately gives the bulk density of states for sample 5 (152). The dashed lines indicate how the raw (DLTS) data in Fig. 16 are used to construct the first trial function (solid line).

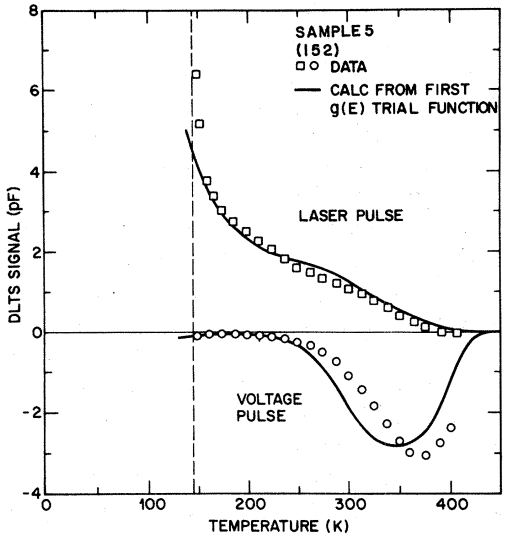

FIG. 18. DLTS spectra for sample 5 (152) calculated from the first $g(E)$ trial function in Fig. 17. The DLTS data points are taken from the spectra in Fig. 16.

18. Note that while the agreement is not perfect at this stage, it is clearly close enough to indicate that a $g(E)$ with the general shape of Fig. 17 is a good starting point for the subsequent iterations. We adjust the shape, magnitude, and band gap of $g(E)$ until the calculated DLTS spectra agree acceptably well with the measured DLTS spectra and with the TSCAP ΔC_e and ΔC_{e+h} trap integral results discussed in Sec. IV B. More specifically, each iteration is performed in the following way. First, the conduction-band tail is adjusted so that the integral of conduction-band-tail states between $E_c - 0.5$ eV and $E_c - E_F$ gives the correct numerically calculated capacitance for the TSCAP zero-bias-fill curve (as shown in Fig. 6). This integral is given as N_s in Eq. (19) and is analogous to the net shallow donor concentration in the crystalline case. While this does not uniquely determine the band tail, it does allow the characteristic energy E_0 to be determined if one *assumes* an exponential band tail such as given in Eq. (25). The second step in each iteration is to adjust the vertical scale factor for the *deep* electron states ($E_c - E_{g/2} < E < E_c - 0.5$ eV) in $g(E)$ so that the numerical capacitance calculation gives the correct steady-state or "baseline" capacitance in the TSCAP measurement. The final TSCAP value for the "laser-fill" C-T curve just above turn-on is fitted by adjusting the scale factor of the deep hole traps ($E_v + 0.5 < E < E_v + E_{g/2}$). These capacitance values depend on *both* the shallow and deep parts of $g(E)$ and, as with all capacitance values, depend on the measurement frequency, bias voltage, and temperature. Such effects are included in our numerical analysis and are discussed at length in the second paper.[35] At this point in the interation procedure the vertical scale of $g(E)$ has been adjusted so that the integrals over $g(E)$ in Eqs. (17)–(19) fit the TSCAP data. As we have discussed, this is the most reliable way to determine the concentration scale of $g(E)$ since complete saturation of the traps is easy to achieve in the TSCAP measurement.

It is now possible to calculate the DLTS spectra from the properly scaled $g(E)$ trial function using the numerical procedure discussed in the following paper.[35] In that paper we discuss at some length the general relationship between the DLTS spectra and the density of gap states for *any* semiconductor, whether crystalline or amorphous. In that discussion we show that DLTS is a *direct* measurement of $g(E)$, provided the nonlinearities due to the dynamic response of the diode are properly taken into account. These nonlinearities are a trivial result of Poisson's equation and are most pronounced for the case where the deep states have a larger concentration than the shallow states, as in *a*-Si:H. Thus the numerical diode response solu-

tion discussed in the second paper, which is important in its own right, allows us to remove the distortions of the DLTS spectra and obtain $g(E)$ directly. Unfortunately, as we are now demonstrating, the procedure must be carried out in the *reverse* order. Namely, the distortions due to the diode dynamics are folded into the numerical calculation of the DLTS spectra from a given $g(E)$. Thus when the properly scaled $g(E)$ trial function produces numerically calculated DLTS spectra which agree with the actual DLTS data, we can consider that $g(E)$ to be the *true* density of states. If not, we must make another iteration, i.e., adjust the shape of $g(E)$, rescale to the TSCAP data, and calculate another set of DLTS spectra. We should emphasize here that the numerical fitting procedure is not in the nature of a fit to a *theory* for $g(E)$ where one should expect detailed justifications for the changes made in $g(E)$ during the various iteration steps. Rather, we are actually iterating the $g(E)$-DLTS relationship to *self-consistency* by removing the inherent nonlinearities and distortions in what is actually a *direct* measurement of $g(E)$. Thus the details of how the final $g(E)$ is obtained are less important than the fact that self-consistency can be demonstrated. The only physical assumptions which affect the validity of this approach are clearly spelled out at the beginning of this section.

The *final* $g(E)$ for sample 5 (152) is shown in Fig. 19. The corresponding calculated DLTS spectra are shown in Fig. 20. In this figure the calculated spectra have been normalized to the measured DLTS spectra at the 370-K electron-emission peak. As discussed in Sec. IV B, this is necessary because the trap concentrations obtained from TSCAP measurements are considered more reliable than the DLTS signal magnitude. The measured DLTS signal for sample 5 (152) is about a factor of 2 larger than the result calculated from the $g(E)$ in Fig. 19. The DLTS shape, on the other hand, is well represented by this density of states. Additional support for the final $g(E)$ may be gained from a comparison of the measured and calculated steady-state properties of the barrier. These will be discussed in Sec. IV B.

In spite of the good fit displayed in Fig. 20 the detail shown in the shape of $g(E)$ would be misleading if the fit were insensitive to changes in $g(E)$. Indeed, the insensitivity to $g(E)$ for an iterative numerical fitting procedure is a serious problem in the case of a-Si:H field-effect data.[7,8] As we discuss in Ref. 35, our analysis of the shape

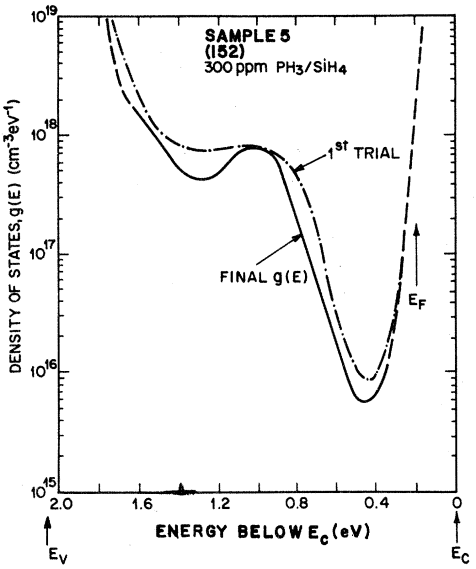

FIG. 19. Final density of states for sample 5 (152). Note that in all $g(E)$ plots the solid line is the energy range corresponding to the DLTS spectra, while the (dashed) conduction-band tail is adjusted to be consistent with both the DLTS spectra and the steady-state capacitance as explained in the text. For comparison, the first trial function in Fig. 17 is shown as the dash-dot line.

of normalized C-V data is equally insensitive to rather substantial changes in the shape of $g(E)$. Fortunately, the DLTS spectra are quite sensitive to the shape of $g(E)$ as is shown convincingly in Figs. 21 and 22. These figures show two of the

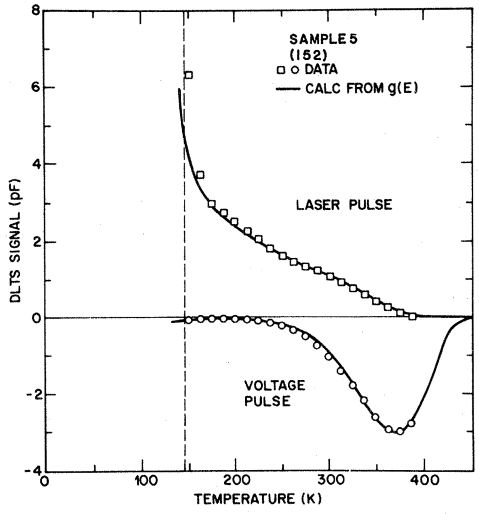

FIG. 20. DLTS spectra for sample 5 (152) calculated from the final $g(E)$ in Fig. 19. The DLTS data points are taken from the spectra in Fig. 16.

