A novel computer based pseudo-logarithmic capacitance/conductance DLTS system specifically designed for transient analysis

William A. Doolittle and Ajeet Rohatgi
School of Electrical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 15 May 1992; accepted for publication 10 August 1992)

A new inexpensive, simple to construct, PC based deep level transient spectroscopy (DLTS) system has been developed that efficiently digitizes and analyzes capacitance and conductance transients by conventional DLTS methods (boxcar, rectangular, and exponential) as well as by several transient methods (nonlinear least squares, modulation function methods, and correlation method of linear predictive modeling). A unique "pseudo-logarithmic" sample storage scheme allows *each transient* to be sampled at more than 11 different rates, permitting 3 to 5 decades of time constants to be observed in one thermal scan allowing the resolution of closely spaced defect energy levels. The high system flexibility allows weighting times (or transient observation times) to be selected between <200 μs to 3 days. This is the first report that details the merits, including faster response and recovery times, of using the new Boonton 7200 capacitance meter in a DLTS system. This is also the first reported use of the pseudo-logarithmic sample spacing to enhance the range of time constants observable in a single thermal cycle as well as the first use of selectable resistance values in conjunction with selectable capacitance values to allow the conductance and capacitance to be sampled simultaneously in high resolution mode (<1 fF and <0.01 μS). It is shown experimentally that the transient analysis capability of this system can separate closely spaced deep levels where the conventional DLTS methods fail.

I. INTRODUCTION

Deep level transient spectroscopy (DLTS)[1,2] is one of the most powerful tools for semiconductor characterization because it reveals information about several characteristics of electrically active defects present in such materials. Many computer controlled systems have been reported in the literature.[3–11] These systems can often be expensive and/or difficult to build. Additionally, conventional DLTS suffers from one very important limitation that it cannot resolve very closely spaced defect energy levels or distributed levels within the band gap. In recent years, many researchers have attempted to extend the power of DLTS beyond this inherent limitation by using spectral fitting methods[12,13] as well as different types of transient analyses including nonlinear least squares fitting,[14,15] spectral analysis DLTS (SADLTS),[16–18] fast Fourier transform (FFT),[19,20] method of moments,[21,22] correlation method of linear predictive modeling,[23] modulation function methods,[24] and mixed methods.[25] Most of the commercial DLTS systems available do not lend themselves to digitation of the transients without some modifications that can be, at best, somewhat cumbersome and, at worst, very restrictive. Specifically, it is difficult and/or time consuming to observe the transient over many decades of time without manual intervention due to the nonintegrated nature of the pulse generation and sampling processes. In this paper, we describe a DLTS system that was designed and built specifically for the fast and efficient acquisition of capacitance/conductance transients for later analysis by any of the above methods. The new system is very inexpensive and easy to construct. It not only represents a significant improvement over the conventional DLTS system due to the incorporation of the Boonton 7200, but its capabilities also include the conductance analogs of DLTS[26] and an extremely efficient transient acquisition system that allows a *single* transient to be sampled at up to 11 (or more) different sample rates without changing the hardware configuration. We will first describe the unique requirements of transient analysis routines and how our pseudo-logarithmic sample scheme eloquently meets these requirements, followed by a description of the hardware and software of the system. Finally, in order to highlight this systems advantages, including the separation of closely spaced deep levels, we present a comparison of the system to one commercially available DLTS system.

II. PSEUDO-LOGARITHMIC STORAGE SCHEME FOR DLTS TRANSIENT ANALYSIS

Most existing methods used in DLTS transient analysis, though extremely diverse in approach, are extremely time consuming. Many are iterative in nature, others, like the correlation method, require many extensive calculations for each temperature. Thus, it is desirable to analyze only enough data to retrieve the information required, specifically, the time constant of the transient at each temperature. If, for example, the capacitance transient time constant at a certain temperature T_1 is 250 μs, there is no need to observe the transient over a 1 s time period. Doing so would entail storing and processing more "zeros" than relevant "nonzeros." However, the time constant at $T_2 < T_1$ may be 0.1 s. At this temperature, one might want to observe the transient over a 1 s time period and all the data points acquired in this period would be relevant. One solution to this might be to manually vary the "observation

window" while the measurement is progressing. This is typically performed by changing the digitation rate of the digitizer. Since we are interested in an automated system, this is not a viable option. Another solution is to set the observation window to 1 s and simply record the zeros. If the digitation is performed fast enough to accurately record the 250 μs time constant, a tremendous amount of nonrelevant data is acquired at temperatures around T_1. Additionally, the hard disk memory required to store such data could easily exceed 500 kbytes per temperature. This is clearly not feasible. Still another solution is to sample at many different rates at each temperature. Since typically, many transients are averaged at each temperature, the time required to average enough transients at several different rates would drastically slow the measurement. A more efficient solution, and the one we have adopted, is to sample at a fast rate for a long time but only store and process the data points corresponding to the different sample rates of interest. For example, if 131 072 data points are taken at a base sample rate of 100 kHz, one could store the first 128 (or 256, 512, 1024, etc.) points then extract another 128 points from the same transient starting from the original starting point but this time taking every other data point. Then take another 128 points at every fourth point, etc. This storage scheme allows the transient to be observed at 100 kHz for 1.28 ms, 50 kHz for 2.56 ms, 25 kHz for 5.12 ms, etc., up to 97.65625 Hz for 1.31072 s. Thus, many decades of time can be sampled from the same transient without permanently storing or processing redundant data. Additionally, only one transient need be acquired greatly accelerating the measurement. If one makes the observation that the first 64 data points of the 50 kHz sample set are included in the 100 kHz sample set and likewise for the other nine sample sets, one can reduce the original 131 072 data points to only 768 without the loss of any essential information. If more than 128 data points in each sample set are desired, than one merely needs to store more than 768 total points (example: a subset size of 256 and a base buffer size of 131 072 would require the permanent storage and processing of 1408 total points).This efficient and smart pseudo-logarithmic storage scheme is an important part of our system.

