NIST Technical Note 1496

A Next Generation Sampling Comparator Probe For The NIST Sampling Waveform Analyzer

O. B. Laug
D. I. Bergman
T. M. Souders
B. C. Waltrip

NIST National Institute of Standards and Technology • U.S. Department of Commerce

NIST Technical Note 1496

A Next Generation Sampling Comparator Probe For The NIST Sampling Waveform Analyzer

O. B. Laug
D. I. Bergman
T. M. Souders
B. C. Waltrip

Electronics and Electrical Engineering Laboratory

June 2008

U.S. Department of Commerce
Carlos M. Gutierrez, Secretary

National Institute of Standards and Technology
James M. Turner, Acting Director

National Institute of Standards and Technical Note 1496
Natl. Inst. Stand. Technol. Tech. Note 1496, 52 pages (June 2008)
CODEN: NSPUE2

Table of Contents

List of Figures

List of Tables

A Next Generation Sampling Comparator Probe for the NIST Sampling Waveform Analyzer

O. B. Laug, D. I. Bergman, T. M. Souders, B. C. Waltrip

Quantum Electrical Metrology Division,
National Institute of Standards and Technology, Gaithersburg, MD 20899

ABSTRACT

The design and performance of a Next Generation Sampling Comparator Probe (NGSCP) is described. It is intended as one of a group of probes designed for use with the NIST Sampling Waveform Analyzer (SWA). The probe design is centered on an application specific integrated circuit (ASIC) analog comparator featuring a bandwidth of 6 GHz. The design considerations of the ASIC analog comparator and the probe are discussed. The probe's performance features are compared against a previously designed ASIC probe. In addition to the design aspects of the probe, a thermal error correction technique is described which shows how high-speed settling performance can be enhanced. This technical note provides complete schematic diagrams of the ASIC comparator and the probe.

KEY WORDS: comparator, sampling comparator, latching comparator, waveform sampler, sampling comparator probe, thermal correction, pulse settling

1. Introduction

Over a span of several years, the staff in the Applied Electrical Metrology Group of the Quantum Electrical Metrology Division (formally the Electricity Division) has pursued the technique of waveform sampling for the measurement of voltage. More specifically, the focus has been on a sampling comparator system that utilizes a single comparator which functions as a 1-bit analog-to-digital converter (ADC) sampling in equivalent-time. The work has led to an in-house development of a wide utility voltage measurement system called a Sampling Waveform Analyzer (SWA). The SWA is being employed as an alternative to thermal converters for the measurement of RMS voltage. In addition, it is routinely used to measure properties of time-domain signals such as pulse amplitude and settling. The performance of the SWA for measuring pulse settling is unmatched by any other type of sampling and digitizing system.

The elegance of the system centers on a single latching analog comparator which forms the heart of the measurement system. While the performance of other parts of the system, such as the timebase and digital to analog converter (DAC), do affect the accuracy, it is the comparator that ultimately determines the accuracy, bandwidth, and noise of the measurement. Several probes containing custom designed integrated circuit comparators as well as discrete designs have been specifically designed for this application and are currently used in support of NIST calibration services. The first NIST designed probe, developed in the early 1990s, utilized an application-specific integrated circuit (ASIC) latching comparator in a probe that we currently refer to as the "ASIC Probe." This document describes a next generation sampling comparator probe (NGSCP) in which an integral part is a new (ASIC) latching comparator featuring a combination of lower noise and wider bandwidth than the previous design. In addition to the design aspects of the NGSCP, this document provides selected performance measurements of the NGSCP that are presented with the first ASIC Probe for comparison purposes. All performance measurements are made with the NIST designed SWA mainframe. A detailed discussion is presented on the issues of thermal errors and techniques to lessen their effects. Data on step settling performance is provided from an ensemble of measurements from various reference pulse sources.

2. Background

As a review, the basic sampling comparator system is presented in Fig. 1. The conversion process is implemented in equivalent-time by means of a latching comparator probe that provides the sampling function as well as the decision function. As shown in Fig. 1, the latching comparator probe is operated in a feedback loop to repetitively sample the input signal at a given instant on the waveform. When the comparator is strobed, it compares the instantaneous value of the signal waveform applied at one input to a reference voltage applied at the second input. Based on the output decision of the comparator, the successive approximation register (SAR) instructs the DAC to increment or decrement the level applied to the reference input of the comparator in preparation for the next comparison. The process is repeated until the reference level reaches the

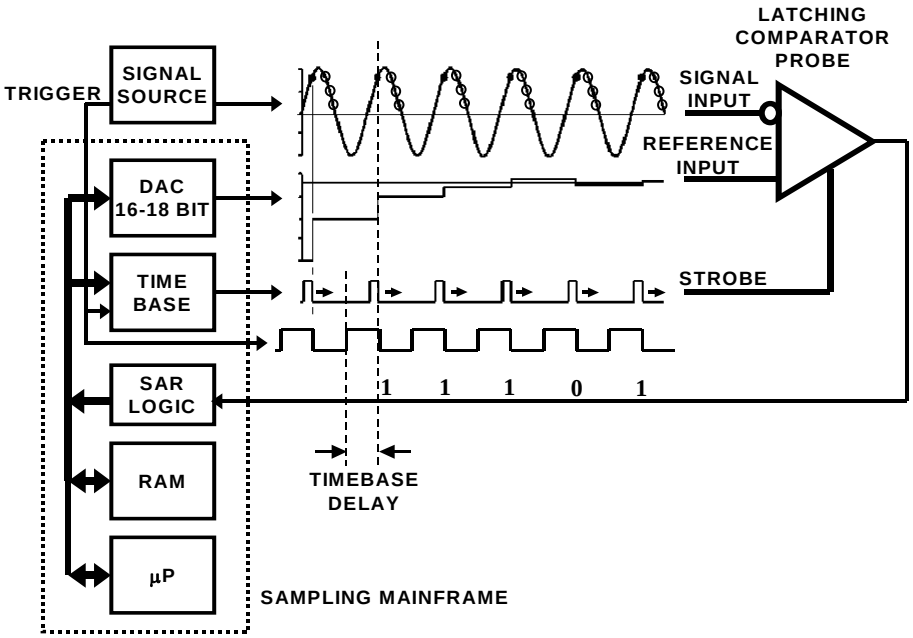


Fig. 1. A basic sampling comparator system with a successive approximation search

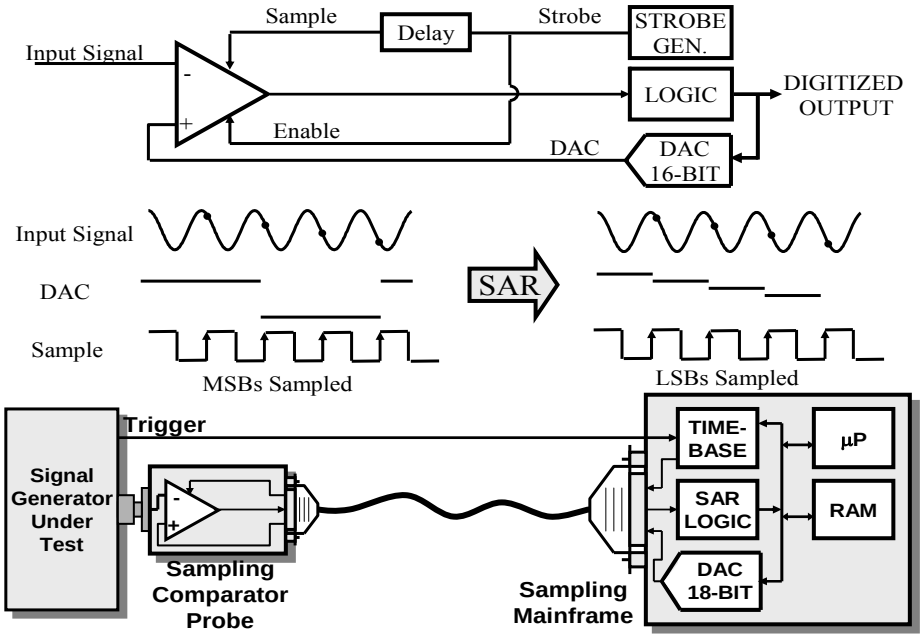


Fig. 2 Equivalent-time, successive approximation digitization. Comparator probe connects to sampling mainframe through umbilical harness.

sampled level of the waveform to the desired level of precision. After each conversion is completed, a timebase increments the strobe timing delay to another point on the waveform, and the successive approximation routine is repeated. The sampling process is continued until a record of samples has been obtained. A more complete description of the sampling system including averaging techniques can be found in [1, 2, 3, and 4].

The actual implementation of the block diagram of Fig. 1 is configured such that the DAC, timebase, SAR logic, memory and processor are contained in a mainframe enclosure as shown in Fig. 2. The latching comparator with various support circuitries is located in a probe head connected to the mainframe by an umbilical cable which allows the input signal connector on the probe head to be brought directly to the measurement plane of interest. This is an important concept because no intervening signal cables with their attendant losses are required between the source under test and the input to the sampling comparator. Fig. 3 shows a photograph of the NIST SWA mainframe together with a selection of probe heads used for various applications. The Next Generation Sampling Comparator Probe shown to the left is the primary subject of this report.

Fig. 3. Picture of SWA system showing mainframe unit, low-noise sampling probe with 100 V attenuator attached (foreground), ASIC sampling probe (nearest to mainframe), and the Next Generation Sampling Comparator Probe (left).

A simplified schematic diagram of a typical latching or sampling comparator shown in block form of Figs. 1 and 2 is shown in Fig. 4. A "front-end" differential stage is followed by a cross-coupled latch stage, level shifter, and logic output drivers. While in the "track" mode the differential stage remains energized by the tail current until the instant the track/latch (strobe) signal arrives. As the switch moves from the track to the latch position, the tail current is rapidly steered from the differential stage to the latch stage. It is during this important time instant that the comparator is operating as a sampler.

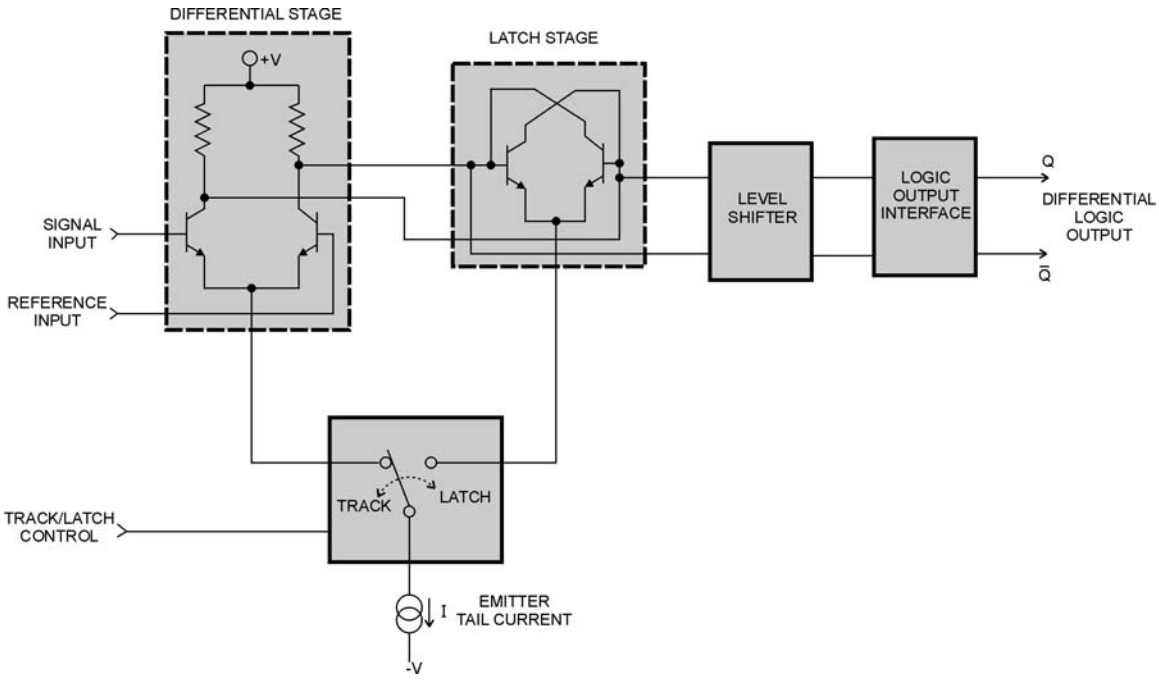


Fig. 4. Basic block diagram of analog latching comparator.

Before discussing the design particulars of comparators for this application it is well to provide a bit of background as to why the basic analog latching comparator as outlined in Fig. 4 works so well in the sampling application. The earliest known application of comparators for analyzing high-speed waveforms was dubbed a "sampling voltage tracker" whereby a comparator was enclosed in an analog integrator loop [5]. Integrated circuit manufacturers have traditionally produced latching comparators for the primary purpose of latching or holding the amplified difference between a time varying signal and a fixed reference. The latched or held decision of the comparator is in turn processed or stored as needed before it is unlatched to track the difference. A latching comparator usually employs a differential front-end stage designed to provide enough amplification to produce a clear and unambiguous indication of the difference. In some applications the comparator is used in the "transparent" mode where the difference signal simply passes through the latch to the output. An example of such an application would be to "square up" or increase the slewing rate of a sinusoidal signal for triggering purposes. Typical latched comparator specifications such as *minimum hold time*, *setup time*,

overload recovery, and *propagation delay*s are important specifications used to judge the speed and quality of a latching comparator. The words strobe and latch tend to be used interchangeably by many comparator manufacturers since the process of strobing and latching are essentially occurring at the same instant. However, in the context of sampling the term strobe in this document denotes the sampling or capturing process and latching means the holding of the final decision.

What was perhaps not always fully appreciated was that these types of latching comparators could be applied as a sampler and digitizer to achieve very high sensitivity without having any significant "front-end" gain. Researchers realized that the latching process produced by positive feedback causes a regenerative reaction initiated by a small difference voltage presented to the latch stage. This phenomenon is somewhat like that of regenerative detectors used in early radio receivers in which the detector stage is set on the verge of oscillation. Under this condition during the exponential build-up of oscillation, a detector stage can easily achieve a gain of over a million resulting in very high sensitivity.

One way of visualizing the process is to liken it to a handoff or relay of the instantaneous differential output of the input stage that is rapidly passed to the latch stage as the latter stage is being powered down. Again refer to Fig. 4. As the tail current to the differential stage begins to diminish its gain starts to drop while the increasing tail current to the latch stage begins to increase its gain. This is a well known effect of bipolar transistors stages where the transconductance or gain is directly proportional to the bias current. As the current to the latch stage increases, the gain will eventually exceed unity causing the stage to regenerate on its own. The latch stage will continue regeneration in the direction set by the differential stage output condition from the moment regeneration begins. The exponential regeneration process if allowed to grow will eventually reach the bounds of the stage. The final bounds of the latch stage (interpreted as a 1 or 0) provide a one bit indication of the difference between the reference voltage and the signal at the strobing instant. Thus, high resolution performance can be achieved using regenerative amplification without input preamplification. Depending on the speed of the devices that compose a latch stage, the exponential regeneration time constant is relatively fast. The smaller the difference voltage presented to the latch at the strobing instant the longer the regeneration process will take to reach the final bound. In other words, time is being exchanged for sensitivity. A general rule is that each decade increase in sensitivity requires about one regeneration time constant in delay for the latch stage to reach its final bound. Because of the equivalent time application, the time required for this exponential growth is of no consequence and is always completed well before the next sample is taken. This whole transfer process described here is referred to as the "aperture time." Aperture time affects the bandwidth and distortion of the sampling process. It is influenced by the transition speed of the strobe signal to the comparator, the speed of the current steering switches, and the regeneration time of the latch stage. The shorter the aperture time the greater the bandwidth and the lower the distortion [6].