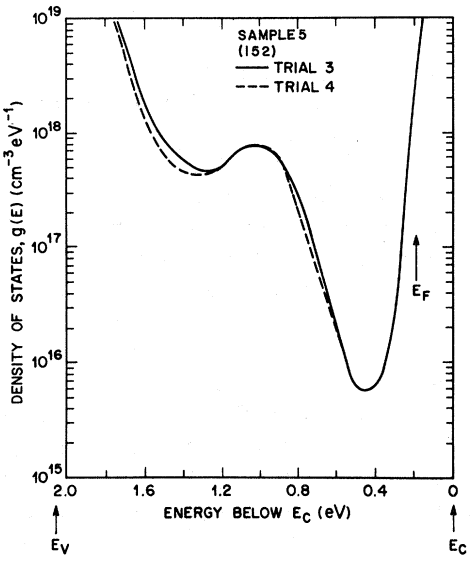

FIG. 21. Two $g(E)$ curves used in the iterative fitting procedure for sample 5 (152).

last stages of iteration for sample 5 (152). Note that while the calculated and measured DLTS spectra in Fig. 22 are clearly different, the two $g(E)$ curves in Fig. 21 differ only slightly from each other and from the best fit in Fig. 19. We also know from our analysis in Ref. 35 of various arbitrary densities of states that the general shape of the DLTS spectra in Fig. 16 can *only* be fitted by a $g(E)$ of the general shape shown in Figs. 17, 19, and 21. The fitting procedure thus leads to a *unique* $g(E)$.

Figure 19 should not, however, be taken as characteristic of all of our PH_3-doped a-Si:H samples. Indeed, we see variations in $g(E)$ from sample to sample which are well beyond our estimated limits of error. This clearly indicates that many of the gap states are *extrinsic* in nature and depend in as yet unknown ways of various impurities and/or sample preparation conditions.

The extrinsic nature of the gap states is best illustrated by showing $g(E)$ for two other samples which have significantly different DLTS spectra from sample 5 (152). The DLTS spectra for sample 4 (139) were used extensively to illustrate the discussion of our DLTS measurements in Sec. V (cf. Figs. 9, 10, 12, and 15). The best-fit $g(E)$ for sample 4 (139) is shown in Fig. 23. The calculated DLTS spectra for this $g(E)$ are compared with the data for two rate windows in Fig. 15. As a second example we show the density of states and DLTS spectra for sample 3 (138) in Figs. 24 and 25. In all of these cases the calculated DLTS spectra are somewhat smaller in magnitude than the measured spectra and have been normalized to the τ_{max} = 100-ms electron-emission peak. As we mentioned above, this discrepancy, which is typically of the order of a factor of 2, is due to the fact that we believe it is more reliable to fix the magnitude of $g(E)$ by the concentration integrals obtained from TSCAP measurements. While this may indi-

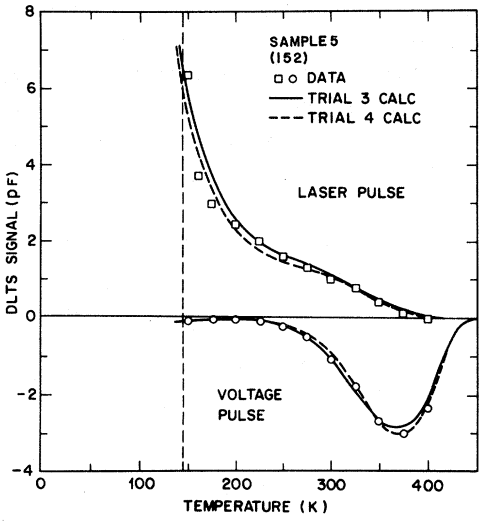

FIG. 22. Calculated DLTS spectra for the two $g(E)$ curves in Fig. 21. The DLTS data points are taken from the spectra in Fig. 16. The solid and dashed curves in this figure correspond to the respective solid and dashed curves in Fig. 21.

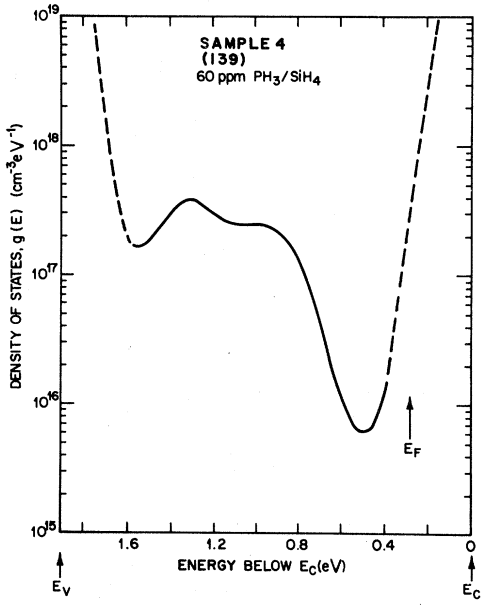

FIG. 23. Final density of states for sample 4 (139).

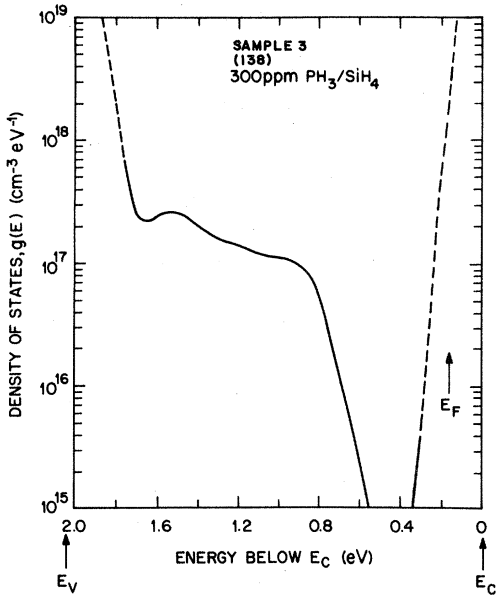

FIG. 24. Final density of states for sample 3 (138).

cate that our *absolute* concentration scale is only accurate to within a factor of 2, the *relative* scale from sample to sample is considerably more accurate. Therefore the differences in the three $g(E)$ curves in Figs. 19, 23, and 24 are significant and reflect real changes in the state density in these three samples. Some of the changes in $g(E)$ may be related to PH_3 doping and will be discussed in Sec. VI C.

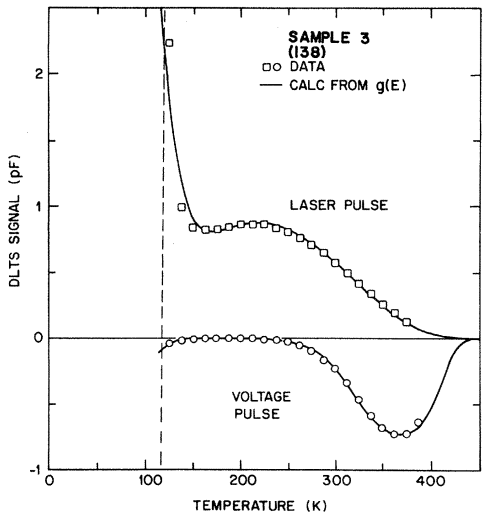

FIG. 25. DLTS spectra for sample 3 (138) calculated from the final $g(E)$ in Fig. 24.