The total number of rates (available subsets of the base sample rate) depends on the size of the total base buffer allocated for the transient (131 072 in the above example) and the number of points desired in each subset (subset size=128 in the above example) and can be obtained by the following equation:

$$\text{total number of rates} = 1+\log_2\left(\frac{\text{total base buffer size}}{\text{subset size}}\right), \quad (1)$$

while the ith sample rate is given by

$$R_i=\frac{\text{base sample rate}}{2^{i-1}}, \qquad i=1,2,\ldots, \text{ total number of rates}. \quad (2)$$

It is important to recognize that the different sample sets at the different rates are for the same transient decay as opposed to resetting a digitizer and acquiring a different transient. The above example is equivalent to sampling one transient by as many as 11 different digitizers operating in parallel, but more or less substrates are possible by adjusting the base buffer and subset sizes.

Clearly, this scheme can be implemented by a digitizer that downloads the data to a computer which then extracts the relevant data. However, digitizers with this much memory (256 kbytes in the above example) are expensive when compared to the solution presented later and rarely have the 12 bit precision used here. Additionally, this large amount of data requires a substantial amount of time to be transferred over the IEEE-488.2 bus even when a direct memory access (DMA) transfer is used. A better solution would be to store the transients directly in the computer memory and bypass the transfer process all together. This is exactly what occurs in our system and will be described later in the hardware section.

The above scheme insures that transient decays that fall within a large time constant spread are optimally sampled without knowing the time constant ahead of time. Another advantage of the scheme should now become evident. Some transient analysis methods work best for nonuniform sample steps. Specifically, SADLTS[18] and nonlinear least squares fitting prefer to have the highest density of points (in time) at the beginning of the transient where the derivative with respect to time is the greatest. This allows the routine to more accurately determine the derivative and aids in the convergence of such routines. For such routines, all 768 points are processed. Most of the other transient analysis methods require uniform sample spacing and thus only the appropriate subset is passed to such routines. Clearly, this scheme of data storage and processing satisfies all the requirements of all the current transient analysis methods while optimizing the overall measurement time, processing time, as well as memory and storage requirements.

Another limitation that commercial DLTS systems impose on transient analysis is the lack of knowledge of the exact baseline capacitance.[27] This baseline capacitance is essential for many transient analysis routines to eliminate the dc component of the decay. In our system, we record the capacitance verses temperature curve as the sample cooled down under reverse bias. Thus, even at temperatures where deep traps emit there carriers very slowly, the baseline capacitance is known to within a few femtoseconds.

III. HARDWARE DESCRIPTION

A. General description

The hardware configuration described below is elegantly simple. The schematic diagram of the system is shown in Fig. 1. The system is based around the new Boonton electronics model 7200 capacitance meter. To our knowledge, this is the first paper detailing the merit of using this meter in a DLTS system. The sample capaci-

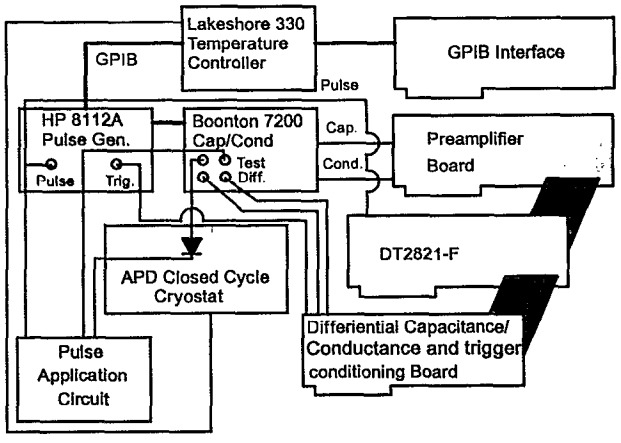

FIG. 1. DLTS system block diagram.

tance and conductance is read by the Boonton 7200 and fed to the analog-to-digital converter inputs of the Data Translation DT2821-F data acquisition board by means of two preamplifiers that extract the dc component, amplify and filter the signal. The DT2821-F board operates at a maximum rate of 150 kHz and minimum rate of 0.5 Hz and stores the data directly into the AT computer's memory via DMA transfers. This eliminates the need for downloading data from a remote digitizer. Furthermore, all processing of the data, including averaging, displaying, and even calculation of the conventional DLTS spectrum points is performed immediately after the data at each temperature is acquired. The DT2821-F board also contains a 16 bit digital output port which is used to output a binary code to a relay bank that switches different capacitor and resistor values from 0.125 up to 256 pF and from 0.25 to 512 μS in 0.125 pF and 0.25 μS steps onto the differential impedance leads of the Boonton 7200. This is used to offset the baseline capacitance and conductance of the sample so that at each temperature, the capacitance meter is operating in the most sensitive range. Johanson precision variable capacitors and cermet variable resistors are used to insure the accuracy of each impedance. The capacitors were calibrated with a Boonton 73-4A precision decade capacitor. The filling pulse, supplied by a Hewlett Packard HP8112A pulse generator, is applied to the sample via a pulse transformer for short pulses (<1 ms) or for longer pulses, a notch filter network that passes the test signal and blocks the 1 MHz component of the pulse. In our setup, the sample was held in an APD Cryogenics Inc. closed cycle helium cryostat and the temperature is read and controlled by a Lakeshore cryotronics model 330 temperature controller. The capacitance meter, pulse generator, and temperature controller are setup by means of a National Instruments general purpose interface bus (GPIB) interface. No data is read from the GPIB bus except for the temperature and the baseline capacitance and/or conductance which are read while cooling the sample and before each initial filling pulse is triggered at each preselected measurement temperature. The former is used to accurately record the baseline capacitance and conductance for trap concentration calculations and for use in some of the transient analysis routines. The latter is used for zeroing the baseline impedances as will be described in Sec. III D. All the transient signals, both capacitance and conductance, are read by the DT2821-F board via a preamplifier. The measurement progresses as follows.