One commercial manufacturer markets such a general-purpose waveform analyzer based on the sampling comparator approach [7]. This instrument utilized an off-the-shelf

commercial integrated circuit latching comparator. The performance is impressive but, it suffers primarily from thermal tail errors in the comparator. While the manufacturer has recognized the thermal problem and provided a software correction algorithm for pulse settling, it does not address the fundamental problem. Moreover, thermal tail errors in the time domain cause errors in the frequency domain which are not corrected by their algorithm. Aside from the thermal tail error problem, most general purpose commercial comparators are not optimized for the sampling application.

2.1 The Thermal Error Problem

Thermal tail error is a time dependent offset voltage that results from signal-dependent differential heating in the front-end differential transistor pair of the comparator. It manifests itself in the time domain by exhibiting an exponential settling tail after a pulse transition requiring up to microseconds to settle to the new level. Thermal errors not only result in poor pulse settling but also translate into errors in the frequency domain.

The problem of signal-induced thermal error can be understood most easily by the effect this error has on the sampler's response to a pulse or step-like signal. The problem is described as follows: For waveform samples taken immediately following a step transition, the DAC reference value under SAR control is nearly equal to the waveform final value when the lesser significant bits are being decided. In order to allow maximum time for the DAC to settle, the DAC value required for the next comparator decision is programmed as soon as the comparator decision for the present bit is available. Fig. 5 illustrates the time-sequential process. For step signals, this means that during the period before the waveform transition, a differential voltage equal to the step signal amplitude is present at the comparator inputs shown in Fig. 5 (top graph). During this period, the two transistors comprising the comparator input differential pair heat unevenly, and a voltage offset error between the two develops. When the step transition occurs, differential power dissipation between the two transistors instantly becomes zero, but the time for thermal equilibrium to be reestablished can be hundreds of nanoseconds or longer. As the thermally induced offset voltage returns to zero, its superposition onto the true signal is sampled by the latch, and an error in the sampled data is manifested as a long exponential tail (Fig. 5, bottom graph).

2.2 Enabling Technique

A unique technique of minimizing thermal tail errors has been designed into all of the NIST comparators for sampling applications. It is referred to as "enabling" whereby the differential and latch stage are maintained in a deenergized state until just prior to the sampling instant [8].

The enabling process begins by keeping the tail current I, of Fig. 4, in the "off" state while the switch is in the "normal" track state. Under this beginning condition both stages (differential and latch) stages are in a deenergized state. Enabling then is invoked by a separate control pulse that turns on the tail current. This allows the differential stage to become activated and begin to track the difference between the signal and the

reference input. After a short interval, a separate strobe control signal arrives that switches the tail current to the latch stage. Thus, the time period between the onset of the tail current to the strobe time is referred to as the "enable time." The enabling process consists of activating the front-end differential stage for a short time period much less than the thermal time constant of the devices. By keeping the "on" time interval short neither junction has time to heat up from imbalanced input signals thereby maintaining both junctions of the differential stage at nearly the same temperature and thus eliminating time dependent V_{be} offset potentials. However, the enabling time must be long enough for the stage to electrically recover from the enabling transient and accurately track the difference signal before the latch stage is energized. Generally, the electrical time constant of the transistors which make up the stage tends to be much faster than the thermal time constant. Studies of pulse settling performance have shown that enabling periods as short as 2 ns effectively eliminate settling errors. Early studies of the enabling process were demonstrated by simply switching on the negative supply voltage to the comparator just before strobing. The effects of enabling related to the general topic of thermal errors turn out to be rather profound and will be discussed in more detail in a later section on modeling and correction of residual thermal error.

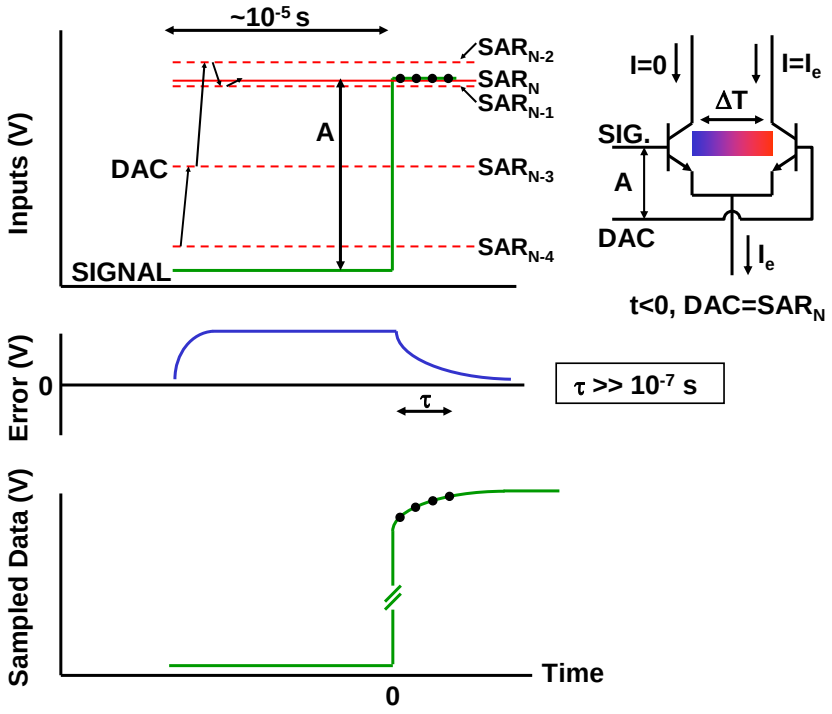


Fig. 5. Successive approximation digitization process with step-like input signal and DAC voltage sequence (top). Differential transistor pair offset error voltage arising from differential heating (middle). Samples taken after step transition show thermal tail error (bottom).

3. Next Generation ASIC Comparator Design Considerations

The first ASIC comparator designed at NIST for the sampling comparator application utilized the QuickChip™ 6 integrated circuit array design method. This scheme was devised at Tektronix Inc. to shorten the time and reduce the cost of designing custom integrated circuits. It uses a fixed IC array of transistors, resistors, and several other types of components, which are connected with a custom metal pattern that defines a specific functional circuit. This first ASIC comparator utilized a recessed-oxide-isolated wafer fabrication process optimized for analog signal design. It features NPN bipolar silicon transistors with a transition frequency (f_T) up to 8.5 GHz. The first NIST-designed ASIC comparator chip, when incorporated into the "ASIC Probe" and controlled by the SWA mainframe, resulted in an equivalent-time -3 dB bandwidth of nearly 2.5 GHz.

The success of the first ASIC comparator prompted the consideration of another custom comparator design with the goal of greater bandwidth. The GST-2 QuickCustom Wafer Fabrication Process offered by Maxim Integrated Products was selected [9]. It represents an upgrade of the previously used Tektronix silicon foundry process which later was acquired by Maxim. GST-2 is a high-speed, self-aligned, double-polysilicon, bipolar process with 3 layers of metal for device interconnect. It features trench isolation for minimum device spacing resulting in the following features:

- NPN $f_T \approx$ 27 GHz @ (V_{CE} = 4 V, I_c = 2 mA)
- Min. transistor emitter area = (0.8 μm X 1.6 μm)
- Min transistor area = (5.4 X 8.1) μm
- 170 Ω/square and 600Ω/square poly silicon resistors
- 1fF/μm^2 MOS capacitors
- 3 layers of gold interconnect (min. pitch = 2.7 μm)

The above process is available on a QuickChip 9-60D, general purpose die containing 36 input/output (I/O) bonding pads, with dimensions of 1.9 mm X 1.78 mm. The die contains over 3,000 devices arrayed into different tiles organized and optimized for various analog or digital designs. Six general purpose core tiles are laid out across the die each including 3 sizes of NPN transistors, lateral PNP transistors, and 400 Ω, 4 kΩ, and 10 kΩ resistors. Schottky diodes as well as capacitors are available. These general purpose tiles are considered the work horse of the GST-2 QuickChip IC design. The polysilicon resistors can be connected in various serial and parallel combinations to make the desired values. All of the resistors have large temperature coefficients and modest absolute tolerance but match and track others of the same implant type. Custom NiCr resistors are also available which offer lower temperature coefficients and allow for custom absolute values.

The success of any design relies on accurate device simulation models. With the GST-2 process Maxim provided an Analog Design System (ADS) which uses "TekSpice," a nonlinear time-domain circuit simulator. Their SPICE simulation models were determined through proprietary closed-form extraction routines. The data were obtained

from direct-on-wafer measurements up to 18 GHz. All of the design work performed at NIST relied heavily on simulations, layout parasitic extraction, and built-in layout error correction software to produce a first-time success with performance very close to predictions.

Simulation studies throughout the design process involved the usual SPICE analysis such as dc, small signal, and transient time-domain analysis. The overall comparator performance was evaluated in the manner that emulated the sampling process by using time-domain analysis to allow the comparator to make a decision for each static reference level compared against a dynamic signal at the strobing instant. The successive approximation process had to be done manually because the software could not be easily modified to automate an external successive approximation loop. The SAR process was done after each decision by incrementing or decrementing the reference level and running a new simulation. While this seems laborious, because prior information was usually available for the waveform under test, the number of simulation iterations could be considerably reduced to gain the specific information needed. For instance, in determining the 10 % to 90 % transition duration of the comparator, the reference level was set to 90 % of the input test step amplitude. By knowing the approximate initial value in time, only a few simulation iterations of strobe delay time were required to bracket the 90 % amplitude time point. The same process was repeated at the 10 % amplitude time point and the resulting time difference between the 10 % and 90 % time points provides the comparator's transition duration.

3.1 Differential Input Stage

The first ASIC comparator's front-end differential amplifier was designed for a nominal gain of 1. While lower front-end gain tends to maximize bandwidth and minimize input capacitance, noise can become a factor. One can qualitatively view the effect of low front-end gain by appreciating that the latch stage has to do all the work of amplifying very small differences through regeneration. During the critical beginning of the regeneration period, noise can randomly influence the direction of regeneration.

The Next Generation Comparator design is patterned after the first ASIC comparator with the primary goal to increase the -3 dB bandwidth to 5 GHz. The basic architecture of the comparator is shown in Fig. 6 which is similar to but differs in some respects from the first ASIC comparator. Not shown are the various tail current sources required to bias some of the stages. One of the main features of this design was to increase the "unsampled gain" of the front-end differential stage primarily to reduce the noise of the comparator. While attempting to achieve a maximum cutoff frequency is an important goal, one must keep in mind that most linear amplifiers can suffer from the effects of non-flatness (ripple, early roll-off, or peaking) in gain in the pass-band at frequencies well below the cutoff frequency. This in part is due to poles and zeros well beyond the cutoff frequency influencing the pass-band flatness. For instance, layout parasitics can often spoil the flatness in what might otherwise appear to be a flat response in simulation. This point must be emphasized in that there is a tradeoff between the amount of un-sampled gain that should be considered for this application and gain accuracy. Emitter

degeneration with resistance is an important feedback technique to help level and flatten the gain over the intended band. Simulation studies indicated that with the GST-2 process, an emitter-degenerated, cascode-differential stage with a gain of 4 will produce a -3 dB bandwidth of nearly 6 GHz. A system having a single pole response and –3 dB frequency of 6 GHz in theory should be attenuated only by about 0.1 % at 200 MHz, whereas simulation studies including layout parasitics indicate a response down by nearly 0.5 % at 200 MHz. This is a clear indication that there are multiple ultra high frequency poles affecting the response. While the bandwidth of the comparator is determined primarily by the bandwidth of the front-end differential stage, the previously mentioned aperture time also plays an important role in the overall sampling bandwidth. The quantitative effects of aperture time are not addressed here except to say that in simulation the small signal bandwidth of the front-end stage always related well to the overall bandwidth of the comparator when simulated in the sampling mode. These simulations confirmed that the front-end stage bandwidth is dominant in establishing the bandwidth of the comparator.

The differential front-end stage employing a cascode stage at the collectors not only improves the bandwidth and gain flatness, but it reduces the so-called Miller capacitance, as seen at the input. The attendant nonlinear effects of capacitance on the source impedance affect the harmonic distortion. Thus, reduction of the Miller capacitance plays an important role in improving the harmonic distortion of the comparator. Also, the cascode stage tends to buffer the strobing "kick out" voltage back to the signal input.

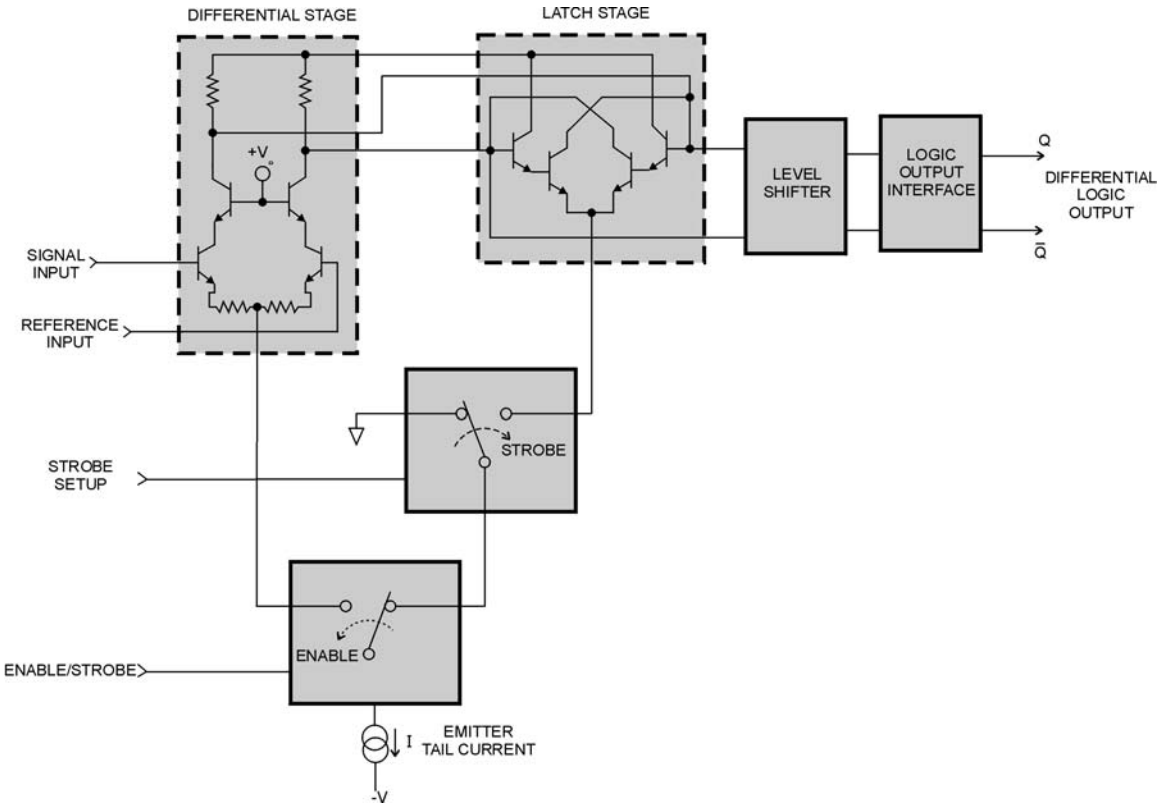


Fig. 6. Basic block diagram of next generation analog comparator with enabling.