B. Prediction of steady-state properties

Obtaining $g(E)$ by fitting the DLTS spectra is the most sensitive means for determining the density of states. However, this procedure has the drawbacks that it requires doped films and rather complex DLTS instrumentation. For this reason it is desirable to examine some of the more simply measured parameters. By our methods of analysis[35] we may calculate either the capacitance or current DLTS spectra as well as the temperature, frequency, and bias dependence of the diode admittance from a given $g(E)$. While it is possible in principle to obtain $g(E)$ from E_F^* to E_F from steady-state admittance measurements alone, as we discuss for undoped films, we will discuss such data primarily as a check on our DLTS fitting procedure for the doped samples. This provides us with added confidence in our results and also makes it possible for others to compare their doped samples with ours by simple C-T measurements.

In Figs. 26 and 27 we show the 10-kHz C-T curves as well as the room-temperature 100-kHz C-V data for sample 5 (152). The theoretical curves in these figures are calculated from the $g(E)$ in Fig. 19 with no adjustable parameters. We know that the steady-state capacitance is very sensitive to interface states; therefore, we do not expect perfect agreement with the data from a calculation based on the bulk $g(E)$. In general, however, the overall agreement is surprisingly good. The lower turn-on temperature of the calculated capacitance is simply due to the neglect of transport effects in the theory. One should only compare theory and experiment *above* turn-on in the regime where capture and emission processes dominate the dynamic response of the junction. Any discrepan-

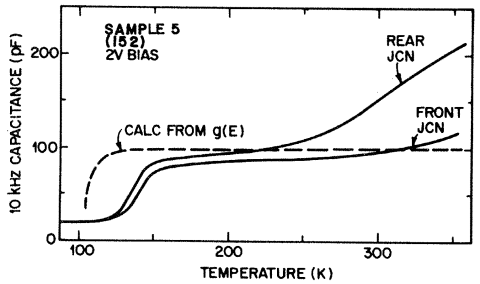

FIG. 26. Steady-state 10-kHz capacitance for the front and rear junctions of sample 5 (152) (solid lines). The dashed line is calculated for the same experimental conditions using the bulk density of states in Fig. 19 with no adjustable parameters.

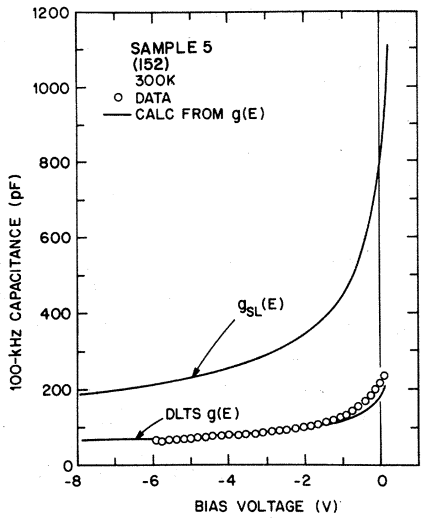

FIG. 27. Steady-state 100-kHz capacitance vs bias voltage at 300 K for the rear junction of sample 5 (152). The curved labeled DLTS $g(E)$ has beeen calculated for the same experimental conditions using the bulk density of states in Fig. 19 with no adjustable parameters. The curves labeled $g_{SL}(E)$ has been calculated for the same junction area, temperature, frequency, and Fermi level using the Spear-LeComber field-effect density of states as shown in Fig. 24(a) of Ref. 35 and in Refs. 1 and 2.

cies between theory and experiment in this regime generally decrease for large reverse bias voltages. This is the behavior characteristic of interface states. In fact, preliminary results on MOS structures show unmistakable evidence for interface states superimposed on the behavior associated with the bulk $g(E)$.

The major discrepancy between our typical shape of $g(E)$ and most other densities of states reported in the literature[1] is the deep minimum in the upper half of the gap. As we demonstrate in the theory paper,[35] this minimum in $g(E)$ corresponds to the characteristic plateau in the C-T curve above the dielectric turn-on temperature for conducting films such as shown in Figs. 2 and 26. Note, however, that while a plateau in C-T is sufficient to prove the existence of a minimum in $g(E)$ in the upper part of the gap, it may be obscured by extraneous effects. This is evident in Fig. 26 where in spite of the fact that both sides of the sample show approximately the same DLTS $g(E)$, only the front junction gives clear evidence of a well-defined plateau in the C-T curve. The origin of the extra increase in the rear junction C-T curve in this case is not clear, but it is most likely related to area-spreading effects and states near the interface as discussed in Sec. III.

The 100-kHz—300-K C-V curve in Fig. 27 fits very well with that calculated from the DLTS $g(E)$ for the same temperature and frequency. The good fit to the shape is not surprising since C-V curves are especially insensitive to the shape of $g(E)$ but are related primarily to various integrals of $g(E)$. Nevertheless, this check of our results is very important because transient measurements such as DLTS may fail to detect gap states if the density is large enough to allow direct hopping conduction out of the depletion layer to take place more rapidly than thermal activation to the band-tail conduction path. Such hopping states would, however, still have a significant effect on the C-V data since the integral of *all* gap states between E_F^* and E_F enters into slow-voltage-scan C-V measurements regardless of the respective conduction mechanisms for the states. Thus the C-V measurement provides a convenient *sum rule* for states between E_F^* and E_F which allows us to verify that the DLTS $g(E)$ includes *all* states. Such a sum rule is important in particular to compare our $g(E)$ to $g_{SL}(E)$ since it might be argued that the large peak in $g_{SL}(E)$ at $E_c-0.4$ eV could escape DLTS detection by just such a hopping mechanism. It is clear from Fig. 27 that $g_{SL}(E)$ cannot possibly fit our data. The only evidence which we see for any states in the vicinity of $E_c-0.4$ eV are admittance effects most likely associated with the interface, as previously discussed.

Note that the type of C-V curve in Fig. 27 is *not* the sort of low-frequency—high-temperature measurement usually associated with the MOS C-V algorithm used for obtaining $g(E)$.[18] In that case the conditions require that deep states at E_F^* must respond fast enough to follow the measurement frequency. In the high-frequency C-V curve in Fig. 27 only the shallow states can follow the measurement frequency so that the capacitance is a measure of the width of the depletion region. This width at a given bias, however, depends on the sum of all deep levels from E_F^* to E_F as long as the dc voltage is changed slowly enough so that the deep-state occupation is at its steady-state value for each capacitance measurement. Such a measurement regime has the advantage that one may discriminate against the effects of interface states provided these states cannot follow the measurement frequency. Such discrimination is not possible in the low-frequency—high-temperature C-V measurement regime at low bias voltages. From Fig. 3 we can see that the condition of 100 kHz at

300 K for sample 5 (152) is relatively free from interface effects, i.e., the C-T curve is still in the plateau regime given by the bulk $g(E)$.

An additional useful role for steady-state measurements is for the study of high-resistivity undoped a-Si:H films. For such films, DLTS and TSCAP measurements are not particularly useful since E_F^* and E_F are both close to midgap. In such a case the C-T curves may be taken as primary data in order to obtain $g(E_F)$. An example is shown in Fig. 28 for an undoped a-Si:H film with $E_c - E_F = 0.87$ eV. The 100-Hz C-T data in part (a) of this figure show a dielectric turn-on between 325 and 350 K. The capacitance increase at higher temperatures is due to increasingly deeper states being able to follow the measurement frequency. As discussed in Ref. 35, it is particulary useful to express such C-T data in the form $C^2(dC/dT)^{-1}$ vs T as in Fig. 28(b). From the slope of such a plot we obtained the density of states at the Fermi level of 1.9×10^{15} cm^{-3}eV^{-1} for this sample.