A menu driven program was written to input all the essential hardware and measurement options. Using the GPIB interface, the instruments are then setup to implement the chosen options. The Boonton 7200 is set to auto range and the differential capacitance/conductance relay board is set to 0 pF/0 μS. In order to cancel out the stray capacitance of the test fixture, the sample probe is lifted slightly above the surface of the sample and the auto zero command is issued to the Boonton 7200. After the meter zero routine is finished, contact is made to the sample and the sample is cooled to approximately 5° below the start temperature and held there for 5 min to allow the cryostat to stabilize and the autotuning temperature controller to adjust its proportional integral differential (PID) control parameters automatically. During the cooling process, the capacitance/conductance verses temperature curve under reverse bias is recorded in order to supply the transient routines with an accurate accounting of the baseline capacitance/conductance. The pulse width is then measured as described in Secs. III B and VI C. The sample is then warmed to the selected start temperature where the transient acquisition begins. The capacitance and/or conductance is measured across the GPIB bus and the appropriate capacitor and resistor is switched onto the meters differential terminals. This cancels the baseline capacitance down to 0.0625 pF and the conductance down to 0.125 μS. The capacitance and/or conductance is again measured, but this time by the analog-to-digital (A/D) converters of the DT2821-F and the appropriate voltage is output through the two digital-to-analog converters (DACs) of the DT2821-F in order to further cancel out the remaining dc component in the signal. By using the preamplifiers to cancel out the remaining dc component no capacitance value less than 0.125 pF need be used on the differential capacitance board. The trigger signal is then initiated over the GPIB bus and when the trigger signal reaches the external trigger input of the DT2821-F, the acquisition begins. As the data buffers (described in the next section) fill, the ATLAB device driver releases them to the program for processing. Thus, directly after each single transient is acquired, the data averaging starts. Only the data points of interest are averaged, increasing the speed dramatically. After the data is averaged, another trigger is issued and this process is continued until the preselected number of transients at each temperature are acquired. The average temperature, temperature standard deviation, baseline capacitance (before the first pulse), and the averaged transient data are stored. The transient is then displayed, and the conventional DLTS spectrum is calculated, displayed, and stored. This process is repeated at preselected temperatures until the final temperature is reached.

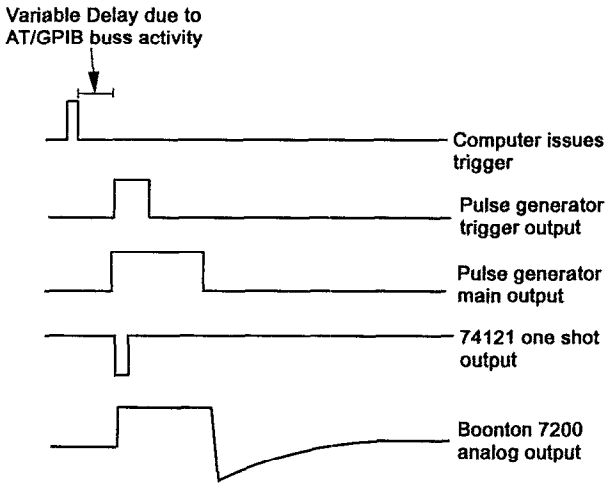

FIG. 2. Timing diagram showing the reduced timing complexity due to software oriented triggering.

B. Timing

One advantage of our system is that the normal complicated timing circuitry has been reduced to one ¢ 60 (TTL) chip. The timing of our system is summarized in Fig. 2 and is described below. The computer issues a trigger command to the pulse generator through the GPIB bus. Since the timing of the acquisition is crucial and many factors such as GPIB bus activity, and even the AT computer bus activity effect the starting time of the data acquisition, there must exist an external trigger circuit to trigger the DT2821-F. This was implemented by a 74121 one shot setup to issue a 1 μs pulse driven by the trigger output of the pulse generator. This one shot insures that the trigger pulse sent to the DT2821-F is of proper polarity and uniform width. This is necessary because in triggered mode, the HP8112A has a trigger pulse width that is a function of the output pulse width and is always longer than the maximum external trigger pulse width allowed by the DT2821-F circuitry. Since the falling edge of the 74121 corresponds to the beginning of the output pulse within a negligible propagation delay the data acquisition is started at the beginning of the injection pulse. However, the transient event we wish to observe begins at the end of the injection pulse. This is easily accounted for by storing this time period of the signal in a different memory buffer than that of the transient signal of interest. This is also an easy way to account for the relative inaccuracy in the pulse width of the HP8112A (5%). This is done by initially sampling the filling pulse on one of the spare inputs of the DT2821-F. The pulse width can then be accurately measured and accounted for. This is further explained in the software section that follows.

C. Merits of the Boonton 7200

The use of the Boonton 7200 capacitance meter as opposed to previous models like the 72A, 72B, and 72BD brings many improvements to our system even for the conventional DLTS measurement. In some devices, such as CdS (Ref. 28) and high resistivity Si and Ge,[26] the series conductance of the sample is not negligible. Thus, the conventional capacitance DLTS can not be properly employed. Conductance DLTS may provide an answer to such problems[26] but to this point has not been thoroughly explored. Moreover, the Boonton 7200 preamplifier has been modified such that it no longer suffers from the slow recovery from an overload condition. In older models, this tuned amplifier required internal modifications such as the addition of a relay to short out the preamplifier input during the filling pulse.[29] Due to the slow nature of such components, the performance was severely limited. This new meter requires no internal modifications, and thus is easier to use, and greatly exceeds the performance of the older models. Additionally, the noise introduced in the 72BD due to the digital switching as reported by Chapell and Ransom[29] has been eliminated by proper shielding of the analog and digital circuitry. Since DLTS requires measurements in a remote test fixture that introduces stray capacitances, the auto zero feature of the 7200 is very convenient.