3.2 The Latch Stage

The latch stage consists of a cross-coupled differential transistor pair with emitter followers inside the loop that serve the dual purpose of increasing the open-loop current gain of the stage and provide level offsetting to prevent reverse biasing of collector-base junctions of the cross-coupled pair. The output level of the latch is followed by a level shifting stage, interface circuit, and output drivers to produce standard ECL levels of the comparator's output decision.

3.3 The Switching Scheme

The switching scheme to implement the enabling and strobing of this comparator required a somewhat different approach than previously described in the first ASIC comparator. Because of the additional operating voltage needed for the cascode stage, the potential of the emitters in the differential stage tend to be closer to the negative rail supply voltage. This in turn uses up some of the negative voltage head room required to implement both strobe and enabling switching transistors. In other words, there would be insufficient headroom voltage for stacking two switching devices; one for the "track/latch" switch, and one for the "enable" emitter tail current switch. As a result, a slightly different switching scheme had to be implemented. Refer again to Fig. 6 and the pair of SPDT switches labeled as Enable/Strobe and Strobe/Setup. This scheme takes advantage of the fact that more operating voltage head room is available for the Strobe/Setup switch because the emitters of the latch stage are at a more positive potential than the emitters of the differential stage. The SPDT switches are always implemented with differential transistor stages that steer emitter tail current to respective nodes through the collector of the "on" transistor. Consider the initial conditions of the two switches as shown. Under this condition the emitter tail current is diverted to ground potential which keeps both the differential and latch deenergized. The Enable/Strobe signal activates the SPDT switch to the enable position which energizes the front-end differential stage. This allows the front-end stage to begin tracking the difference between a fixed input reference level and the signal. Shortly after the Enable/Strobe switch is activated the Strobe/Setup" switch is activated. This provides a path for the emitter tail current to the latch stage when the Enable/Strobe switch returns to its original state. After a short period when the latch has fully regenerated to its bounds, the Strobe/Setup control signal returns to its original state ready for another sample. The return of the Strobe/Setup" switch to its original state activates a slave latch (not a part of the comparator) to capture and hold the state of the latch before it is denergized. Fig. 7 shows a detailed timing sequence with typical timing intervals referenced to various I/O pads of the integrated circuit schematic diagram (Appendix A). The timing sequence occurs as follows.

1. A differential strobe signal from the mainframe switches the Enable/Strobe switches from the strobe position to the enable position.
2. This action allows the tail current to flow to the emitters of the front-end differential stage. The 4 mA of tail current is divided between the emitters according the balance of the stage as the stage begins to track the difference between the signal and reference voltage.

3. In the meantime, approximately 300 ps after the enable/strobe signal, the Strobe/Setup switches are activated providing a future path to the emitters of the latch stage.
4. When the Enable/Strobe signal returns to its original state, the path set up by Strobe/Setup switch now allows the emitter tail current to activate the latch stage. At this instant of time the emitter tail current in the front-end differential stage is diverted to the latch stage allowing the stage to regenerate.
5. Finally, when the control signal to the Strobe/Setup switch returns to its original state, a slave latch stage is clocked to capture the state of the comparator latch stage as its emitter current is powered down.
6. The above timing sequence constitutes one sample and is repeated for every sample. Note the very short duty cycle of the comparator which is active for about 4 ns for every sample period. The relatively short "on" period of the differential and latch stage in relation to the sampling interval creates a very low power duty factor for the comparator.

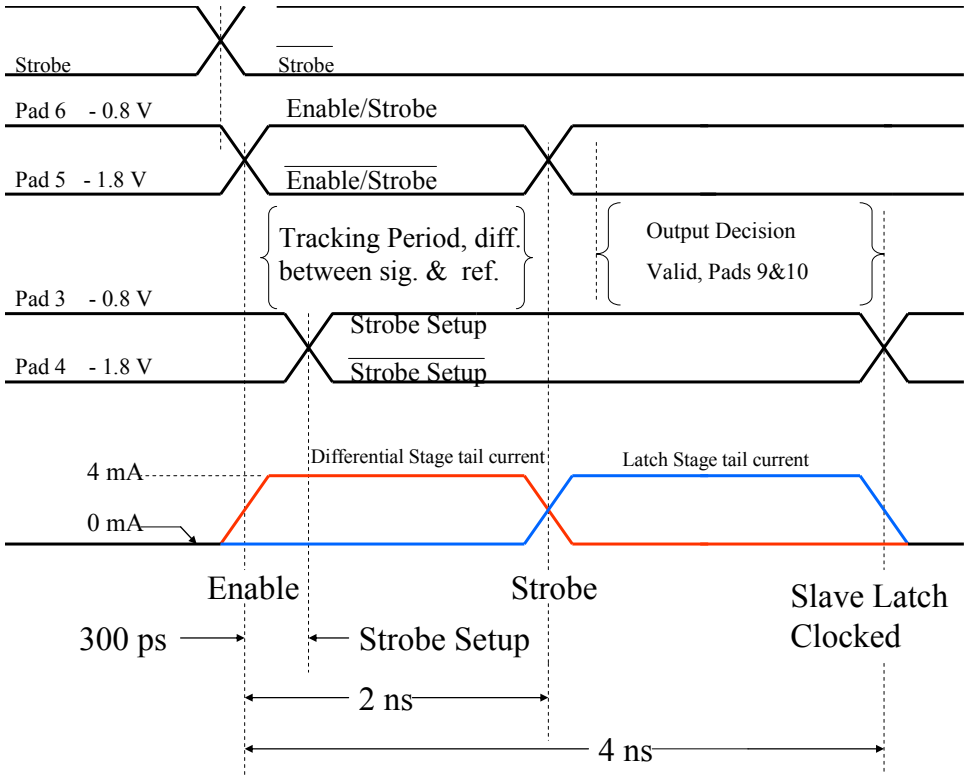


Fig. 7. Timing diagram of next generation analog comparator with enabling.

Refer to the schematic diagram in Appendix A which is a direct copy taken from the ADS 5 schematic capture editor of the MAXIM simulation program showing the details of the design including source generators, supply voltages, and external parasitic inductances for purposes of simulation. All of the resistors used were 400 Ω polysilicon resistors having a temperature coefficient of resistance of -900 μΩ/Ω per deg. C.

Numbers beside the resistor symbol indicate how many of one type of resistor are connected in parallel to obtain the value. Note that Electro-Static Discharge (ESD) protection in the form of back-to-back diodes is included at every bond pad. These diodes, in conjunction with ESD Supply Shunts cells at all supply voltage bond pads are the foundation of effective ESD protection of the die. The one exception to ESD protection is that none is provided on the reference and signal inputs. A decision was made not to include ESD protection on the analog signal input lines because additional capacitance loading coupled with off-chip bond wire inductances tends to compromise the bandwidth and gain-flatness performance. While the signal and reference inputs to the bare die are susceptible to ESD, it was determined that once the die was connected with 50 ohm termination resistors in the probe housing, the die would be reasonably protected except for inadvertent over-voltage input signals in excess of the power supply voltages.

3.4 The Probe Design

One of the more critical parts of the probe design is referred to as the reference network. The reference network serves as the interface between the reference voltage receiver amplifier and the reference input to the comparator in the probe. The reference voltage receiver is a differential amplifier that receives the voltage output from the DAC in the SWA mainframe.

Simulation studies have consistently shown the importance of maintaining the same source impedance to both inputs of the comparator particularly over the upper frequency band of the comparator. Maintaining similar source impedances over the high frequency band is essential to keep the stage balanced and help it recover quickly after the enable signal. Also, at high frequencies there is a tendency for some of the input signal to leak through to the reference input and combine with the reference level to cause an error. To overcome these problems, the reference source impedance should ideally be on the order of 25 Ω which is the Thévenin equivalent of the 50 Ω signal source and 50 Ω input termination located at the signal input connector. However, this requirement places an undue burden on the reference amplifier which would have to maintain up to 2 V into 25 ohms throughout the entire sampling interval.

Fig. 8 shows the reference network used in the first ASIC probe. Its purpose is to present low source impedance at high frequencies to the comparator reference input and to lessen the output current burden requirements of the reference voltage receiver. The network is a low-pass “T” network with a reference voltage charging time constant of about 600 ns; requiring about 6 μs for the capacitor to charge to within the resolution of the DAC. A second benefit to the low-pass network is the additional filtering effect on the noise voltage from the reference receiver. Even this seemingly simple network must be designed with care. The quality of the capacitor must be high to avoid settling errors in the reference voltage. Low grade capacitors with high equivalent series resistance are to be avoided.

While the reference network used in Fig 8 worked well for the first ASIC probe, experience showed that the same network could not be used in the NGSCP. Attempts at using the low-pass "T" network were undermined primarily by excess step response ringing. The problem was narrowed down to the parasitics of the capacitor and its self-resonance frequency within the bandwidth of the system. Experiments with high frequency, high quality capacitors, multiple paralleled capacitors, and resistive damping techniques did not alleviate the problem. The network that gives the best response is shown in Fig. 9. Here the reference input to the comparator is connected directly to a 50 Ω resistor to ground. A high frequency ferrite bead provides isolation between the reference voltage receiver/buffer and the comparator input. Any high frequency currents from the reference input of the comparator are directly shunted through the 50 Ω resistor and prevented from impinging on the reference amplifier by virtue of the ferrite bead's high impedance at high frequencies. The disadvantage to this approach is that the reference voltage receiver must be able to drive up to ± 40 mA into 50 ohms for a ± 2 V reference voltage range. A buffer amplifier inside the loop of the differential receiver amplifier provides the required current. A complete schematic diagram of the Next Generation Sampling Comparator Probe is provided in Appendix B.

The NGSC chip is housed in a probe assembly that contains companion circuitry to interface the comparator to the SWA mainframe. Fig. 10 shows the probe with the lid removed to expose the layout. The circuit is divided into two compartments. The comparator chip is mounted as a bare die on a substrate board with wire bonds from the die to the board. Bare die mounting was selected in order to minimize package parasitics and permit the comparator to be physically mounted as close to the input connector as possible. As previously mentioned the probe contains an ECL-type slave latch to capture the comparator's decision before it is shut down by the Strobe/Setup pulse returning to its normal state. This allows the data to be accessed any time during the interval between samples while the important stages of the comparator are in a deenergized state. All timing intervals and delays are established by general purpose gates, resistors, and capacitors. Timing intervals can be adjusted by changing capacitor values.

.

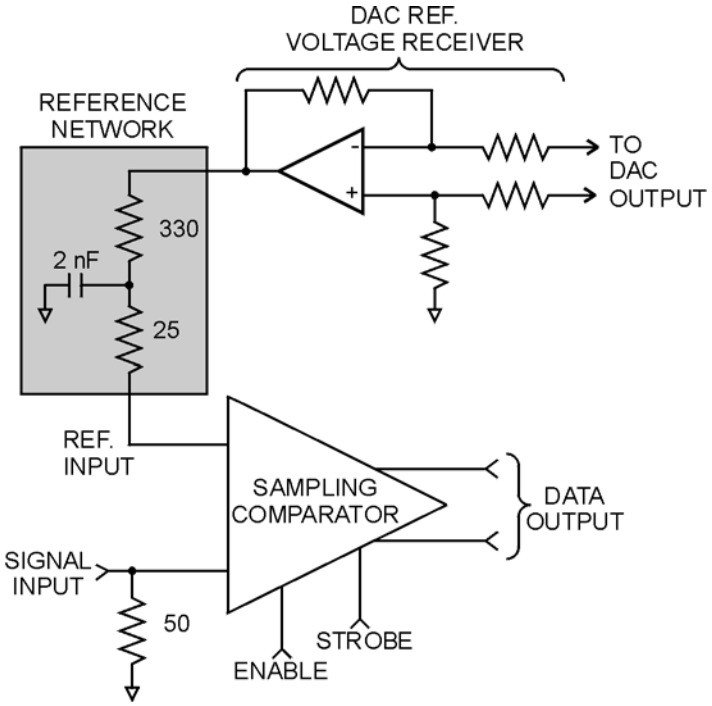


Fig. 8. ASIC Probe reference network interface

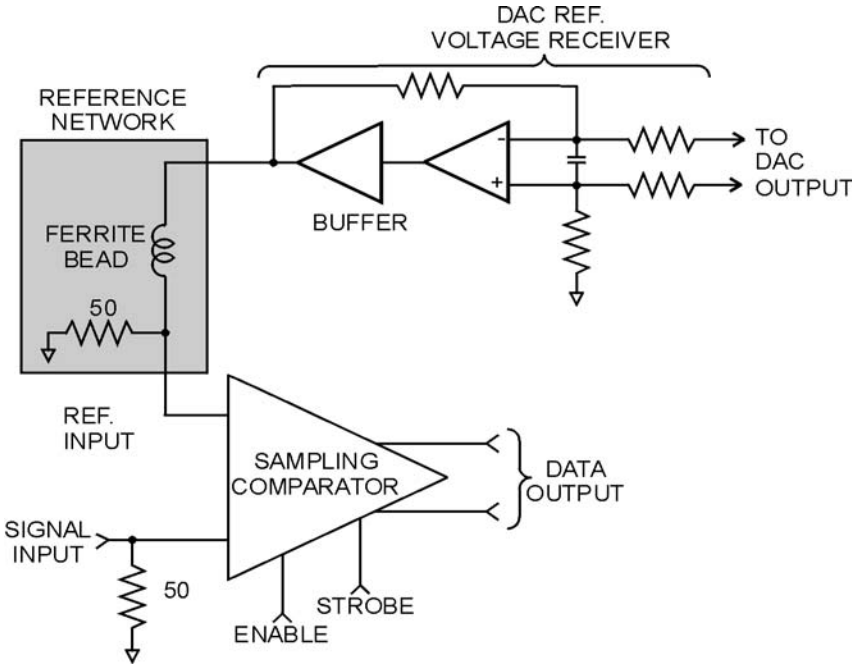


Fig. 9. NGSCP reference network interface.