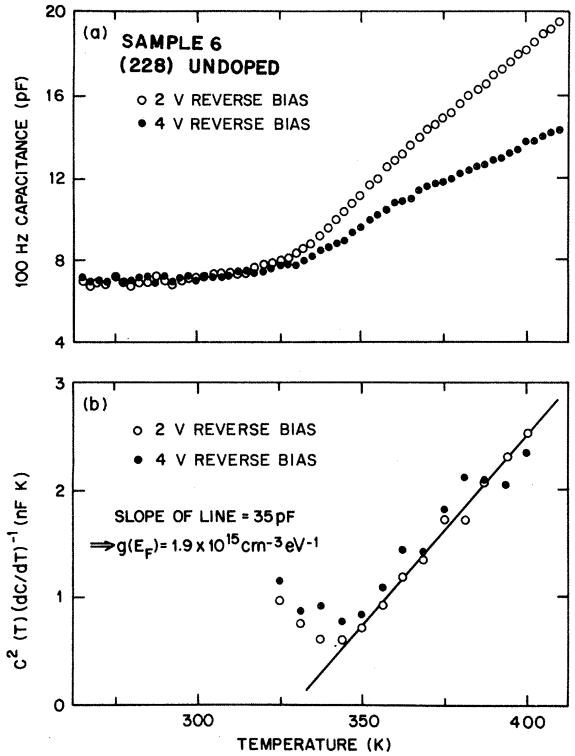

FIG. 28. Illustration of a technique for obtaining $g(E_F)$ in undoped films. (a) shows 100-Hz capacitance vs temperature data for two bias voltages in a typical high-resistivity undoped a-Si:H film. (b) shows replot of the above data in the form $C^2(dC/dT)^{-1}$ vs T. The slope of a line drawn through the data gives the density of states at the Fermi level as explained in Ref. 35.

Furthermore, the fact that the $C^2(dC/dT)^{-1}$ data are independent of bias voltage verfies that the resulting $g(E_F)$ value is representative of the *bulk* of the film.[35]

C. Effects of PH_3 doping

It is clear from a comparison of the densities of states for the three PH_3-doped samples in Figs. 19, 23, and 24 that extrinsic defects or impurities are playing an important role in determining $g(E)$ for a particular sample. The role of the PH_3 doping itself in determing the deep-state density may be assessed by comparing these films with ones deposited *without* PH_3 doping. In Fig. 29 we show a composite of the densities of states for the three PH_3-doped films in Figs. 19, 23, and 24, as well as two undoped films, one of which, sample 6 (228), is a typical high-resistivity film (see Fig. 28), while the other, sample 1 (106), was unintentionally doped with a donor of unknown origin (see Sec. II). Clearly, the addition of 60–300 ppm of PH_3 during growth results in a dramatic increase in the deep-state density.

Since it is impractical to generate a detailed

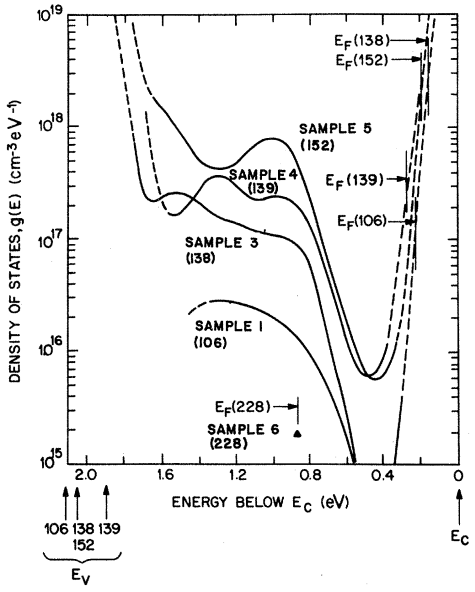

FIG. 29. Summary of the density-of-states results for five different a-Si:H films, the properties of which are given in Table I. The notion "E_F (sample number)" indicates the position of the bulk Fermi level in each film. The films have slightly different band gaps, and for purposes of comparison the curves are all normalized to the conduction-band edge E_c

$g(E)$ for each sample, the results for a large number of films are best presented in terms of the integrals over $g(E)$ which may be obtained from TSCAP data. Let us consider three trap concentration integrals, namely,

$$N_T(h)=\int_{E_v+0.5\text{ eV}}^{E_v+E_g/2} g(E)dE\ , \tag{17}$$

$$N_T(e)=\int_{E_c-E_g/2}^{E_c-0.5\text{ eV}} g(E)dE\ , \tag{18}$$

$$N_S=\int_{E_c-0.5\text{ eV}}^{E_c-E_F} g(E)dE\ . \tag{19}$$

The relationship of these integrals to measured parameters such as ΔC_e and ΔC_{e+h} from TSCAP is somewhat complex, but may be directly obtained by our numerical analysis.[35] For the specific case of 10-kHz TSCAP at -2-V bias for 2×10^{-3} cm^2 samples with a $g(E)$ of the same general shape as those in Fig. 29 we have derived the following empirical relationships which are valid to within 20%,

$$N_S=(7.7\times10^{12})C_1^2\ , \tag{20}$$

$$N_T(e)=N_S\left[\frac{3\beta}{3-\beta}\right]\ , \tag{21}$$

$$N_T(h)=\gamma N_S\left[\frac{6+\beta}{6-2\beta}\right]\ , \tag{22}$$

expressed in cm^{-3}, where

$$\beta\equiv\frac{N_T(e)}{N_S+N_T(e)/3}=4.9\left[\frac{C_0^2-C_1^2}{C_0^2}\right]\ , \tag{23}$$

$$\gamma\equiv\frac{N_T(h)}{N_S+N_T(e)/2}=5.4\left[\frac{C_2^2-C_0^2}{C_0^2}\right]\ . \tag{24}$$

The capacitance values (in pF) correspond to TSCAP values at 180 K where C_0 is the baseline curve, $C_1=C_0-\Delta C_e$ is the zero-bias-fill curve, and $C_2=C_0+\Delta C_{e+h}$ is the laser-fill curve (see Fig. 6). Thus we do not need to construct and numerically integrate a detailed $g(E)$ fit for each sample, but may obtain the deep-hole-trap concentration $N_T(h)$, the deep-electron-trap concentration $N_T(e)$, and the shallow-tail-state concentration N_S directly from simpler TSCAP measurements.

The results for a number of samples are presented in Fig. 30. Note the general trend that both $N_T(e)$ and $N_T(h)$ increase approximately linearly with the shallow-state density and tend to saturate for $N_S>10^{16}$ cm^{-3}. Among the PH_3-doped films the N_S values tend to increase monotonically with the [PH_3/SiH_4] concentration ratio. Aside from the dramatic increase over undoped films, however,

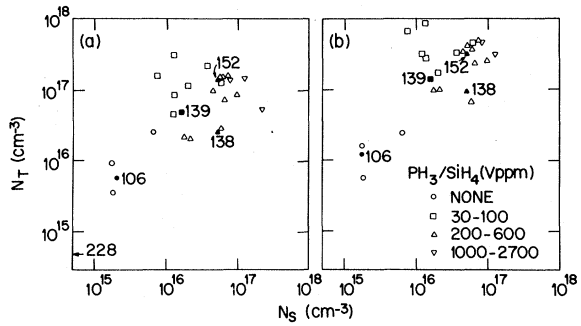

FIG. 30. Plot of deep gap-state integrals $N_T(e)$ and $N_T(h)$ vs the integral over shallow gap states N_s as obtained from TSCAP measurements and defined in the text. (a) is a plot of $N_T(e)$ for a number of samples while (b) is a plot of $N_T(h)$ for the same samples. The samples for which $g(E)$ is given in Fig. 29 are shown as solid symbols labeled with their respective sample numbers. For sample 6 (228) TSCAP is not possible. Thus, the value of $N_T(e)$ indicate by the arrow is estimated by scaling the $N_T(e)$ result for sample 3 (106) by the density-of-states ratio at $E_c-0.87$ eV in Fig. 29.

the relationship of $N_T(e)$ and $N_T(h)$ to the PH_3 concentration is not particularly well defined, and any trend is less than the sample-to-sample variations in deep-state density. Apparently there is a large increase in deep states (from less than 10^{15} to greater than 10^{17} cm^{-3} in the energy range $E_v+0.5$ eV to $E_c-0.5$ eV) corresponding to the introduction of a small amount of phosphorous (less than 30 ppm [PH_3]/[SiH_4]) with most additional phosphorous (up to 2700 ppm [PH_3]/[SiH_4]) being incorporated into band-tail states outside this energy range. More work is in progress to relate the specific changes in $g(E)$ to various impurities as well as to changes in the growth or annealing conditions.