Finally, the response time and overload recovery time of the 7200 is superior to previous models. This allows a smaller delay from the end of the filling pulse to the beginning of the data acquisition which improves the detection limit for shallow traps and traps with low concentrations. When the filling pulse is small enough to keep the capacitance within the 2 pF range, the total time for full recovery (settling to within 0.1% of its final value) is simply equal to the instruments 50 μs response time [see Fig. 3(a)]. If the pulse drives the capacitance out of the 2 pF range more time is required to recover from the overload. Under these conditions, the time required to fully recover is 175 μs as shown in Fig. 3(b). When the pulse drives the sample into forward bias, the amplifier clamps and requires 500 μs to recover fully [see Fig. 3(c)]. This nonideal behavior when pulsing into forward bias has been observed by other researchers.[30] These recovery times are significantly shorter than the 1 ms recovery times of the modified older models.[29,31]

D. Preamplifier

The capacitance and conductance signals are passed through a preamplifier (consisting of OP-27 opamps) that subtracts out the remaining dc offset and amplifies the signal by a factor of 5. This makes full use of the entire ± 10 V input range of the DT2821-F. Additional gain factors of 1, 2, 4, and 8 are software selectable on the DT2821-F to further enhance the resolution of the system. The dc offset is measured and the appropriate offset voltage is output by the DT2821-F DAC that "cancels" out the dc component in the signal that is passed to the ADC converters. In this way, the full dynamic range of the converter is used to record the transient signal and is not "wasted" on the dc component without the need for ac coupling the input. This avoids the possible waveform distortion that could occur by ac coupling. All of the transient fitting routines would interpret the waveform distortion as the existence of

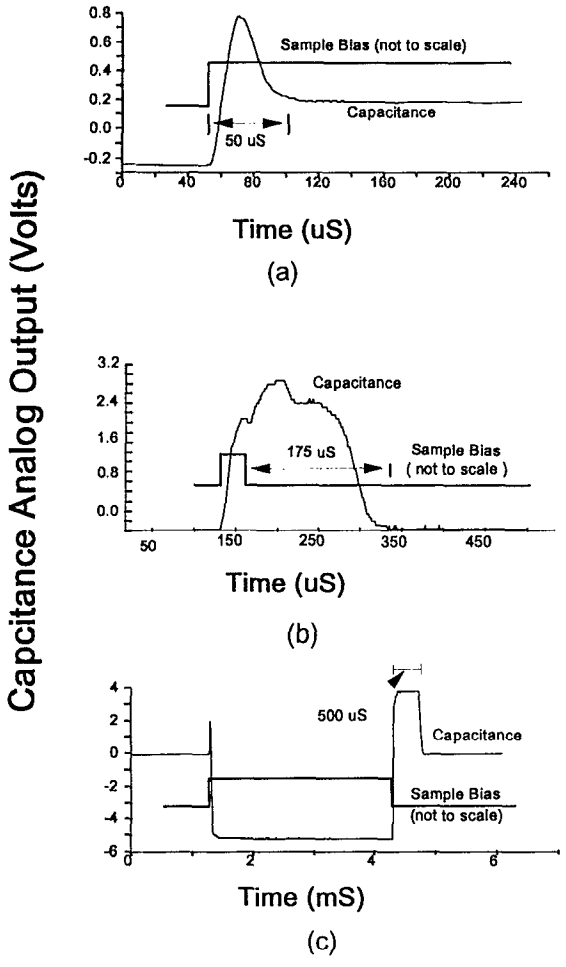

FIG. 3. Boonton 7200 response times. (a) 50 μs when pulse is of small enough magnitude as to keep the capacitance within the 2 pF capacitance range, (b) 175 μs when pulse outranges the capacitance, and (c) 500 μs when pulse forward biases the sample.

an additional transient in the decay and thus, introduce errors in the transient analysis.

Finally, the DT2821-F board digitizes the signals and stores the data directly in the computer memory. The software then extracts the relevant data points in each transient and averages them with successive transients in floating point arithmetic in order to increase the signal to noise ratio. The data is then stored, displayed, and then the conventional DLTS spectrum is calculated by numerically integrating the product of the transient with a selected weighting function (boxcar, rectangular/lockin or Miller's correlator[32]). The conventional DLTS spectrum is then stored, and displayed. Since the data being manipulated is only the 768 points not the 131 072 points, the processing of the data is almost instantaneous. The temperature is then incremented to the next temperature and the process is repeated. The system is software intensive and can perform a variety of different measurements simply by changing the control programs. A description of the software that controls the system is given in the following section.

IV. SOFTWARE DESCRIPTION

Most of the real work to acquire the DLTS transients is done in the software. All the software was written in C language since the use of pointers is crucial. Since disc operating system (DOS) programs are limited to 640 kbytes of memory and a substantial part of this memory is devoted to buffers allocated for temporarily storing the transients (256 kbytes), the code was written to conserve memory. The interface to the DT2821-F board was performed using the ATLAB device driver package supplied by Data Translation. The software developed for this system is broken down into four sections: measurement/hardware setup, memory allocation, acquisition control, and post measurement analysis (either transient or conventional).

A. Measurement and hardware setup

The measurement and hardware setup portion is a menu driven routine that selects the parameters that are to be implemented. Some of the many parameters, particularly those related to the timing, allow the user to trade off measurement time for number of substrates to be sampled. The measurement/hardware related parameters include which channels to sample (capacitance and/or conductance), the base acquisition rate, gain of the system, the base buffer size, and the subset size. The base rate can be chosen up to 150 kHz and is adjusted to the closest allowable frequency available on the DT2821-F board. This adjusted rate is defined as $f=4$ MHz/N, where N is an integer greater than 25. If both the capacitance and conductance are sampled simultaneously, the actual sample rate for each transient is half of the adjusted base rate. Total system gains of 5, 10, 20, or 40 can be selected from the menu. These high gains when combined with signal averaging can increase the resolution in terms of trap concentration significantly. The measurement time can be controlled by adjusting the base sample rate or the base buffer size. By choosing a base buffer size less than the nominal value of 131 072, the measurement time can be decreased at the expense of the total number of rates stored for each transient. Additional hardware/measurement options that can be selected include the reverse bias, the filling pulse voltage and width, starting and ending temperatures, temperature increment, capacitance test signal level (15, 30, 50, or 100 mV), the number of averages at each temperature, and the delay time defined as the time period after the end of the filling pulse before data acquisition begins. The GPIB buss is used to setup the instruments with the selected parameters.

B. Memory allocation

Certain blocks of memory must be allocated for use with the DT2821-F board. These blocks will hold the data samples converted by the DT2821-F and thus cannot be used by the program while acquisition is taking place. However, unlike the case when the transients are converted by a remote digitizer, the data is available for program use directly after the DT2821-F board finishes sampling.