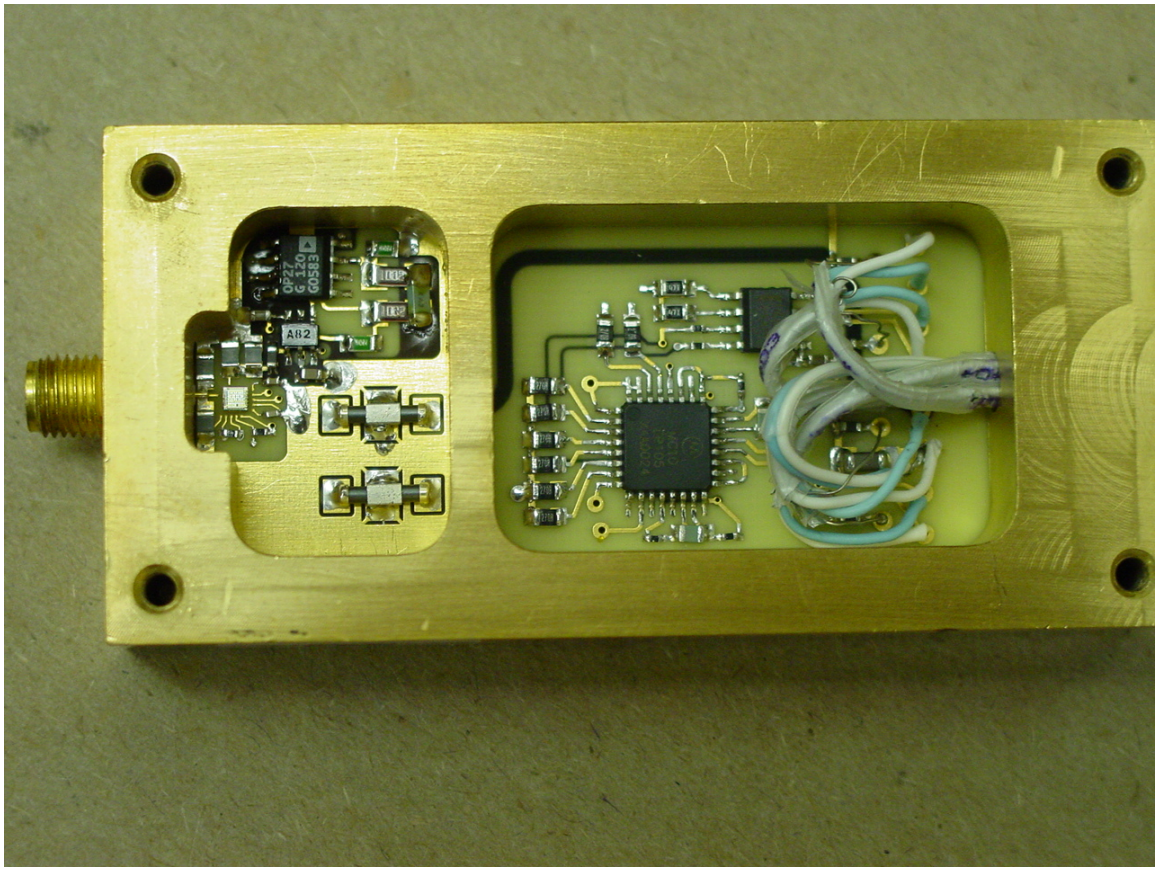


Fig. 10. Photograph of NGSCP with top cover removed.

4. Thermal Error in the NGSCP

As previously discussed, the thermal tail error problem of a comparator when used as a sampler, has been effectively overcome by the "enabling" technique. The enabling method assumes that the thermal time constant of the differential stage junctions can be modeled by a single thermal time constant from junction to ambient and that the thermal decay is effectively truncated by limiting the average power dissipation. This assumption has essentially been shown to be correct for the first ASIC comparator and other discrete comparator designs. Extensive measurements on the first ASIC comparator on pulse amplitude settling with enabling times between 2 ns and 4 ns have confirmed that any errors due to thermal effects can be virtually eliminated by the enabling technique.

In early measurements of the NGSCP however, it was observed that different settling responses occurred in the first 10 ns following a step, depending on a small range of enabling durations used. Although the probe's "enabling" feature was designed to eliminate so-called "thermal tail" error, the observed effect suggested that some residual short-term thermal tail might still exist. It was therefore postulated that residual differential heat is produced in the front-end transistor pair over the window defined by the enabling time, and this differential heating results in a temperature-dependent differential offset voltage that decays with time as the input transistors reach thermal equilibrium. The GST-2 process uses very small transistor geometries with deep-trench isolation which increases their thermal insulation and thus tends to slow the dispersion path of heat to the substrate. The thermal model is more complex involving perhaps more than one thermal time constant. The literature suggests that depending on the structure, a more accurate model should be chosen that consists of multiple and progressively larger thermal time constants modeled as lumped elements in the direction of heat propagation [10]. Fig. 11 shows one possible lumped element electrical analog thermal model consisting of two thermal impedances in series from each junction to the ambient. One is the thermal impedance from the junction to the substrate and a second is the thermal impedance from the substrate to the ambient. The dissipated power P1 is represented by a current source while temperatures are modeled as node voltages. The shorter time constant model represents the greater insulating effect of trench isolation of the transistor structure. The relatively large mass of the substrate would have a longer time constant. Also modeled in Fig. 11 is the coupling impedance through the substrate which transfers some of the heat to the adjacent transistor. The degree of coupling and any attendant time constant is unknown but it is believed that these particular high frequency transistors buried in deep wells with oxide barriers have high coupling thermal impedances to adjacent devices. The literature suggests that any significant lateral coupling to adjacent device is considered negligible [11].

In the above discussion of what appears to be a short time constant residual thermal error, the obvious solution would be to simply shorten the enabling period well below the thermal period. However, from a practical standpoint this approach is not realizable. Experience both in simulation and with experiments indicates that a very short enabling

period applied to the front-end stage does not allow the stage to recover accurately to perform the electrical difference tracking function.

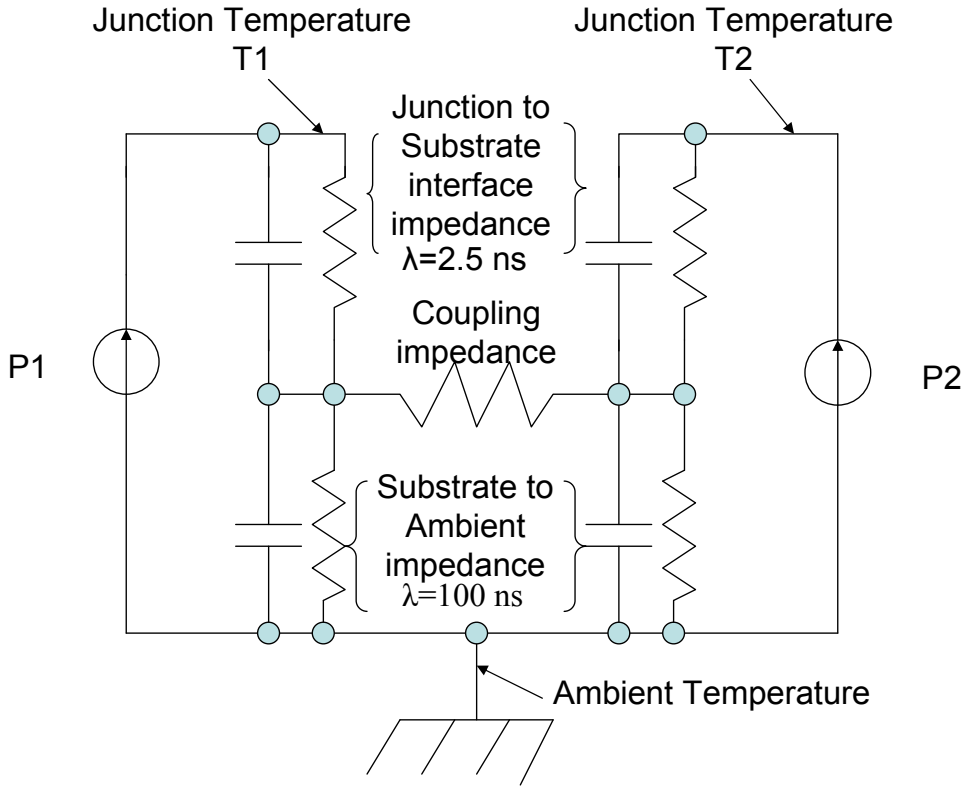


Fig. 11. Electrical equivalent model of the thermal impedances of a differential transistor pair.

4.1 Thermal Error Correction during the Enable Period

In an attempt to verify this theory and provide a method of correction, we simulated the effect of signal-dependent differential heating in the NGSCP and compared the predicted effect with the observed data. In the initial work, the power dissipation curve was assumed to be a linear function of differential input voltage, which is reasonable for small signals, e.g., signals on the order of 0.25 V or less. Using this assumption, Fig. 12 shows plots of probe step-settling response over a 10 ns epoch, for four different enabling times. Note that the divergence among the plots increases as the enabling time increases. In these and subsequent plots showing real data, the step transition occurs at 1 ns, and the data has been filtered with a Gaussian low-pass filter having a -3 dB bandwidth of 6 GHz. The input signal to the probe is provided by a Tektronix 067-1338-00 calibration step generator output that makes a transition from minus 0.25 V to 0 V in approximately

15 ps. The plotted data has been normalized to unit step amplitude, and only the settling region of the response is displayed.

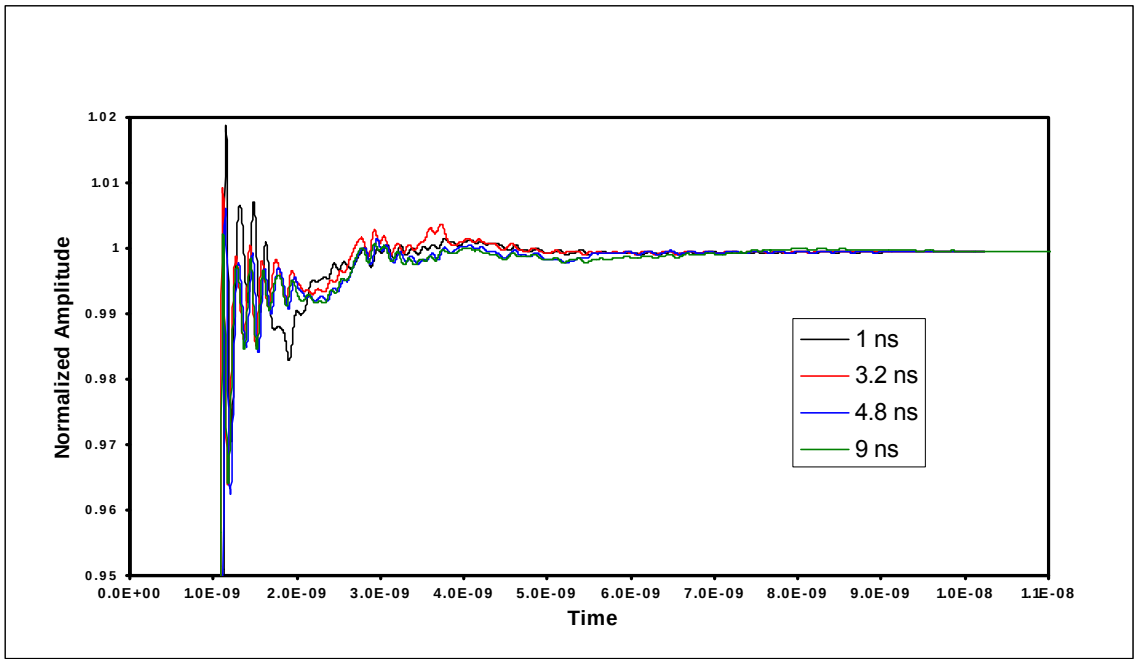


Fig. 12. NGSCP Step Waveforms with Different Enabling Times.

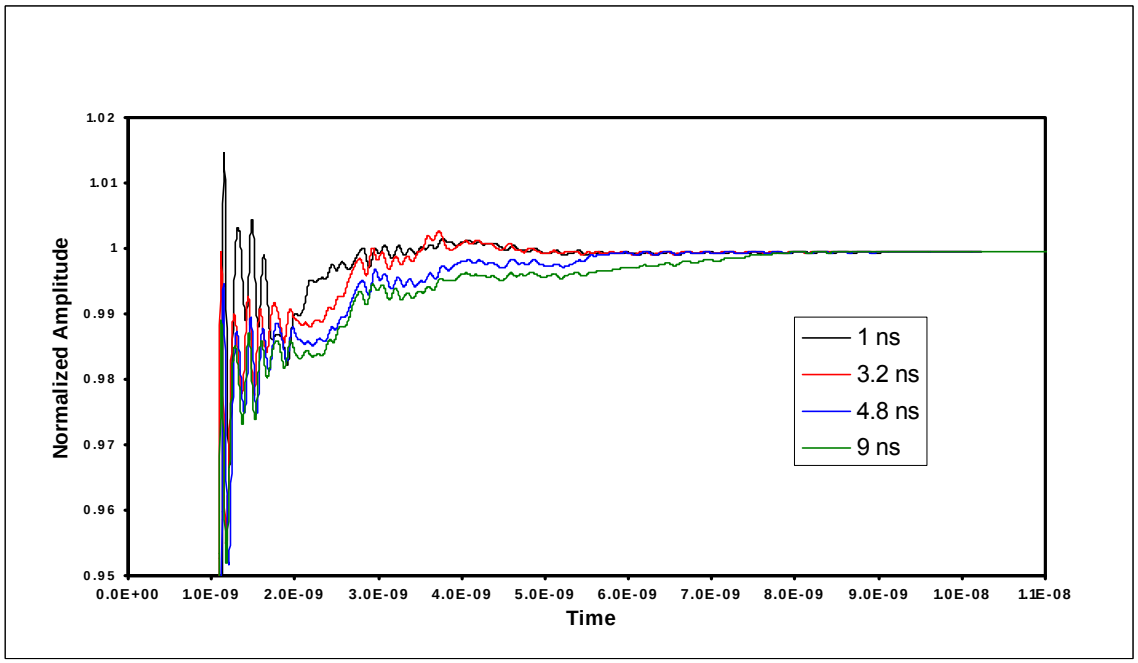


Fig. 13. Corrected NGSCP Step Waveforms with Different Enabling Times.

Fig. 13 shows the same results after correcting the data using a differential heating model. Despite aberrations in the vicinities of the enabling times (following the step transition), the divergence among the plots is now much smaller. The model uses three constants: the enabling time t_e, the time constant for thermal equilibrium τ, and a constant C that relates integrated differential power to differential offset voltage. For the results in Fig. 13, t_e is set to the actual value used and the remaining two constants were chosen to bring all four plots into the best alignment. The same values of τ and C were used for all four plots, i.e., $\tau = 2.5$ ns and $C = 1.4 \times 10^8$ V/J. Fig. 14 shows a comparison of the corrected NGSCP response to the Tektronix calibration step generator output, and the response of a 50 GHz Agilent (HP) oscilloscope to the same step. The HP data has been fitted to the probe data using least-squares with respect to time alignment, offset and gain. The fitting is done to correct for the known large, long-term settling error in the HP instrument, which causes apparent offset error immediately following the step. The greater HP "dribble up" in the last half of the record is evidence of changing settling error. Ignoring these long-term settling problems, the agreement as detailed in Fig. 14 is quite good. In Fig. 15, the corresponding response from a 50 GHz Tektronix oscilloscope has been added for comparison. The agreement between the NGSCP and the Agilent responses is in striking contrast to the 4 % divergence that occurs in the first nanosecond with respect to the Tektronix response.