Finally, the introduction of the integrals in Eqs. (17)–(19) allow us to give a more quantitative discussion of the conduction-band tail. Recall that we said $g(E)$ was adjusted to give the correct value of the steady-state capacitance just above turn-on. More specifically, this means that the integrals over $g(E)$ [Eqs. (17)–(19)] must satisfy the TSCAP data [Eqs. (20)–(24)]. In particular, the conduction-band tail is very sensitive to the experimental values of N_S and E_F in Eq. (19). The densities of states in Fig. 29 give values of N_S and E_F that are consistent if the conduction-band tail has the form

$$g(E)=10^{21}\exp(-E/E_0)\ , \tag{25}$$

in units of cm^{-3}eV^{-1}, where $E<0.3$ eV, and with

values of the characteristic energy parameter E_0 ranging between 25 and 34 meV for different samples. Note that such a rather steep exponential band tail is necessary not only to give a consistent fit to N_S for a given E_F but also to match up with the DLTS data at $E_c - 0.3$ eV. In addition, the upturn seen at low temperatures in the voltage pulse DLTS data for sample 3 (138) (see Figs. 24 and 25) provides rather direct evidence for an exponential band tail with $E_0 \simeq 25$ meV in that sample.

VII. COMPARISON WITH OTHER RESULTS

There have been many results in the literature in the past few years on the density of gap states in *a*-Si:H. The various techniques were discussed in the Introduction. The results may be roughly grouped into three categories. Namely, (i) those which give a qualitative shape for $g(E)$ over a large part of the gap, (ii) those which give a quantitative value for $g(E_F)$ for only a very narrow range of energy near E_F, and finally (iii) those which are quantitative over a wide range of energies. Our $g(E)$ results are one of the very few in the last category.

We may briefly summarize the relationship of our $g(E)$ to those in the literature as follows. In general, *all* quantitative measurements[2,6,7,9,18–20,24,25,55,56] give state densities at midgap in the range $10^{15} - 10^{18}$ cm^{-3}eV^{-1} for various samples produced and doped in different ways. This is precisely the range of values which we observe in our various samples (see Fig. 29). These differences in density are clearly extrinsic effects relating to sample purity and preparation conditions, just as is the case for deep states in crystalline semiconductors. The intrinsic-state density at midgap must therefore be in the range of 10^{15} cm^{-3}eV^{-1} or less.

On the other hand, the few results[2,6,7,18,55,56] giving $g(E)$ over a wide range of energies do not agree well with our general shape of $g(E)$ over the entire gap. In particular, most published results for $g(E)$ in the upper half of the gap differ dramatically from our results. We typically find a deep minimum ($< 10^{16}$ cm^{-3}eV^{-1}) at about 0.4 to 0.5 eV below E_c while others[2,6,7,55,56] often find a *peak* ($\gtrsim 10^{18}$ cm^{-3}eV^{-1}) in this energy range.

With the exception of ESR spin-density measurements, nearly all of the reliable quantitiative techniques involve junction space-charge layers in some way. However, as we have repeatedly stressed throughout this paper, the various types of space-charge measurements differ markedly in their ability to differentiate between bulk and interface effects. Historically, the first approach was to neglect surface effects in the analysis of field-effect[6] and *C-V* (Ref. 18) data. However, there is now considerable evidence that surface or interface effects can influence a wide variety of experiments.[10–15] In particular, a number of different quantitative studies[23,57,58] have shown that the interface state density is typically on the order of 10^{13} cm^{-2}eV^{-1}. This corresponds, for example, to a bulk density of 10^{18} cm^{-3}eV^{-1} over a depth of 1000 Å.

The surface-sensitive tunnel diode admittance measurements of Balberg[26–28] also support the idea that the field-effect density of states $g_{SL}(E)$ is dominated by surface or interface states. In this context we should keep in mind that the surface is not simply the outermost monolayer of atoms, but rather is an interfacial transition region of more or less oxidized *a*-Si:H which may also be affected by adsorbates.[11,12] The interface states may in fact be a large concentration of bulk states in a thin (~ 1000 Å) interfacial layer.[13] Balberg used very thin *a*-Si:H layers (1100 Å) and the classical crystalline MOS surface-state analysis to obtain a dominant peak in $g(E)$ in the vicinity of $E_c - 0.45$ eV. He argued on the basis of the many experiments[2] supporting $g_{SL}(E)$ that in spite of the clear surface sensitivity of his method the 0.45-eV peak in $g(E)$ should be attributed to the bulk as well as to the surface. We do not believe that Balberg's results can be extended to the true bulk properties farther than 1000 Å from the surface. In fact, one can argue that Balberg's results along with our bulk $g(E)$ are solid evidence that the 0.4-eV E_x peak in $g_{SL}(E)$ is due primarily to interface states.

While most of the discrepancies between our results and those obtained by field effect and *C-V* are thus most readily explained as being due to interface effects, there are other experiments,[2] apparently not subject to the interface-state problem, which also seem to support a density of states rather similar to $g_{SL}(E)$. One such class of experiments is the shift of the Fermi level with doping[2] and with ion implantation.[2,55] It has been pointed out[1] that the analysis of such doping data rests on two major assumptions, namely, that the efficiency of doping is independent of the doping level, and that the doping procedure itself does not modify $g(E)$. However, our results imply that these assumptions are not valid, i.e., we show that $g(E)$ changes considerably with doping (see Sec. VI C). The major

effect of donor dopants on states in the lower half of the gap can also be inferred from other results in the literature.[56,59–62] Indeed, given that the doping results agree with $g_{SL}(E)$ assuming the motion of E_F through a *rigid* $g(E)$, one can argue that the *increase* of $g(E)$ below midgap, which we observe with added phosphorous, proves that the density of states must be substantially lower than $g_{SL}(E)$ in the upper half of the gap. In fact, our $g(E)$ does agree with the motion of E_F with doping when one takes into account the increases in deep states shown in Fig. 30.

When one measures the temperature dependence of the Fermi level rather than its shift with added dopants, the results imply a density of states which is strikingly similar to our $g(E)$. Beyer *et al.*[59] and Vanier *et al.*[63] have analyzed the statistical shifts in the Fermi level with temperature in a variety of doped and undoped *a*-Si:H samples grown by glow discharge and come to the conclusion that there must be a very pronounced minimum in $g(E)$ at 0.4 to 0.5 eV below the conduction-band edge. The same minimum at $E_c - 0.4$ eV is also inferred from an analysis of the room-temperature photoconductivity of a large number of samples.[63] In addition, from infrared and thermal quenching of the photoconductivity, Vanier *et al.*[63] also find strong evidence for three distinct peaks in $g(E)$ in the lower half of the gap. The positions of these peaks do not agree with the three peaks which we see in Fig. 29, but as we discuss shortly, one should not necessarily expect good agreement between optical and thermal emission energies. Therefore, the observation of *three* types of defects by both measurements should be redarded as significant.