Four data buffers are dynamically allocated by the measurement control program. The first is an integer buffer of an appropriate length to contain the entire pulse and delay time. This length is given by

buffer size 1 = adjusted rate

$$\times (1.06 \times \text{pulse time} + \text{delay time}), \quad (3)$$

where the factor of 1.06 is included to insure that even if the pulse width is longer than the maximum inaccuracy of the HP8112A (5%), the pulse buffer is long enough to hold the entire pulse, hence allowing the pulse width to be measured. The buffer is then allocated and the pulse is sampled in order to accurately measure its width. The exact pulse width is then determined by finding the maximum negative derivative of the sampled pulse. This corrected pulse time is then used in Eq. (3) but with the 1.06 factor replaced by 1.0, to reallocate the first buffer to the correct size. The data sampled into this buffer is not stored and is only used for timing purposes.

The second integer buffer is the base buffer mentioned above and contains the actual transient data points we are interested in. This buffer length is selected in the measurement/hardware setup portion of the code but is always a power of 2 and is less than or equal to 131 072 points. This requires up to 256 kbytes of memory and since the DT2821-F cannot transfer data across a physical 128 kbyte memory page, must be broken up into smaller buffers that straddle the memory page boundaries. Each smaller buffer is then added to the DT2821-F buffer transfer list and are filled sequentially. A larger base buffer size is possible if the buffers are allocated in expanded memory but this requires a DMA transfer to lower memory in order for the program to access the data. Since this bottlenecks the acquisition in the same way that the remote digitizer does, this procedure was deemed unnecessary.

TABLE I. Summary of results from a commercially available system, literature references [see also Song *et al.* (Ref. 34) and Su *et al.* (Ref. 36)] and our new system for the analysis methods discussed in Sec. V.

12 MeV electron irradiated $p+$ n diode				
Trap designation DLTS System/ analysis	T_1 E_1 (eV)	T_2 E_2 (eV)	T_3/T_4 E_3 (eV)	T_5 E_4 (eV)
	σ_1 (cm^2)	σ_2 (cm^2)	σ_3 (cm^2)	σ_4 (cm^2)
SPC Elect./ rectangular	0.392	0.335	0.243	0.174
	3.8×10^{-16}	3.2×10^{-16}	6.9×10^{-15}	1.4×10^{-14}
Miller *et al.*[a]	0.39	0.36	0.23	0.18
	4.0×10^{-15}	...	2.0×10^{-16}	1.0×10^{-14}
Di Zitti *et al.*[b]	0.40	0.36	0.195	0.170
	1.0×10^{-17}	1.0×10^{-16}	1.0×10^{-16}	2.2×10^{-14}
Chantre and Kimerling[c]	0.44	...	0.21	0.17
	...	...	2.0×10^{-16} 0.23 4.0×10^{-16}	...
Present sys./ rectangular	0.409	0.355	0.252	0.166
	1.3×10^{-15}	8.2×10^{-16}	1.6×10^{-14}	7.4×10^{-15}
Present sys./boxcar	0.414	0.354	0.256	0.164
	1.4×10^{-15}	8.0×10^{-16}	2.2×10^{-14}	5.9×10^{-15}
Present sys./linear predictive modeling	0.419 1.6×10^{-15}	0.352 8.0×10^{-16}	0.205 7.0×10^{-16} 0.249 5.3×10^{-15}	0.167 7.5×10^{-15}
Present sys./ rectangular long weighting times	0.409 1.1×10^{-15}	0.352 7.5×10^{-16}	0.202 9.3×10^{-16} 0.258 2.3×10^{-14}	0.172 1.8×10^{-15}

[a]Reference 2.
[b]Reference 33.
[c]Reference 35.

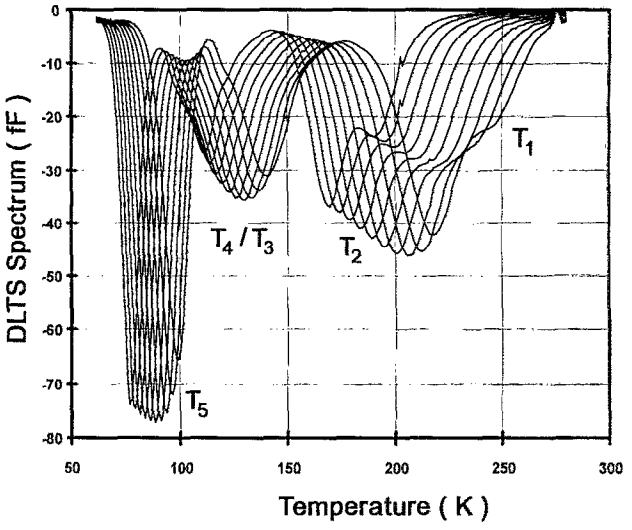

FIG. 4. Ten rectangular weighting function DLTS spectrums taken in one thermal scan. The rightmost spectrum is for a weighting time of 2.56 ms while the leftmost spectrum is for 1.31072 s. Note that there is no resolution of the T_3/T_4 peaks around 107 K.

The third buffer allocated is a floating point buffer used for averaging the transients. By using floating point arithmetic averaging, the "effective" resolution is enhanced beyond the 12 bit digitization of the A/D converter. As mentioned before, to use floating point averaging in conjunction with a remote digitizer, one would have to download each transient from the digitizer requiring an enormous amount of time.

The fourth buffer is an array of pointers that references the desired memory locations within the base buffer without having to loop through the entire buffer. In this way, the averaging is done only over the points to be extracted, greatly increasing the speed of the averaging process. When sampling only the capacitance or only the conductance, if L is the number of points in the subset (128 in the above example) then the first L points in the pointer array "point" to the first L points in the base buffer. The next $L/2$ points in the pointer array point to $L/2$ points occurring every other point in the base buffer starting at the $L+1$th point. This continues until the data structure described in Sec. II is established. When you wish to sample both capacitance and conductance at the same time, the pointer array is similar except that it accounts for every other data point in the base buffer corresponding to a conductance value and the intermittent points being capacitance values. The two transients are separated into and averaged in different arrays.