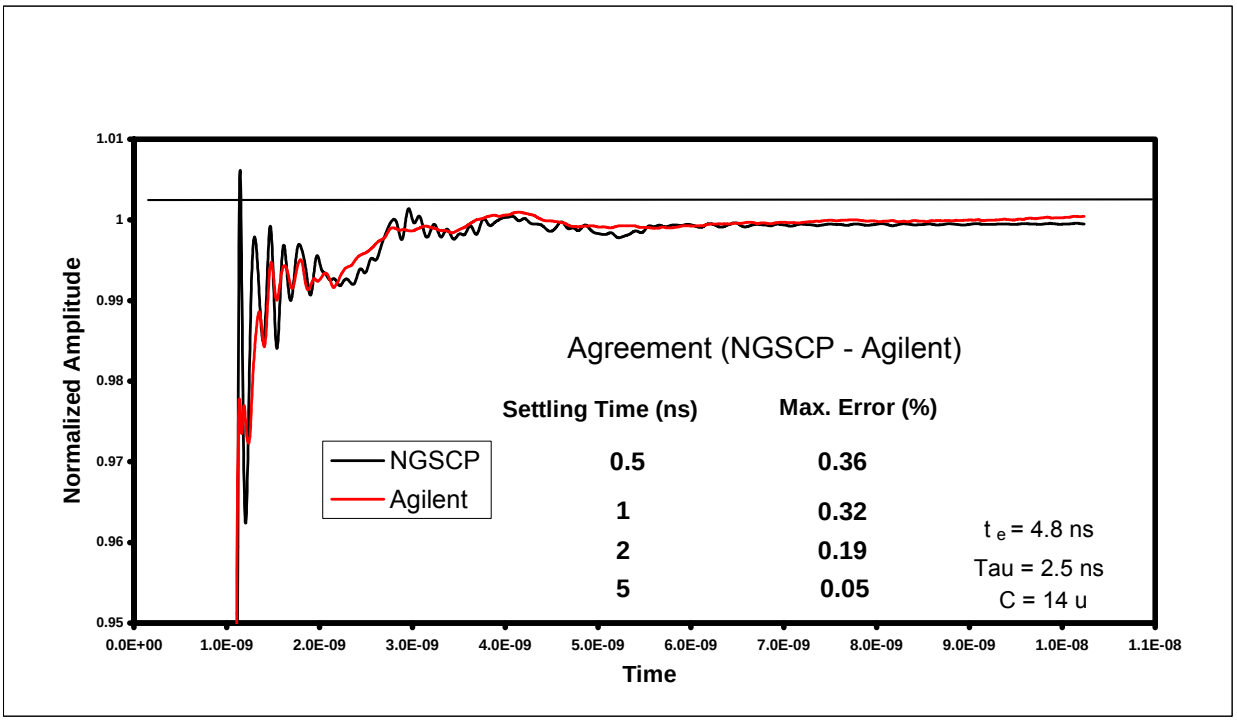


Fig. 14. Corrected Step Waveforms.

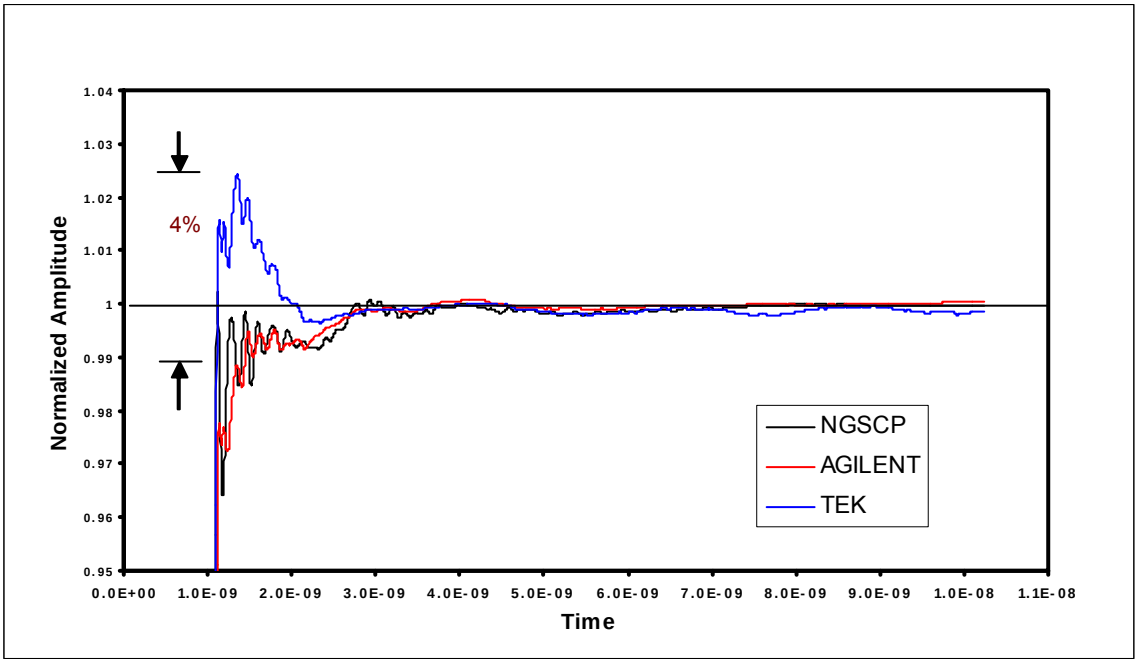


Fig. 15. Corrected Step Waveforms.

The NGSCP design also incorporates a front-end "enabling" feature as previously introduced. The front-end differential transistor pair normally remains off and is then turned on at time t_a just prior to each sampling instant. The duration that the pair is energized before the sampling instant is called the enabling time, t_e, which can be typically set in the range of 2 ns to 10 ns. Therefore, at the sampling instant, t_s, the relevant thermal history of the pair is determined only by the input and reference voltages that occurred during the preceding interval, t_e, i.e., from time t_a to t_s. In operation, the reference voltage, which is static during this interval, is set to equal the instantaneous input voltage, $V(t_s)$, at the sampling instant. This is a consequence of the equivalent-time successive approximation algorithm that drives the probe as previously discussed. If we assume that the differential temperature between the two transistors equilibrates through heat exchange that follows a simple exponential decay with a single time constant, we get the following expression for the differential offset voltage, ΔV, at the sampling instant, t_s:

$$\Delta V(t_s) = C \int_{t_s - t_e}^{t_s} \Delta P(t)\, e^{-(t_s - t)/\tau} dt \qquad (1)$$

where $\Delta P(t)$ is the differential power at time t (the power differential between the input and reference transistors with voltage $V(t)$ on the signal input and $V(t_s)$ on the reference input), τ is the thermal time constant, and C is a constant relating voltage offset to energy (integrated $\Delta P(t)$). Of the variables in (1), t_e is known by design and the function $\Delta P(t)$ is known by analysis (see below). On the other hand, there is no *a priori* knowledge of C and τ. However, from the earlier linear modeling work discussed above and presented in Figs. 12-15, C and τ were estimated from a rather extensive set of measurement data

taken at low-amplitudes, and with four different values of t_e used, yielding enough independent equations for the determination. (In that work, $\Delta P(t)$ was assumed to be directly proportional to $V(t)$- $V(t_s)$, which the subsequent analysis bears out for low amplitude signals.)

As explained above, a linear heating model was used in the first trials with results shown in Figs. 12 to15, since it is believed that the process is essentially linear for small signals, e.g., the 0.25 V step that was used in the measurements. It is realized that for larger signals the differential power becomes a nonlinear function of the signal. This can be visualized in the extreme case where the signal is large enough to completely switch all of the emitter tail current to one device while the other device is turned completely off. Subsequently, a more realistic nonlinear heating model was developed which should be applicable for any signal within the ± 2 V input signal range of the NGSCP. A numerical relationship for the differential power dissipated in the transistor pair was determined by using the MAXIM ADS simulation software and their device models to develop a parametric plot of the differential power dissipation of the two transistors as a function of the signal and reference voltage over a ±2 V range. Representative plots of the relationship are shown in black in Fig. 16, where four differential power curves (each corresponding to a specific reference voltage) are plotted as a function of input voltage. (The reference voltage corresponding to each curve is identified as the input voltage at which the differential power is zero for that curve.) Thus, as one might expect when the reference voltage equals the input voltage over the entire common-mode range the differential power is zero. The piece-wise linear approximation to these curves shown in red will be discussed later.

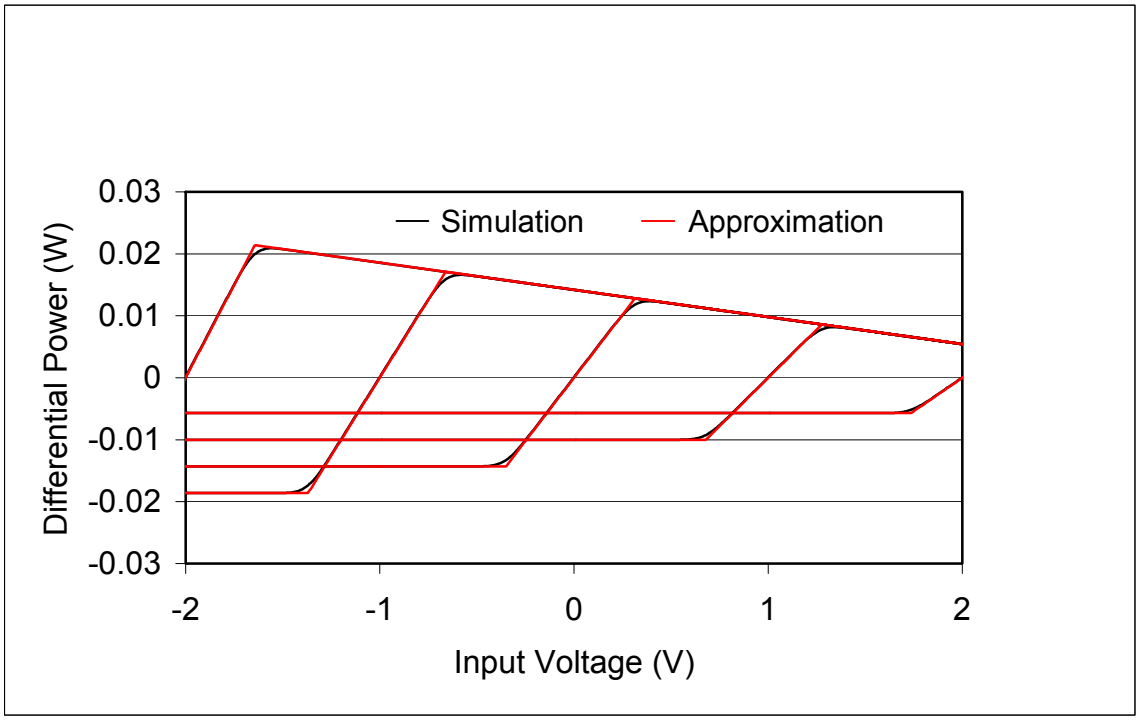


Fig. 16. Differential Power vs. Input Voltage (at Selected Reference Voltages).

The numerical relationship for $\Delta P(t)$ as a function of the signal and reference voltages as determined by simulation in Fig. 16 can be reasonably approximated by three piecewise linear segments with the appropriate values for their slopes and intercepts. Such an approximation was used in an HP VEE program to generate the plots shown in red in the same figure. As can be seen from the figure, these are quite close to the originals except for the small regions where derivative changes occur. Here, the approximation for each curve is defined by a sloping top line, a bottom line, and a connecting ramp. The top line is given by *0.0142 – 0.00438*A*, the bottom line is given by the constant *–0.0143 + 0.0043*B*, and the ramp is given by *(1 – 0.16*(B+2))*(-0.06*B + (0.06)*A)*, where *A* is the input voltage and *B* is the reference voltage. This approach was chosen in part because of its fast execution time.

The approximation just described, along with a discrete-time version of (1) was used to compute the thermal offset error at each sample point of a sampled waveform. Some notes on the computational implementation are given in Appendix C. If the thermal tail error in the sampled waveform, $V(t_{s,i})$, is small (where *i* is the sample index) and the model given by (1) is accurate, then to first order, $V(t_{s,i}) - \Delta V(t_{s,i})$ should return the true input signal for all *i.*

It is not feasible to test the accuracy of the model directly since the true input signal is not known independently with sufficient accuracy. However, we can indirectly test the validity of the model by applying it to two or more sampled (output) waveforms for which the true input signals differed only by a constant scale factor. According to the model, each output waveform will exhibit different thermal offset errors since $\Delta P(t)$ is a nonlinear function of amplitude. Consequently, the output waveforms will not just be scaled versions of each other. However, if the predicted error for each waveform is corrected by subtracting the results given by (1), then the resulting waveforms should again be appropriately scaled versions of each other.

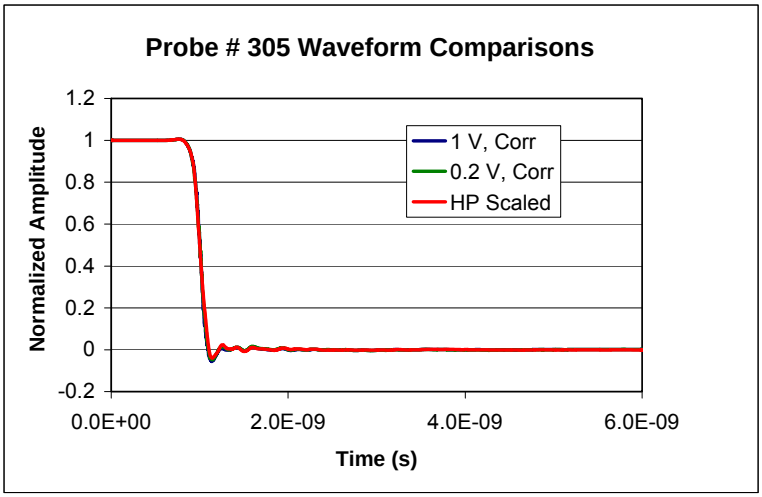
The nonlinear power dissipation model was tested using a NIST developed prototype 1 V step generator with and without a wideband 14 dB attenuator, producing output step amplitudes of 1 V and 0.2 V, respectively with nearly identical wave shapes. The transition duration of the prototype generator is approximately 145 ps which is not a factor in testing the model. The appropriate measurements were made on five NGSCP sample devices in conjunction with the NIST SWA, each of which was configured with a nominal enabling time, t_e, of 4.8 ns. In addition, for comparison purposes, independent measurement of the step waveform was made using an Agilent 50 GHz, sampling oscilloscope. Because of limitations in the Agilent sampler's full-scale-range, a wideband 10 dB attenuator was employed for this measurement.