Further support for our density of states in the energy region of 0.3 to 0.5 eV below E_c comes from the results of Crandall[32] on PH_3-doped and undoped films. These room-temperature current transient measurements give density-of-state information *only* in the range of 0.3 to 0.5 eV below E_c, but agree almost exactly with our range of typical values for $g(E)$ at these energies; namely, he finds 5×10^{14} cm^{-3}eV^{-1} in undoped films and $2-9\times10^{15}$ cm^{-3}eV^{-1} in PH_3-doped films. He originally called these values merely lower bounds on $g(E)$. However, as a result of our theoretical analysis and our experience with current transients as well as capacitance transients and steady-state admittance, we find that these values are most likely accurate to within a factor of about 2 to 4 for the true bulk $g(E)$. We believe that the apparently incomplete nonexponential trap saturation for long filling times which he reported is probably due to the unavoidable filling of interface states and the high sensitivity of current transients to the interface (see Sec. IV A). We also see evidence for interface effects in our current-transient DLTS data, but the proper analysis of such work is very complex and beyond the scope of this paper.

Support for our $g(E)$ also comes from the recent drift mobility measurements reported by Tiedje[64] which imply the existence of an exponential conduction-band tail in *a*-Si:H with a characteristic energy as small as 25 meV [see Eq. (25)]. Vanier *et al.*[63] also find a characteristic band-tail energy of 0.02 eV from their analysis of photoconductivity. As we discussed in Sec. VI C, a narrow band tail of this sort is necessary to be consistent with our overall results. We find characteristic energies between 25 and 34 meV for the samples in Fig. 29.

We next should like to comment on the qualitative densities of states proposed on the basis of optical measurements.[61,65] Both the $g(E)$ models of Anderson *et al.*[61] and the Street[65] bear a family resemblance to $g_{SL}(E)$ in that all three are dominated by a rather large density of hole-trap states in the lower half of the gap and have a distinct but somewhat smaller electron trap peak in the region of 0.3 to 0.5 eV below E_c. To resolve the apparent disagreement between thse optical results and our $g(E)$ one must conclude either that (i) the electron trap peak in $g(E)$ is not a valid conclusion from the optical data, or that (ii) we have failed to observe by admittance and DLTS an important characteristic feature of $g(E)$. Let us discuss these two points in more detail.

First, an obvious problem with optical absorption or photoluminescence measurements in general is that one cannot *a priori* determine from the spectra alone whether subband-gap features are due to states near E_c or E_v. This is to be contrasted with capacitance DLTS spectra where the hole- or electron-trap nature of the emission process is obvious from the sign of the signal. Thus, while both Anderson *et al.*[61] and Street[65] have gone to some lengths to locate their electron trap peak with ancillary experiments, we do not believe that their assignments are as definitive as could be made with DLTS. On the other hand, we see no obvious flaw in their arguments and so must consider the second alternative; namely, could DLTS miss such a peak? This might be possible if the electron trap peak were an *excited state.* We can see by our techniques only the density of ground states (or

long-lived metastable states) of the system. Therefore, if the initial state responsible for the 0.9-eV defect luminescence band[61,65] were actually a short-lived excited state of the defect involved, we would not see such a state in our $g(E)$. Indeed, *none* of the transport or space-charge measurements cited[2] to justify $g_{SL}(E)$ would see such a state. Therefore, the question of comparing our $g(E)$ to the $g_{SL}(E)$-like densities of states based on optical measurements is quite distinct from the comparison discussed above for $g_{SL}(E)$ as it relates to transport and dynamic electrical response.

The optical absorption and spectral dependence of photoconductivity for PH_3-doped samples, on the other hand, are actually in good agreement with our form of $g(E)$.[61,62,66,67] From the typical $g(E)$ curves in Fig. 29 we would predict a subband-gap optical absorption shoulder in PH_3-doped films with a threshold of roughly 0.7 to 0.9 eV. Furthermore, this absorption shoulder should increase with PH_3 doping. If we use an optical cross section typical of most deep levels in crystals[68] ($\sigma^0 \simeq 10^{-16}$ cm^2), we would predict an absorption coefficient $\alpha = \sigma^0 N_T$ of 10 cm^{-1} for a deep level concentration of $N_T \simeq 10^{17}$ cm^{-3}. Considering the crudeness of the argument it indicates surprisingly good agreement between our typical values for N_T in doped films and the shape and magnitude of the optical absorption in doped films. Indeed, Jackson and Amer[67] obtain a value of $\sigma^0 \sim 10^{-16}$ cm^2 from the relationship between subband-gap absorption and ESR spin density.

We wish to point out, however, that detailed comparisons between optical and transport properties should be undertaken with some care. In particular, one should always keep in mind that different types of energy scales are relevant to different types of measurements. There are actually three different kinds of energy measurements for deep levels, namely, (1) thermal equilibrium energies, (2) thermal activation energies, and (3) optical energies. The thermal equilibrium energy is the proper quantity to use in a Boltzmann factor or Fermi function for the distribution of state occupations in equilibrium. This is the sort of energy which is obtained from field-effect, steady-state C-V, or dc transport measurements. The thermal activation energy is the proper quantity is use in a Boltzmann factor governing rate or reaction processes. This is the energy relevant to DLTS measurements. The equilibrium and activation energies for a deep level are rigorously related according to detailed balance (microscopic quantum-mechanical reversibility) by the energy barrier to capture.[30] This capture barrier is given by the activation energy of the capture cross section. The optical energy of a deep level depends on the degree of lattice relaxation associated with the initial and final states (Stokes shift). The optical zero-phonon line for ionization of a deep-level defect is the same as the thermal equilibrium energy. In general, however, the measured optical energies may be quite different from either type of thermal energy.

An additional point concerning energy scales is the thermodynamic question of whether one is measuring the Gibbs free energy G or the enthalpy H, with $G = H - TS$, where S is the entropy.[49] The proper energy to use in a Boltzmann factor for either case (1) or (2) is G. However, the energy actually measured by taking the slope of an Arrhenius plot is H.[30,49] The enthalpy H is often referred to as the free energy extrapolated to $T=0$; this may be verified by a trivial thermodynamic identity.[30,49] Thus, the value of E_F inferred from the logarithmic slope of the temperature dependence of the conductivity as well as the DLTS energy scale are really *enthalpies* in the thermodynamic sense. Optical energies are always Gibbs free energies. The difference between G and H depends on temperature and for the case of deep levels in crystals is equal to or less than the temperature shift of the band gap.[30,49,69,70]

According to the capture data in Sec. V (Fig. 10) we know that the capture barrier for electron states is rather small (<0.1 eV), if such a barrier exists at all. This is less than the overall 10% level of uncertainty in our absolute energy scale and is therefore negligible for our present state of sophistication. Thus, we may safely compare the DLTS energy scale [type (2)] and field-effect energy scale [type (1)] to within our stated accuracy. On the other hand, we have no idea whether lattice-relaxation effects are large enough to affect the relationship of our $g(E)$ to an optically determined density of states. It is comforting, however, to note that our values for E_g (1.9–2.1 eV $\pm 10\%$) are basically in agreement with those obtained from optical absorption by the rather empirical quadratic extraplation technique (1.7–1.9 eV).

Finally, let us compare our thermal emission parameters with those typical of deep levels in crystals. Our value of $\nu_0 \sim 10^{13}$ s^{-1} is about an order of magnitude larger than the typical ν_0 measured for most deep levels in crystalline silicon or III-V compound semiconductors. However, the

range of ν_0 values which has been seen in crystals is typically between about 5×10^{10} and 2×10^{13} s^{-1}, so that the value we measure here is not unreasonable. The somewhat large value of ν_0 in *a*-Si:H could be explained by the 10% uncertainty in the activation energy used to obtain ν_0, but the difference may have a more fundamental significance. In particular, according to the definition of ν_0 in Eq. (12), this prefactor could be somewhat larger in an amorphous material simply because one might expect the density of states N_c to be larger at the mobility edge of a disordered material than at the bottom of a well-defined band in a crystal. Similarly, one might also reasonably expect the matrix element for free-to-bound transitions to be somewhat larger in disordered materials.