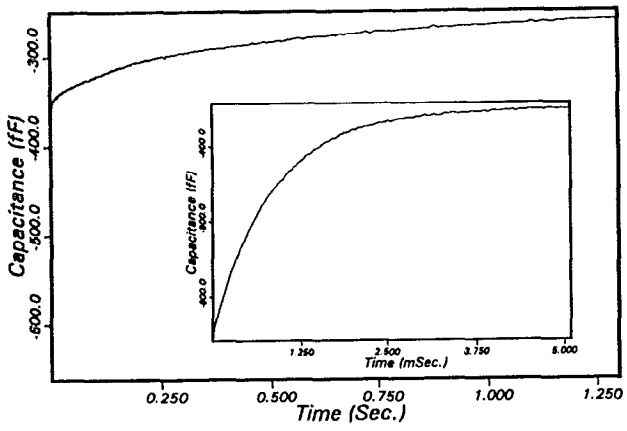

FIG. 5. Typical transient with three exponential components, one fast and two slow. The insert is an exploded view of the same transient showing the smooth nature of the faster component due to the pseudo-logarithmic sample spacing.

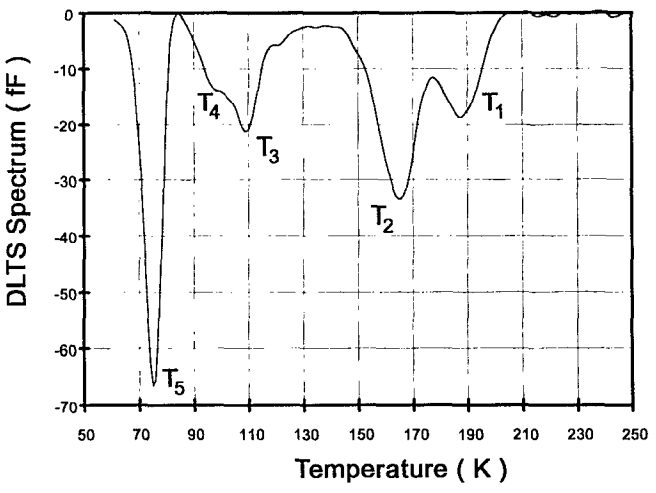

FIG. 6. Boxcar DLTS spectrum ($t_1/t_2=1.31072/0.65536$ s/s) showing a partially resolved shoulder eluding to the presence of more than one level for the T_3/T_4 peaks.

V. ANALYSIS

The analysis can be done by either conventional DLTS spectrum methods or transient methods. Conventional DLTS support is included to facilitate fast analysis when the time consuming transient analysis is not required. Boxcar, rectangular ("lockin style"), and the Miller correlator[32] weighting functions are supported. The methods of transient analysis currently being used are nonlinear least squares,[14] modulation functions,[24] and the correlation method of linear predictive modeling.[23]

Three types of analysis were performed on a 12 MeV electron irradiated $p+n$ silicon diode in order to demonstrate the capabilities of our system. Table I summarizes a comparison of the parameters for the traps obtained by our system using the different methods. These data are compared with the data acquired from a commercially available SPC Electronics Inc. system as well as several traps reported in the literature for similarly prepared samples. All the methods agree well for traps designated as T_1, T_2, and T_5. However, most techniques cannot distinguish a difference between the levels marked T_3 and T_4. As will be demonstrated shortly, the correlation method of linear predictive modeling technique was able to resolve these levels while conventional DLTS methods failed.

Figure 4 shows the unsmoothed, rectangular weighting function conventional DLTS spectrum of a $p+n$ electron irradiated diode. Note that all ten weighting functions were constructed from one thermal cycle and that the T_3/T_4 peak around 107 K appears to be only one peak. Figure 5 shows a typical capacitance transient at 101.25 K. This transient contains three exponential components corresponding to the 0.17, 0.20, and 0.25 eV traps. In the insert of Fig. 5, the same transient is observed over a shorter time period. Note the smooth nature of the transient even in the shorter time period due to the pseudo-logarithmic spacing of the sample points. Without this pseudo-logarithmic time point spacing, the linewidth in the conventional DLTS spectrum in Fig. 4 due to errors in the numerical integration would be considerably larger. Additionally, the limited time range available for the transient analysis would limit the accuracy of the system.

Figure 6 shows a counterpart boxcar DLTS spectrum ($t_2/t_1=1.31072/0.65536$ s/s) obtained from the same transients used to construct the rectangular DLTS spectrum in Fig. 4. This spectrum clearly shows a partially resolved shoulder around the peak at 107°, thus hinting to the existence of at least two energy levels. In the rectangular weighting function analysis, this was not even resolvable and the analysis yielded just a single erroneous trap at 0.252 eV. Though this boxcar spectrum (only at this unusually long weighting time) indicated the presence of this extra level, it could not resolve the peaks well enough to obtain accurate energy levels.

The correlation method of linear predictive modeling analysis, which does not require prior knowledge of the number of transient components, was used to enhance these subtle spectral features and resolve the closely spaced peaks at 107 K. The Arrhenius plot for the correlation method of linear predictive modeling is shown in Fig. 7. The peaks that were unresolvable by the above methods were clearly resolvable by this transient analysis. The T_3 energy was found to be $E_c-E_t=0.202$ eV while the T_4 level was $E_c-E_t=0.258$ eV. It should be noted that in the narrow temperature range between the 125 and 143 K (x axis values from 7 to 8 in Fig. 7) the correlation method of linear predictive modeling resulted in erroneous and unphysical results ($E_c-E_t=0.399$ eV and $\sigma=1\times10^{-8}$ cm^{-2}). Note that these points have been intentionally eliminated from Fig. 9 for clarity. This may at first appear to be a limitation of the technique, but in fact this is easily defeated by simple physical reasoning. Indeed, all results from any automated analysis should always be interpreted in terms of it's physical likelihood and even in our commercially available automated DLTS system, such unphysical results sometimes are observed.