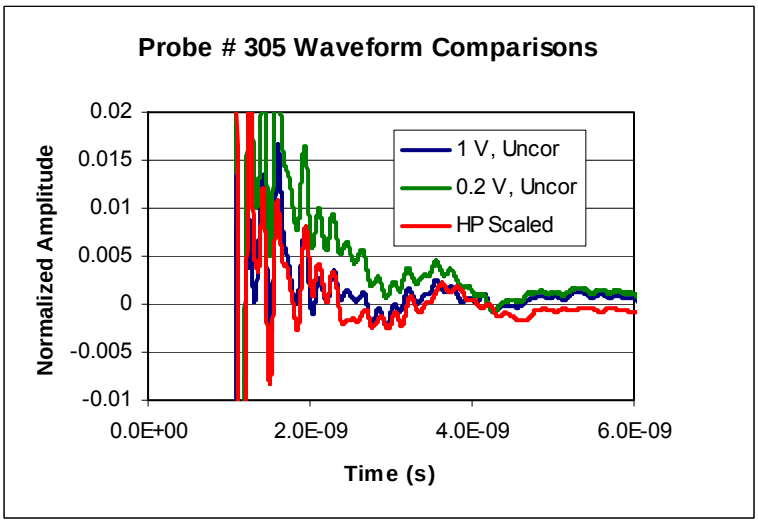
Data for all five probes have been analyzed per the method outlined above. The results show that the model-based correction improves the linearity of the tested probes by a

factor of about three with respect to RMS error in the 0.5 ns to 5 ns epoch following the transition (the time period of most concern). These results are shown graphically for the four probes tested in Figs. 17 – 21, and the RMS improvement achieved for each of the five probes is summarized in Table I.

Table I RMS improvement achieved for each of the five probes

Improvement in RMS Error Over 0.5 ns to 5 ns Interval following Transition			
Probe No.	**Uncorrected (%)**	**Corrected (%)**	**Improvement Ratio**
305	0.380	0.102	3.73
306	0.431	0.155	2.78
307	0.344	0.099	3.47
308	0.371	0.200	1.86
309	0.339	0.108	3.14
Average	0.373	0.133	**3.00**





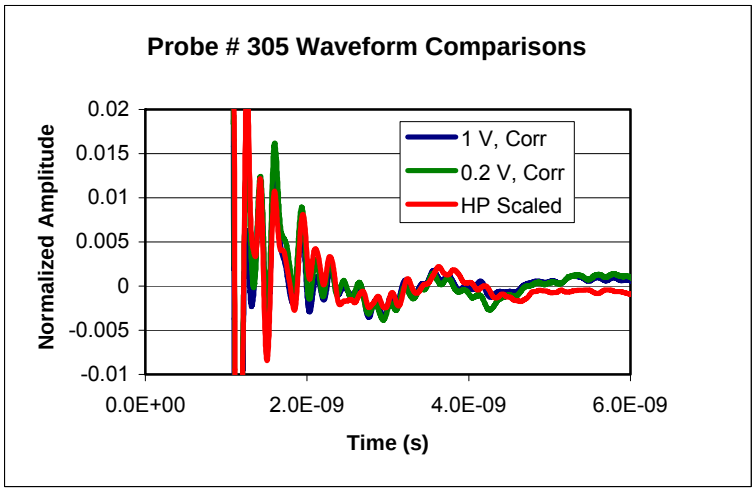
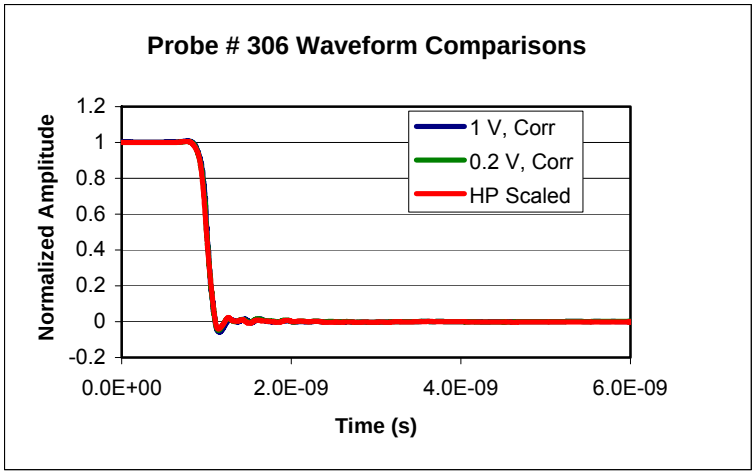
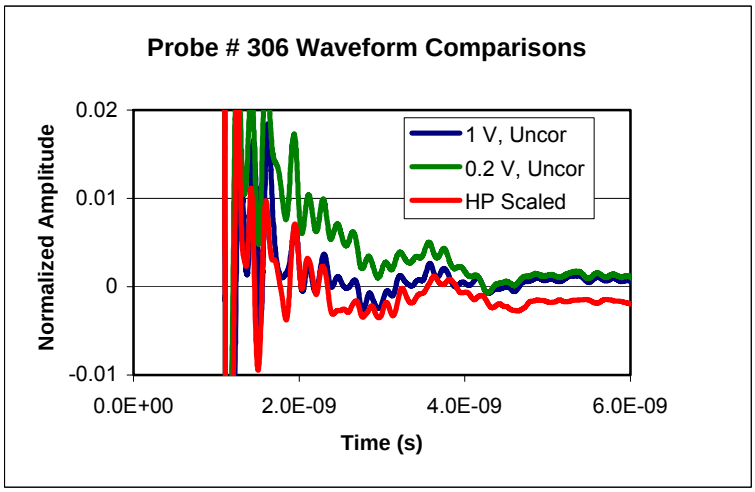


Fig. 17a, b, c Probe # 305 Waveform Comparisons.





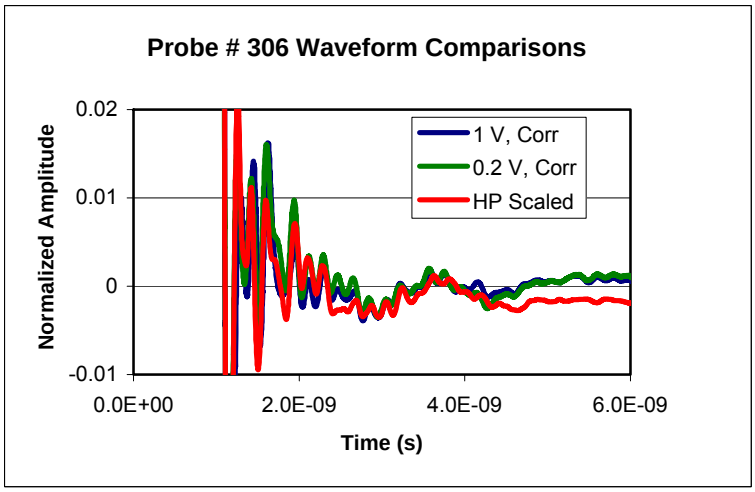
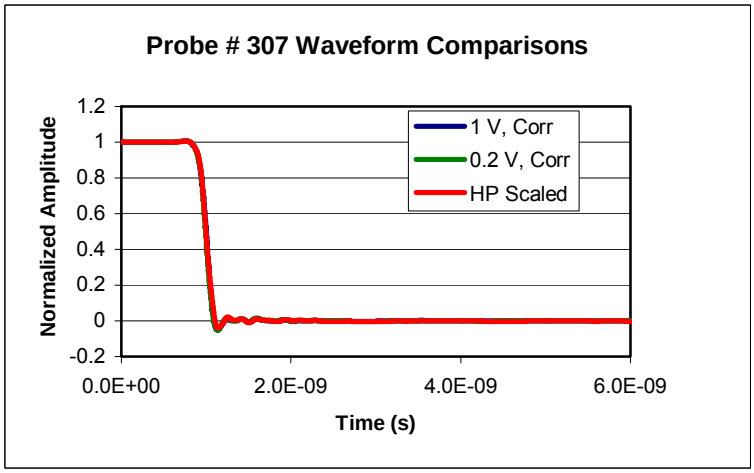
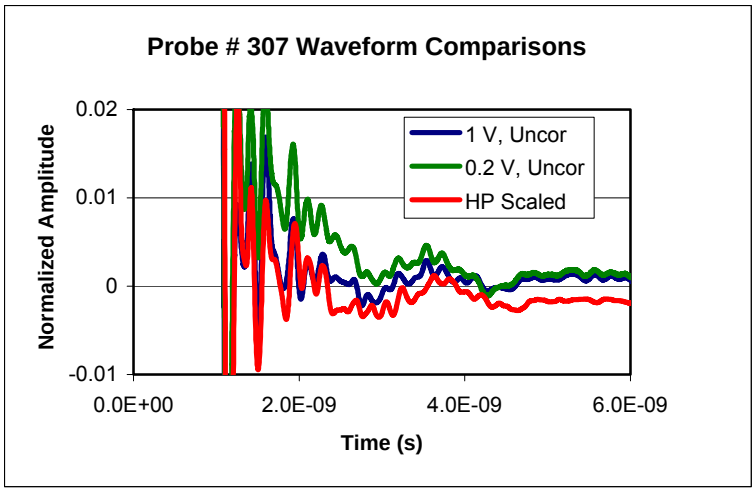


Fig. 18a, b, c Probe # 306 Waveform Comparisons.





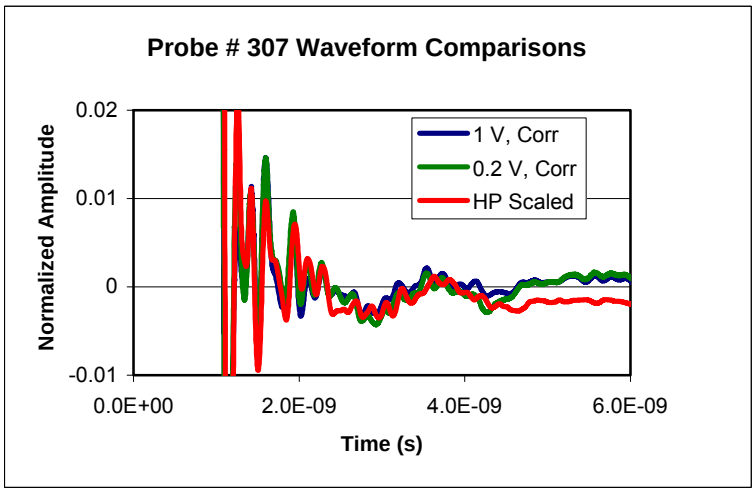
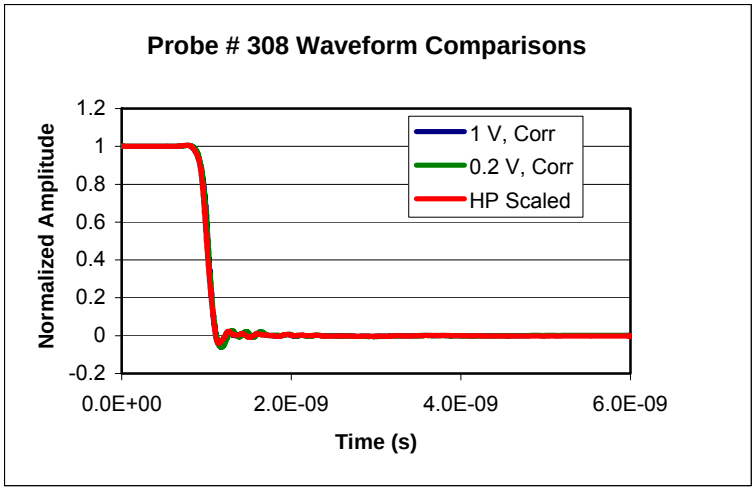
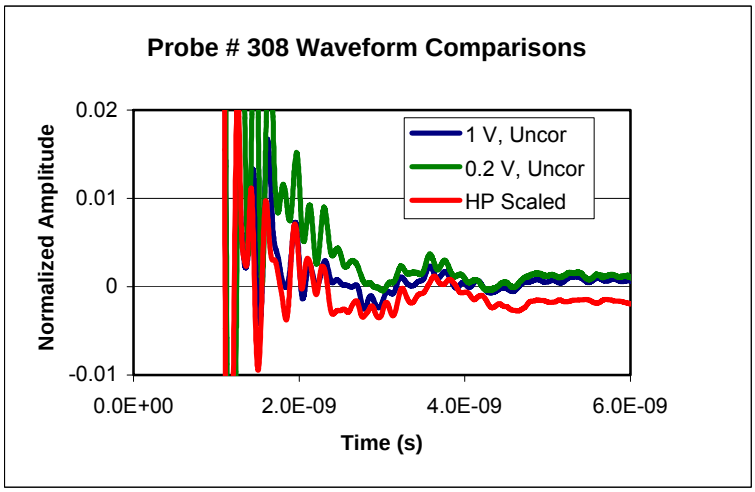


Fig. 19a, b, c Probe # 307 Waveform Comparisons.





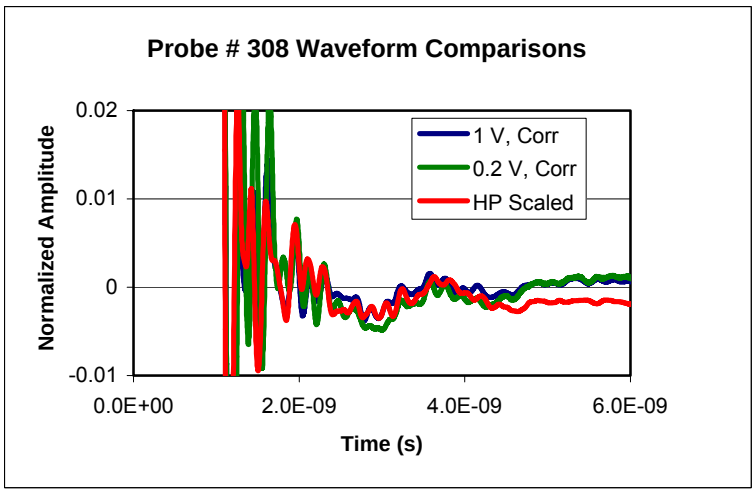
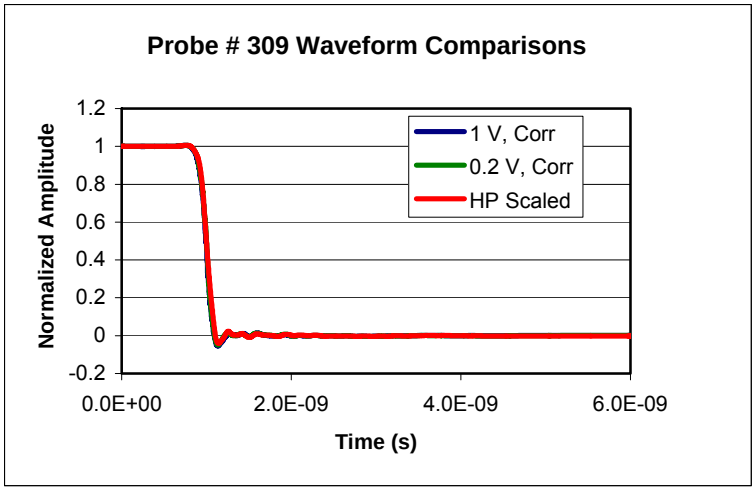
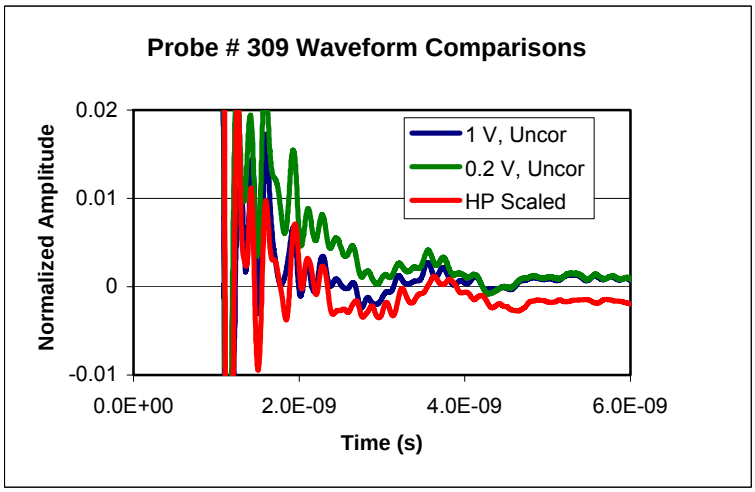


Fig. 20a, b, c Probe # 308 Waveform Comparisons.