The broader question of the relationship of the defects in *a*-Si:H to those in crystalline Si is a very interesting one which has only been discussed in very general terms in the literature. We hope that with the application of the powerful measurement techniques discussed in this paper it will someday be possible to experimentally relate the defects in these two similar, yet very different, types of system.

VIII. SUMMARY AND CONCLUSIONS

We have given a comprehensive discussion of our studies of the gap states in undoped and PH_3-doped *a*-Si:H. To obtain the density of gap states $g(E)$ we have used a wide variety of experimental techniques to study Schottky-barrier diodes fabricated from *a*-Si:H. These rather diverse meaurements may be viewed in a unified way, however, as a result of our theoretical work in the adjoining paper. By using this theory we can model the full dynamic response of a generalized Schottky-barrier structure formed on any material with an arbitrary spatial and energy variation of the density of gap states. Although the numerical analysis techniques are rather sophisticated, the basic physics of the theory rests on two very simple ideas: namely, (1) Poisson's equation is used to describe the spatial variation of the electron potential in terms of the distribution of trapped charge, and (2) changes in metastable trapped charge distributions are governed by thermal emission of trapped electrons to an electron conduction path and trapped holes to a hole conduction path.

By applying this general dynamic response theory to the case of *a*-Si:H we have been able to consistently fit a wide variety of steady-state and transient admittance measurements in terms of a bulk density-of-states function $g(E)$ for the particular sample being studied. The $g(E)$ functions obtained in this manner give a detailed picture of the bulk density of states over nearly 70% of the gap in many cases. The shape of $g(E)$ for a wide variety of *n*-type *a*-Si:H samples differs in detail and in overall concentration from sample to sample, but a general trend nevertheless is apparent. The general shape of $g(E)$ is dominated by a broad band of gap states extending up from the valence band to approximately midgap. The concentration range of this band of states in various samples is typically from 10^{17} to 10^{18} $cm^{-3}eV^{-1}$ in PH_3-doped films, and as low as 2×10^{15} $cm^{-3}eV^{-1}$ in undoped films. Incipient structure within this band of hole trap states suggests the existence of several different types of defects of as yet unknown structure and origin. The upper half of the gap is by contrast relatively free of deep states so that a pronounced minimum ($<10^{16}$ $cm^{-3}eV^{-1}$) exists in all of our $g(E)$ curves between about 0.3 and 0.6 eV from the electron conduction path. The conduction-band tail is found to be relatively narrow and consistent with an exponential having a characteristic energy of 25—35 meV.

An important fitting parameter in our dynamic response model is the thermal emission band gap E_g. This is the energy separation between the respective conduction paths for thermally emitted electrons and holes in the barrier depletion layer. For our PH_3-doped samples, the thermal emission gap is found to be in the range of 1.9 to 2.05 eV with an overall $\pm10\%$ uncertainty. In view of the good agreement between this gap and the optically determined gap typical of *a*-Si:H grown under similar conditions, we believe that the conduction paths for the thermally emitted carriers in our measurements are near the mobility edges and not due to hopping conduction in bands of deep states.

In the course of this work we have made nearly every possible type of dynamic response measurement. Specifically, we have measured the steady-state complex admittance (capacitance and conductance) as a function of temperature (50—400 K), frequency (100 Hz—1 MHz), and bias voltage (−10— + 10 V). We have also explored metastable transient effects by measuring isothermal capacitance and current transients over the time scale of a few ms to a few h. Finally, we have utilized the various thermal scan techniques of TSCAP, TSC, and current and capacitance DLTS.

While many of these measurements have been used previously in *a*-Si:H, no one has attempted to correlate the results of a number of different measurement methods into a single consistent picture for the density of gap states.

As a result of our work we have reduced the very large number of possible measurements to a selected few which have the appropriate complementary strengths to provide a reliable set of data to be fit by our general dynamic response model. Our main measurement tool is capacitance DLTS. This is chosen because the raw data is in a form which is very close to the shape of $g(E)$ and the calculated DLTS spectra are very sensitive to changes in $g(E)$. Thus the DLTS spectra are the primary data used to obtain the shape of $g(E)$ and the energy scale. An important complementary measurement is a set of thermally stimulated capacitance (TSCAP) scans. These provide reliable experimental measures of various integrals over $g(E)$ and are therefore the primary data used to determine the concentration scale. A set of capacitance-versus-temperature scans for various frequencies is used to obtain the Fermi-level position E_F and an integral over the conduction-band tail states below E_F. Finally, a high-frequency C-V curve is used to experimentally measure the overall integral of $g(E)$ between midgap and E_F in a mode which is particularly insensitive to interface states. None of these measurements taken alone would be sufficient to obtain $g(E)$ with the level of confidence which we very conservatively set as $\pm 10\%$ on the absolute energy scale and within a factor of 2 on the absolute concentration scale.

The importance of distinguishing between bulk and interface states is a central concern in our work. Therefore we have paid special attention to fundamental differences among the various dynamic response measurement techniques in their relative sensitivity to the interface. According to this criterion the measurements can be classed into two general groups: (1) capacitance transients and dC/dV measurements have a vanishingly small sensitivity to the interface, while (2) current transient and steady-state low-frequency capacitance measurements have maximum sensitivity at the interface. The characteristic differences which we observe between measurements of these two general types thus indicate a large concentration of deep states at or near the interface. Such states are the best explanation for the significant differences between our bulk $g(E)$ and earlier models for the density of states in *a*-Si:H.

We believe that the power of these types of dynamic response measurements to determine $g(E)$ is quite clear from this work. However, since our results show that the bulk density of states in doped *a*-Si:H films is determined by extrinsic defects or impurities, the real utility of our methods will be in the characterization of a wide variety of different samples in an effort to understand the origins of these extrinsic parts of $g(E)$. Hopefully, such studies will enable us to ultimately understand the fundamental processes which control such phenomena as luminescence, photoconductivity, subband-gap optical absorption, doping efficiency, and other optical and transport effects.

ACKNOWLEDGMENTS

We wish to thank A. M. Sergent for assistance in constructing apparatus and recording data and A. J. Williams and A. Savage for assistance in sample growth and diode fabrication. We also gratefully acknowledge helpful discussions and suggestions from V. Narayanamurti.

[1]For a recent review, see N. F. Mott, J. Phys. C 13, 5433 (1980), and references cited therein.

[2]P. G. LeComber and W. E. Spear, in *Amorphous Semiconductors,* Vol. 36 of *Topics in Applied Physics,* edited by M. H. Brodsky (Springer, New York, 1979), p. 251.

[3]J. D. Cohen, D. V. Lang, J. C. Bean, and J. P. Harbison, J. Non-Cryst. Solids 35-36, 581 (1980).

[4]J. D. Cohen, D. V. Lang, J. P. Harbison, and J. C. Bean, Solar Cells 2, 331 (1980).

[5]J. D. Cohen, D. V. Lang, and J. P. Harbison, Phys. Rev. Lett. 45, 197 (1980).

[6]W. E. Spear and P. G. LeComber, J. Non-Cryst. Solids 8-10, 727 (1972).

[7]N. B. Goodman, H. Fritzsche, and H. Ozaki, J. Non-Cryst. Solids 35-36, 599 (1980).

[8]H. Fritzsche, Solar Cells 2, 289 (1980).

[9]D. E. Carlson and C. R. Wronski, in *Amorphous Semiconductors,* Vol. 36 of *Topics in Applied Physics,* edited by M. H. Brodsky (Springer, New York, 1979), p. 287.