Additionally, for the transient analysis, as many as 63 temperature observations as opposed to ten in Fig. 7, con-

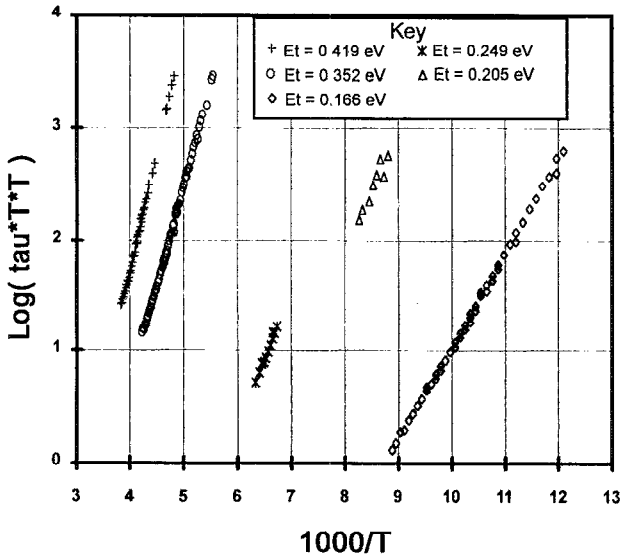

FIG. 7. Arrhenius plot from the correlation method of linear predictive modeling. Note the resolution of the T_3/T_4 traps as well as the high accuracy of the T_1, T_2, and T_5 traps due to the number of data points contributing to the Arrhenius plots.

tribute to the Arrhenius plot of single trap, hence making the knowledge of the energy position of the stronger traps extremely reliable. The transients were modeled 18 times per substrate and 11 substrates per transient for a total of 198 models per temperature. This extremely large number of models per temperature was needed in order to resolve the finer spectral anomalies centered around the peak at 107° in Fig. 6 but is normally not needed. The total automated analysis time was 8 h for temperatures from 60 to 280 K and was performed over night. The analysis time for eight models per substrate and 11 substrates only takes 45 min and easily resolves the 0.41 and 0.35 eV traps but not the T_3 and T_4 levels. It is estimated that optimization of the root finding routines used in the analysis could result in

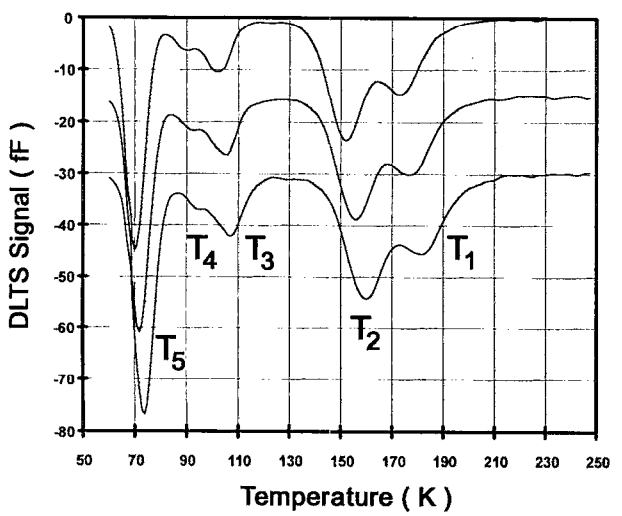

FIG. 8. Three rectangular DLTS spectrums from a very slow scan showing resolution of the T_3/T_4 peaks. The top spectrum is for a weighting time of 20.97152 s, the middle is 10.48576 s and the lower spectrum is for 5.24288 s. Each spectrum has been offset by 15 fF for clarity.

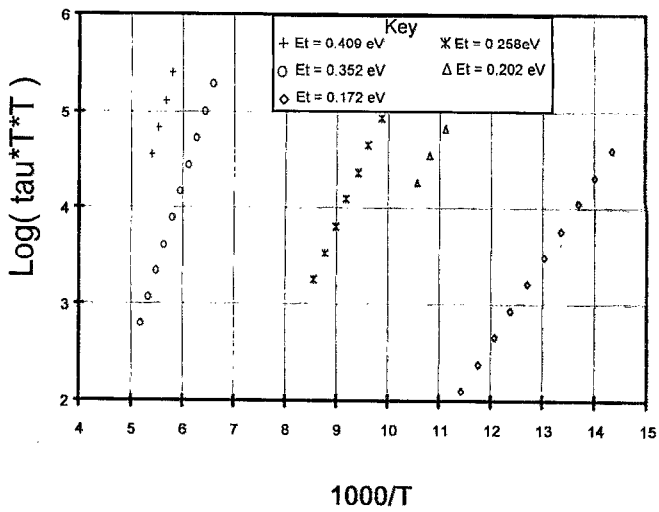

FIG. 9. Arrhenius plots for the slow scan DLTS spectrums shown in Fig. 8. Note the excellent agreement with Fig. 7.

a 50% decrease in the analysis time. Considerable time savings can also be obtained by limiting the complicated analysis to the smaller region of temperature where the peaks are unresolved.

It is well known that closely spaced traps can be separated by performing conventional DLTS scans using long weighting functions. Such slow scans require an enormous amount of measurement time. For example, if the same signal to noise ratio (obtained by averaging 100 transients per temperature) found in the scans above were desired for a slow scan weighting time of 20.97 s a measurement time of 128 h would be required. Consequently, such slow scans suffer from the incorporation of more noise than faster scans due to lower signal averaging. This slow scan approach was used in order to offer confirmation of the accuracy of the linear predictive transient analysis technique used in this system. This scan was performed at extremely long weighting times up to 20.97 s by changing the sampling rate to 6.25 kHz. Three of these spectrums are shown in Fig. 8. Due to the large scan time involved, it was unreasonable (but not impossible) to average enough transients to allow use of the boxcar analysis technique due to the above noise restrictions. Only nine transients were averaged per temperature. Thus, we opted to use the rectangular weighting function analysis. The Arrhenius plots obtained are shown in Fig. 9 and the defect energy levels summarized in Table I confirm the validity of the correlation method of linear predictive modeling. It was found that the energy levels for the slow scan DLTS agreed to within 10 meV of those levels found in the transient analysis while the transient analysis required considerably less total measurement time.