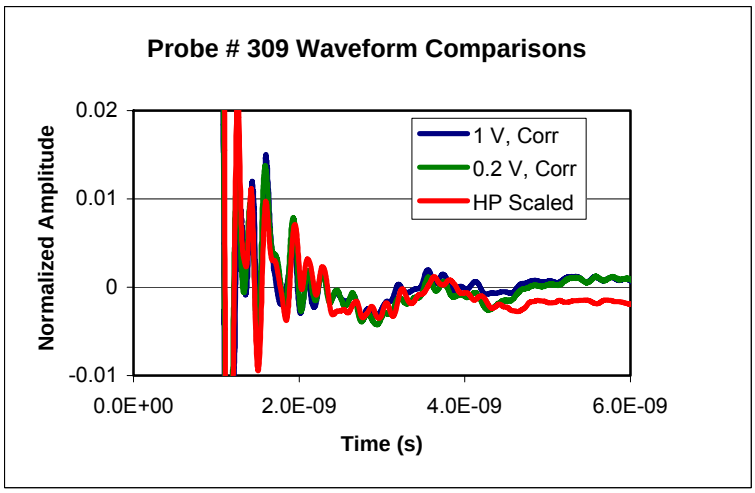


Fig. 21a, b, c Probe # 309 Waveform Comparisons.

4.2 Thermal Error Effects in the Frequency Domain

While the theory and correction-routine are in principle applicable to all types of repetitive waveforms, e.g., sinusoids, we have yet to demonstrate the efficacy for broader classes of signals. The impact of the model in the frequency domain was investigated analytically using both time and frequency domain approaches. This was undertaken since the probes will be used for both time and frequency domain measurements, e.g., for settling measurements of step-like waveforms and for RMS gain flatness measurements, among other applications. In the first approach, the model was applied to fast, ideal Gaussian steps with step-amplitudes ranging from -2 V to +2 V. The resulting step responses were transformed into the frequency domain to investigate the effect of the thermal-tail errors on spectral flatness. The results are given in Fig. 22. Note that a crossover frequency of about 30 MHz is predicted, above which significant thermally induced error occurs and is amplitude dependent. (The upturn shown at higher frequencies for the lowest amplitudes is an artifact of the numerical processing; in fact no upturn actually occurs.) In the second approach, the model was applied directly to pure sinusoidal waveforms at seven frequencies ranging from 10 MHz to 500 MHz. At each frequency, the amplitude was swept from near 0 V to 2 V, peak. The amplitude error and total harmonic distortion (THD) were computed in each case, and are given in Figs. 23 and 24. Note that the THD decreases dramatically at low amplitudes for all frequencies. This is because the error model is essentially linear at low amplitudes, causing only magnitude and phase error, but no harmonic distortion. Surprisingly, the spectral flatness calculated with the two methods agreed reasonably well, despite the fact that the model is nonlinear.

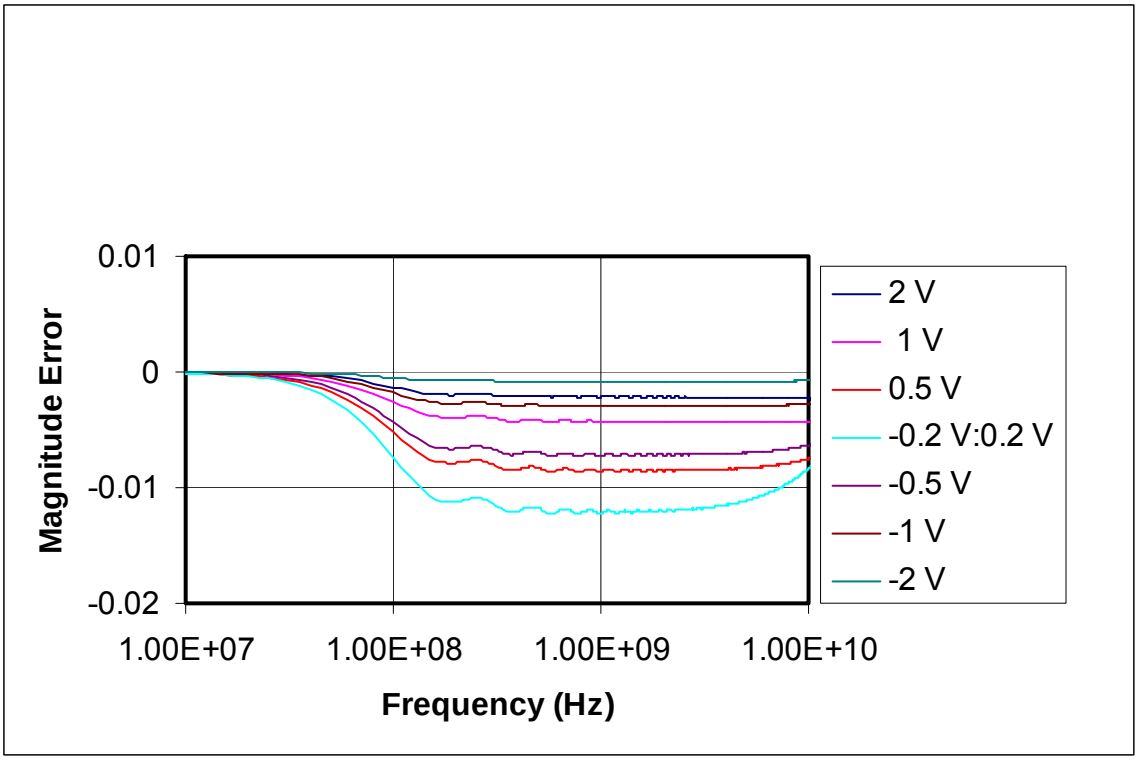


Fig. 22 Frequency Response from Step Response (Selected Step Amplitudes).

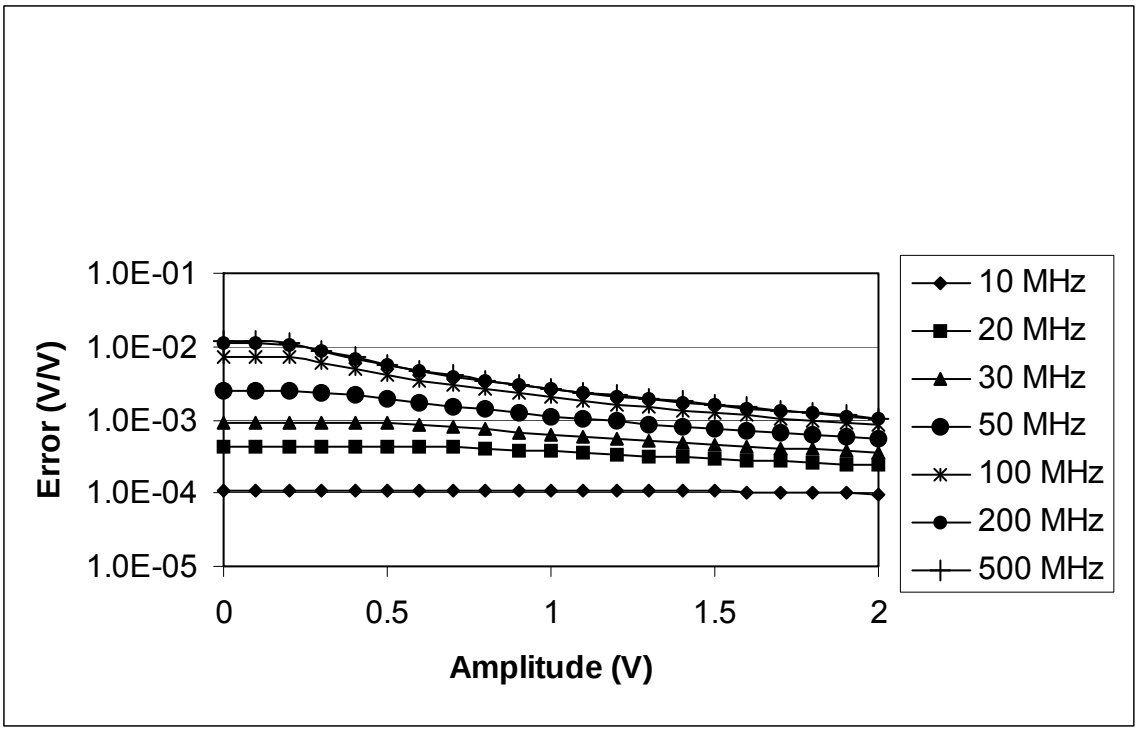


Fig. 23 Sinewave Magnitude Error vs. Amplitude (Selected Frequencies).

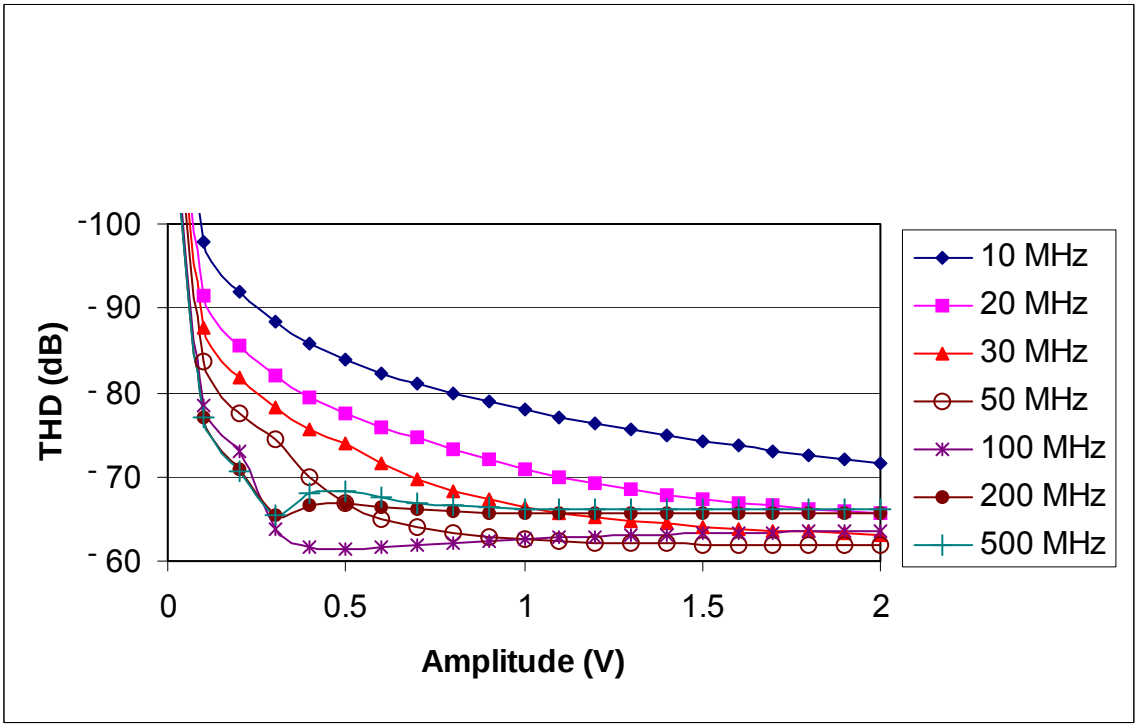


Fig. 24 Total Harmonic Distortion vs. Amplitude (Selected Frequencies).

The nonlinear thermal tail model discussed previously predicts that sinusoids measured with the NGSCP will exhibit harmonic distortion and gain error, examples of which are shown in the previously presented figures. To corroborate these predictions, measurement data was accumulated on 20 MHz filtered sine wave signals, with peak amplitude of about 1.6 V. Probe #305 was used. The frequency and voltage level were chosen to produce distortion levels that we hoped would be large enough to detect unambiguously, in the presence of noise, jitter and other sources of distortion. The data record was then processed using a 4-parameter sinewave curve fitting routine to determine the fit residuals, i.e., a time-domain representation of the harmonic content of the signal. Fig. 25 shows the fitted sine wave, and Fig. 26 shows the residuals of the fit. The data record was also processed using the thermal-tail error model, to determine the predicted error due to thermal tail. This is shown in Fig. 27. Next, the predicted thermal tail error record (Fig. 27) was subtracted from the original data record to produce "corrected" data, and a sine fit was again applied, giving the residuals shown in Fig. 28. Finally, in Fig. 29 we show the residuals of the corrected data, after removing the first three even harmonics, i.e., after removing the 2^{nd}, 4^{th}, and 6^{th} harmonics from the corrected data.

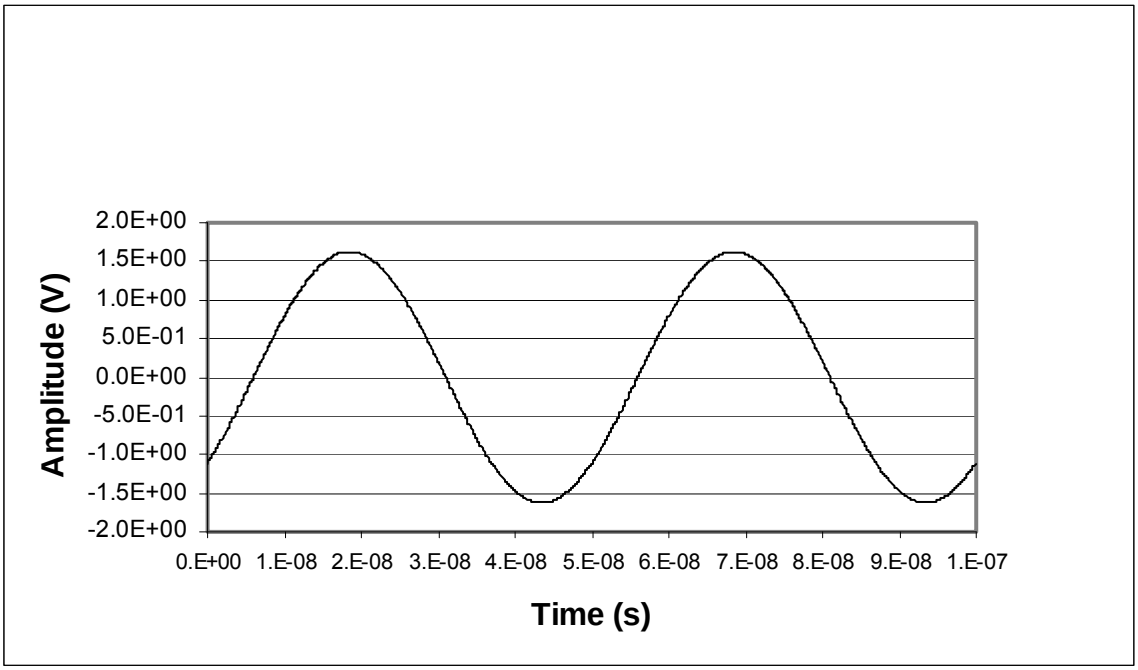


Fig. 25 Fitted Sinewave.