[10]I. Solomon, T. Dietl, and D. Kaplan, J. Phys. (Paris)

39, 1241 (1978).
[11]M. Tanielian, H. Fritzsche, C. C. Tsai, and E. Symbalisty, Appl. Phys. Lett. 33, 353 (1978).
[12]M. Tanielian, M. Chatani, H. Fritzsche, V. Smid, and P. D. Persans, J. Non-Cryst. Solids 35-36, 575 (1980).
[13]D. G. Ast and M. H. Brodsky, J. Non-Cryst. Solids 35-36, 611 (1980); Philos. Mag. B 41, 273 (1980).
[14]I. Solomon and M. H. Brodsky, J. Appl. Phys. 51, 4548 (1980).
[15]W. Rehm, R. Fischer, and J. Beichler, Appl. Phys. Lett. 37, 445 (1980).
[16]C. N. Berglund, IEEE Trans. Electron Devices ED-13, 701 (1966).
[17]I. Solomon (private communication).
[18]M. Hirose, T. Suzaki, and G. H. Dohler, Appl. Phys. Lett. 24, 234 (1979).
[19]P. Viktorovitch and D. Jousse, J. Non-Cryst. Solids 35-36, 569 (1980).
[20]J. Beichler, W. Fuhs, H. Mell, and H. W. Welch, J. Non-Cryst. Solids 35-36, 587 (1980).
[21]A. J. Snell, K. D. Mackenzie, P. G. LeComber, and W. E. Spear, J. Non-Cryst. Solids 35-36, 593 (1980).
[22]M. Shur, W. Czubatyj, and A. Madan, J. Non-Cryst. Solids 35-36, 731 (1980).
[23]P. Viktorovitch, J. Appl. Phys. 52, 1392 (1981).
[24]T. Tiedje, C. R. Wronski, and J. M. Cebulka, J. Non-Cryst. Solids 35-36, 743 (1980).
[25]P. Viktorovitch and G. Moddel, J. Appl. Phys. 51, 4847 (1980).
[26]I. Balberg and D. E. Carlson, Phys. Rev. Lett. 43, 58 (1979).
[27]I. Balberg, J. Non-Cryst. Solids 35-36, 605 (1980).
[28]I. Balberg, J. Electronic Mater. 9, 797 (1980); Phys. Rev. B 22, 3853 (1980).
[29]D. C. Tsui, in *Handbook of Semiconductor Physics,* edited by W. Paul (Academic, New York, 1980).
[30]D. V. Lang, in *Thermally Stimulated Relaxation in Solids,* Vol. 37 of *Topics in Applied Physics,* edited by P. Brauhlich (Springer, Berlin, 1979), p. 93.
[31]D. V. Lang, J. Appl. Phys. 45, 3023 (1974).
[32]R. S. Crandall, J. Electron. Mater. 9, 713 (1980).
[33]L. Vieux-Rochaz and A. Chenevas-Paule, J. Non-Cryst. Solids 35-36, 737 (1980).
[34]W. Fuhs and M. Milleville, Phys. Status Solidi B 98, K29 (1980).
[35]J. D. Cohen and D. V. Lang, following paper, Phys. Rev. B 25, 5321 (1982).
[36]J. C. Knights, in *Structure and Excitation of Amorphous Solids, Willamsburg, Virginia, 1976,* Proceedings of an International Conference of Structure and Excitation of Amorphous Solids, edited by G. Lucovsky and F. L. Galeener (American Institute of Physics, New York, 1976), p. 296; Philos. Mag. 34, 663 (1976).
[37]J. C. Knights, G. Lucovsky, and R. J. Nemanich, J. Non-Cryst. Solids 32, 393 (1979).
[38]All gases purchased from Matheson Gas Products, East Rutherford, NJ 07073.
[39]J. C. Knights and R. A. Lujan, Appl. Phys. Lett. 35, 244 (1979).
[40]S. R. Ovshinsky and A. Madan, Nature (London) 276, 482 (1978).
[41]R. C. Henderson, J. Electrochem. Soc. 119, 772 (1972).
[42]D. L. Staebler and C. R. Wronski, J. Appl. Phys. 51, 3262 (1980).
[43]D. V. Lang, J. D. Cohen, J. P. Harbison, and A. M. Sergent, Appl. Phys. Lett. 40, 474 (1982).
[44]D. V. Lang, J. D. Cohen, and J. P. Harbison, Phys. Rev. Lett. 48, 421 (1982).
[45]G. L. Miller, D. V. Lang, and L. C. Kimerling, Ann. Rev. Mater. Sci. 7, 377 (1977).
[46]D. L. Losee, J. Appl. Phys. 46, 2204 (1975).
[47]S. M. Sze, *Physics of Semiconductor Devices* (Wiley-Interscience, New York, 1969).
[48]D. S. Day, M. Y. Tsai, B. G. Streetman, and D. V. Lang, J. Appl. Phys. 50, 5093 (1979).
[49]Equation (12) neglects entropy and degeneracy factors since they are beyond the scope of our current understanding of amorphous semiconductors. The interested reader is referred to the recent discussion of defect thermodynamics by D. V. Lang, H. G. Grimmesiss, E. Meijer, and M. Jaros, Phys. Rev. B 22, 3917 (1980).
[50]D. V. Lang, in *Radiation Effects in Semiconductors 1976,* edited by N. B. Urli and J. W. Corbett (Institute of Physics, London, 1977), p. 70.
[51]L. C. Kimerling and J. R. Patel, Appl. Phys. Lett. 34, 73 (1979).
[52]R. A. Street, Philos. Mag. B 37, 35 (1978).
[53]J. C. Knights, J. Non-Cryst. Solids 35-36, 159 (1980).
[54]R. J. Nemanich, D. K. Biegelsen, and M. P. Rosenblum, J. Phys. Soc. Jpn. 49, Suppl. A, 1189 (1980).
[55]P. G. LeComber, W. E. Spear, G. Muller, and S. Kalbitzer, J. Non-Cryst. Solids 35-36, 327 (1980).
[56]H. Overhof and W. Beyer, J. Non-Cryst. Solids 35-36, 375 (1980).
[57]J. C. Knights and D. K. Biegelsen, Solid State Commun. 22, 133 (1977).
[58]C. R. Wronski and D. E. Carlson, Solid State Commun. 23, 421 (1977).
[59]W. Beyer, H. Mell, and H. Overhof, in *Proceedings of the Seventh International Conference on Amorphous and Liquid Semiconductors,* edited by W. E. Spear (Edinburgh University Press, Edinburgh, 1977), p. 328.
[60]B. Von Roedern and G. Moddel, Solid State Commun. 35, 467 (1980).
[61]D. A. Anderson, G. Moddel, and William Paul, J. Non-Cryst. Solids 35-36, 345 (1980).
[62]G. Moddel, D. A. Anderson, and William Paul, Phys. Rev. B 22, 1918 (1980).
[63]P. E. Vanier, A. E. Delahoy, and R. W. Griffith, in *Tetrahedrally Bonded Amorphous Semiconductors, Carefree, Arizona, 1981,* edited by R. A. Street, D. K. Biegelsen and J. C. Knights (American Institute of

Physics, New York, 1981), p. 227.
[64]T. Tiedje, Bull. Am. Phys. Soc. 26, 330 (1981); G. D. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. 47, 1480 (1981).
[65]R. A. Street, Phys. Rev. B 21, 5775 (1980).
[66]S. Yamasaki, K. Nakagawa, H. Yamamoto, A. Matsuda, H. Okushi, and K. Tanaka, in *Tetrahedrally Bonded Amorphous Semiconductors, Carefree, Arizona, 1981,* edited by R. A. Street, D. K. Biegelsen, and J. C. Knights (American Institute of Physics, New York, 1981), p. 258.
[67]W. Jackson and N. M. Amer, in *Tetrahedrally Bonded Amorphous Semiconductors, Carefree, Arizona, 1981,* edited by R. A. Street, D. K. Biegelsen, and J. C. Knights (American Institute of Physics, New York, 1981), p. 263.
[68]H. G. Grimmeiss, Ann. Rev. Mater. Sci. 7, 341 (1977).
[69]C. D. Thurmond, J. Electrochem. Soc. 122, 1133 (1975).
[70]J. A. Van Vechten and C. D. Thurmond, Phys. Rev. B 14, 3539 (1976).