It should be noted that even though no other rapid method of analysis was able to discern the difference in the levels around this peak at 107°, the correlation method of linear predictive modeling when combined with the extended range of observation time allowable by our system permitted the resolution of this lower concentration trap. Additionally, the 0.20 ($\pm$5 meV) and 0.25 ($\pm$8 meV) eV

levels obtained match well with the literature values[2,33–36] for electron irradiated silicon under different annealing conditions. The 0.20 eV level is most likely one configuration of the metastable defect in electron irradiated silicon[34–36] previously reported as 0.21 eV. The 0.25 eV level may be a different configuration of the same defect (0.26 eV)[34] or some other unknown defect. Other commercial systems and even some conventional analysis techniques in our system, found an inaccurate energy level between the two actual energy values. The systems used by other researchers may have suffered from the same type of resolution errors. This suggests that the reported defect level at 0.23 eV (Refs. 34–36) may result from a combination of the two defect levels at 0.20 and 0.25 eV. More work is needed to resolve this issue. This example clearly shows the advantages of the transient analysis and the extended observation times implemented in our system.

Future work will involve demonstrating the systems ability to minimize the truncation errors associated with fourier transient analysis techniques and implementing the current system to explore electronic defects at resolutions in energy yet to be obtained by previous systems.

[1] D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).
[2] G. L. Miller, D. V. Lang, and L. C. Kimerling, in *Annual Reviews of Material Science*, edited by Robert A. Huggins (Materials Research Society, Pittsburgh, PA, 1977), p. 337.
[3] E. K. Evangelou, A. D. Horevas, G. E. Giakoumakis, and N. G. Alexandropoulos, Solid State Commun. **80**, 247 (1991).
[4] K. Holzlein, G. Pensl, M. Schulz, and P. Stolz, Rev. Sci. Instrum. **57**, 1373 (1986).
[5] H. S. Woon, H. S. Tan, and S. C. Ng, IEEE Trans. Instrum. Meas. **37**, 86 (1988).
[6] K. Asada and T. Sugano, Rev. Sci. Instrum. **53**, 1001 (1982).
[7] H. G. Maguire and A. Marshall, IEEE Trans. Instrum. Meas. IM **35**, 313 (1986).
[8] T. R. Jervis, W. M. Teter, T. Cole, and D. Dunlavy, Rev. Sci. Instrum. **53**, 1160 (1982).
[9] M. D. Jack, R. C. Pack, and J. Henriksen, IEEE Trans. Electron. Devices ED **27**, 2226 (1980).
[10] E. E. Wagner, D. Hiller, and D. E. Mars, Rev. Sci. Instrum. **51**, 1205 (1980).
[11] C. Y. Chang, W. C. Hsu, C. M. Uang, Y. K. Fang, and W. C. Liu, IEEE Trans. Instrum. Meas. IM **33**, 259 (1984).
[12] R. Langfeld, Appl. Phys. A **44**, 107 (1987).
[13] Z. Su and J. W. Farmer, J. Appl. Phys. **68**, 4068 (1990).
[14] T. R. Hanak, R. K. Ahrenkiel, and M. L. Timmons, J. Appl. Phys. **67**, 4126 (1990).
[15] M. Henini, B. Tuck, and C. J. Paull, J. Phys. E **18**, 926 (1985).
[16] J. Morimoto, M. Fudamoto, K. Tahira, T. Kida, S. Kato, and T. Miyakawa, Jpn. J. Appl. Phys. **26**, 1634 (1987).
[17] J. Morimoto, M. Fudamoto, S. Tashiro, M. Arai, T. Miyakawa, and R. H. Bube, Jpn. J. Appl. Phys. **27**, 2256 (1988).
[18] J. Moritomo, T. Ikda, and T. Miyakawa, Mater. Res. Symp. Proc. **69**, 343 (1986).
[19] S. Weiss and R. Kassing, Solid State Electron. **31**, 1733 (1988).
[20] M. Okuyama, H. Takakura, and Y. Hamakawa, Solid State Electron. **26**, 689 (1983).
[21] P. D. Kirchner, W. J. Schaff, G. N. Maracas, L. F. Eastman, T. I. Chappell, and C. M. Ransom, J. Appl. Phys. **52**, 6462 (1981).
[22] K. Ikossi-Anastasiou and K. P. Roenker, J. Appl. Phys. **61**, 182 (1987).
[23] F. R. Shapiro, S. D. Senturia, and D. Adler, J. Appl. Phys. **55**, 3453 (1984).
[24] C. M. Ransom, T. I. Chappell, J. L. Freeouf, and P. D. Kirchner, Mater. Res. Symp. Proc. **69**, 337 (1986).
[25] J. Morimoto, T. Kida, Y. Miki, and T. Miyakawa, Appl. Phys. A **39**, 197 (1986).
[26] N. Fourches, Appl. Phys. Lett. **58**, 364 (1991).
[27] H. K. Kim, T. E. Schlesinger, and A. G. Milnes, J. Electron. Mater. **17**, 187 (1988).
[28] L. C. Isett, J. Appl. Phys. **56**, 3508 (1984).
[29] T. I. Chappell and C. M. Ransom, Rev. Sci. Instrum. **55**, 200 (1984).
[30] D. V. Lang, in *Thermally stimulated relaxation in solids*, edited by P. Braunlich (Springer, New York, 1979), p. 93.
[31] *Instruction Handbook for wafer analyzer 17D060* (SPC Electronics Corp., Chofu City, Tokyo, Japan, 1985).
[32] G. L. Miller, J. V. Ramirez, and D. A. H. Robinson, J. Appl. Phys. **46**, 2638 (1975).
[33] E. Di Zitti, G. M. Bisio, P. G. Fuochi, B. V. Passerini, and M. Zambelli, J. Appl. Phys. **66**, 1199 (1989).
[34] L. W. Song, B. W. Benson, and G. D. Watkins, Phys. Rev. B **33**, 1452 (1986).
[35] A. Chantre and L. C. Kimerling, Appl. Phys. Lett. **48**, 1000 (1986).
[36] Z. Su, P. G. Wald, and J. W. Farmer, J. Appl. Phys. **67**, 4249 (1990).