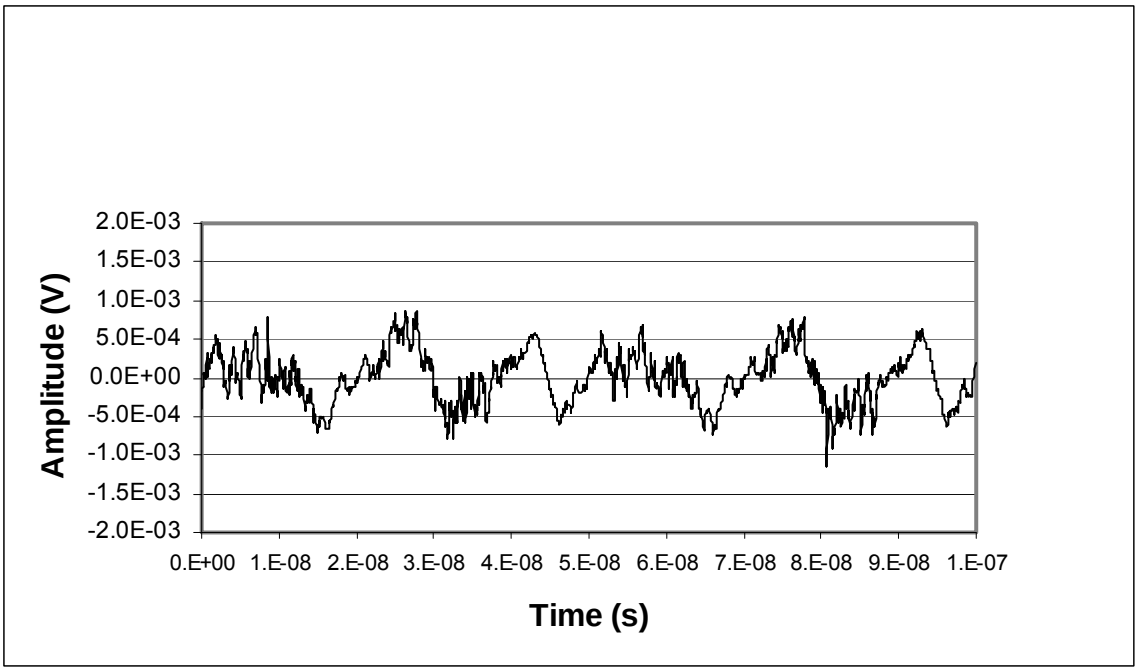


Fig. 26 Uncorrected Residuals.

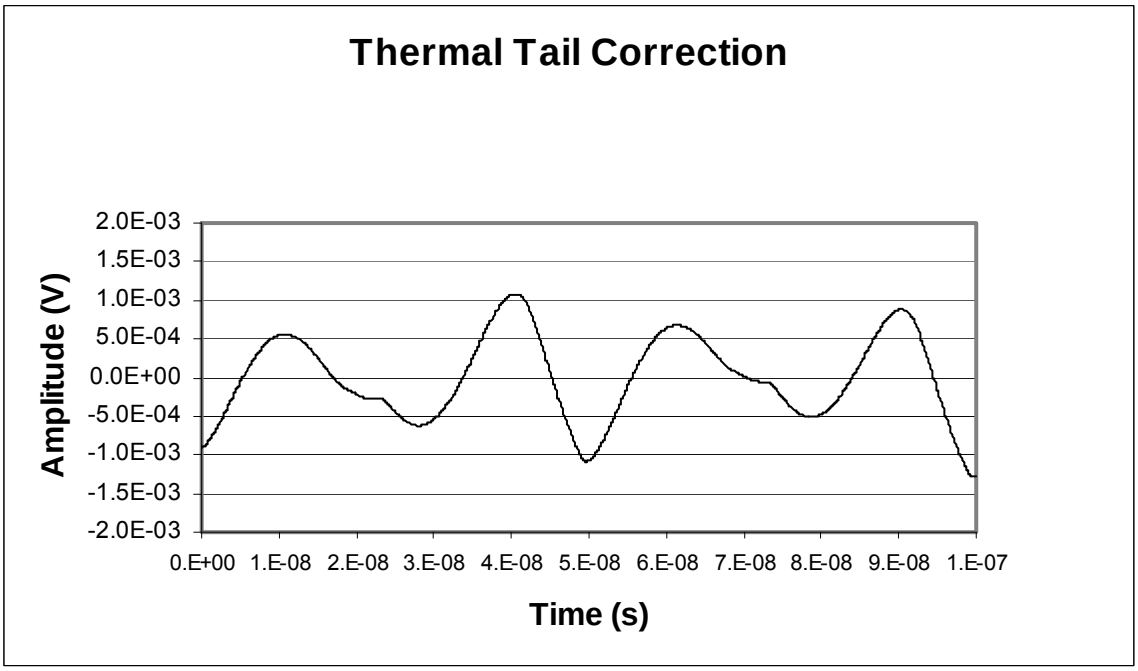


Fig. 27 Thermal Tail Correction.

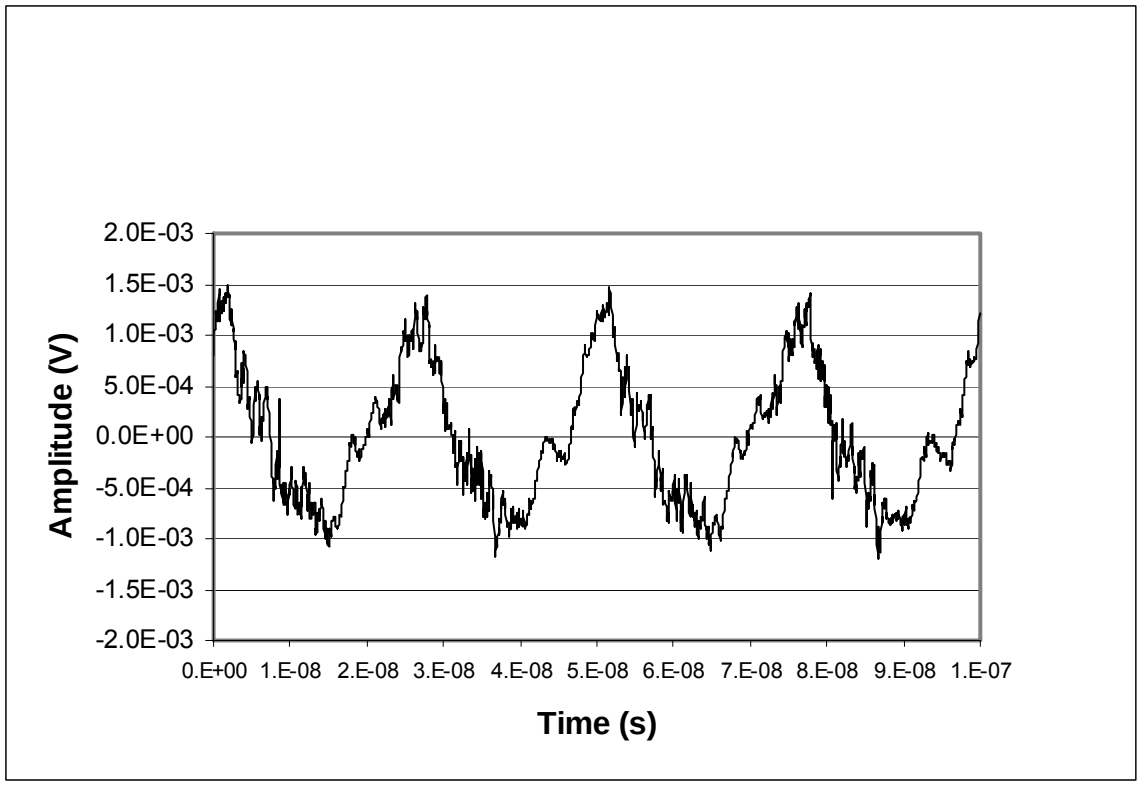


Fig. 28 Corrected Residuals.

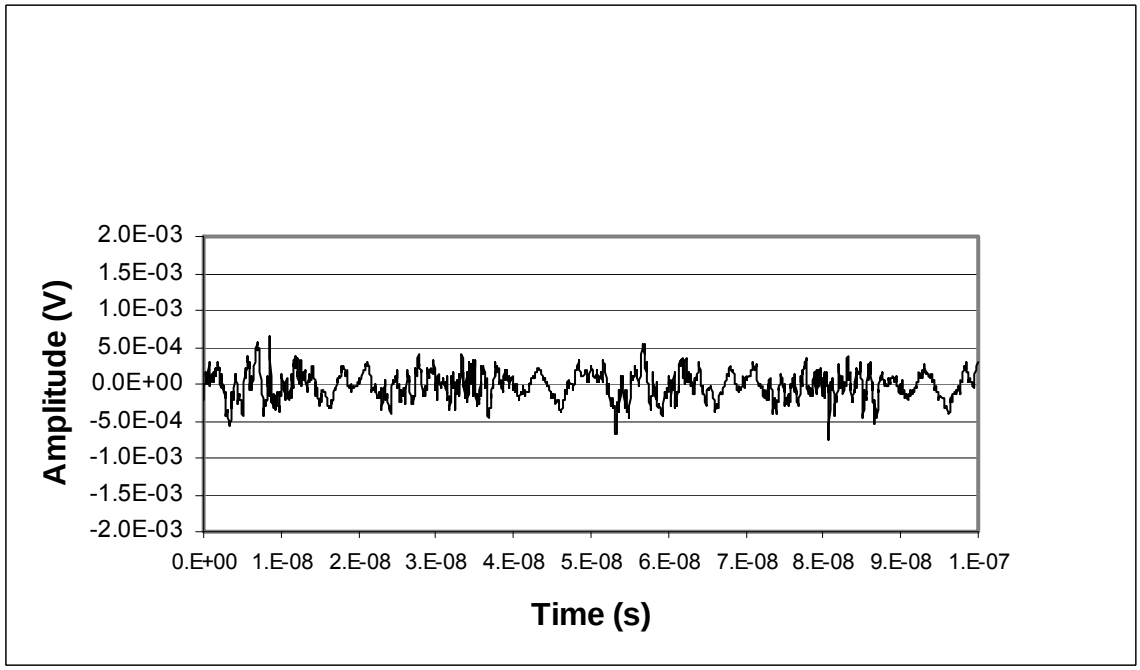


Fig. 29 Residuals of Fit Corrected (less 2nd, 4th, and 6th harmonics).

Table II summarizes the total harmonic distortion (THD) relative to the fundamental component found in each case (in mV/V as well as in dB), along with the relative levels of 2^{nd}, 3^{rd} and 4^{th} harmonic components.

Table II Summary of the total harmonic distortion (THD) relative to the fundamental

Source	THD (mV/V)	THD (dB)	2^{nd} (mV/V)	3^{rd} (mV/V)	4^{th} (mV/V)
Raw Data	0.301	-70.42	0.104	0.202	0.0084
Corrected Data (thermal)	0.573	-64.84	0.535	0.011	0.078
Correction Only	0.478	-66.41	0.434	0.212	0.025
Corrected less 2^{nd}, 4^{th} & 6^{th}	0.166	-75.60	-	0.011	-

As is evident from Figs. 26 and 28, as well as Table II, applying the thermal error correction actually increased the THD. Taken at face value, this suggests that the correction is not appropriate, which in turn implies that the underlying model is not correct. On the other hand, the correction almost entirely eliminates low order odd harmonics, leaving distortion that is predominantly even order. However, even-ordered distortion is believed to typify the harmonic content of the probes at higher frequencies.

5. Next Generation Sampling Comparator Probe Performance

The primary performance characteristics of the ASIC and NGSC probes are summarized in the table shown below. Following the table is a brief description of the parameter and the method of measurement.

Table III Electrical Performance Characteristics of ASIC Probe and NGSC Probe

PARAMETER	ASIC PROBE	NGSC PROBE
Input Range	±2 V	±1.8 V
Maximum Input Voltage	±2.5 V	±2.0 V
Transition Duration (10 % to 90 %)	150 ps	63 ps
RMS Noise (referred to input)	425 μV	250 μV
Total Harmonic Distortion (1.5 V peak)	-35 dB @ 100 MHz	-54 dB @ 100 MHz -33 dB @ 1 GHz
DC Linearity	≤ 50 μV/V	≤ 50 μV/V
Pulse Settling Error (referred to final DC value)	see Figs. 30 & 31	see Figs. 30 & 31

The Input Range is set by the full-scale reference voltage applied to the reference input to the comparator. The maximum input voltage is established by the maximum recommended reverse base-emitter voltage of the transistors of the differential input stage. For the NGSC Probe the worst case reverse voltage of one of the junctions occurs when either the reference or signal input Vbe junction is biased at the opposite extreme of the other. With a full-scale reference voltage of ±1.8 V applied at the reference input, ±2.0 V applied at the signal input, and taking into consideration the voltage drop across one emitter degeneration resistor, the worst case reverse base-emitter voltage across the off transistor results in about 2.7 V. The manufacturer does not recommend exceeding this value in order to avoid what is called forward beta degradation. This phenomenon is referred to as a "soft" failure resulting with unusually high base currents. The devices used in the ASIC comparator of the ASIC Probe are a bit more forgiving of reversed Vbe junction potentials. Voltages applied to the signal input of either probe which exceed the chip biasing voltages of ±5.0 V will most assuredly destroy the comparator. Neither comparator inputs (reference and signal) are protected.

Transition duration was measured with a Tektronix 067-1338-00 step generator. This particular step generator has a -0.25 V to 0 V step amplitude with a nominal transition duration of 15 ps. The transition duration reported above for each of the probes results from taking the root-difference-square of the pulse generator's transition duration from the measured waveform transition duration.

RMS Probe Noise was measured by acquiring a record of data with no signal applied to the probe. Since the intrinsic noise of the probe is approximately five times greater than the size of a code bin, the root-mean-square value of the data record is a good estimate of the probe's intrinsic noise.

Total Harmonic Distortion of the probe was measured by applying an 80 % full-scale signal to the probe and taking a record of data containing precisely 11 signal periods in 4096 data samples. An FFT of the data was computed, and Total Harmonic Distortion was computed from the first 10 harmonics in the spectrum.

DC Linearity of the probes is measured by applying a known dc source to the input over the entire input dynamic range and observing the averaged results. In very demanding applications the effects of nonlinearity are compensated for by applying a correction algorithm within the SWA mainframe.

Pulse Settling Error, ε, at time instant t_x is defined as the maximum absolute difference between the waveform's final value and any value of the waveform occurring over the interval from t_x to the end of the data record [12]. The final value is typically specified as the value of the waveform 1 s after the transition. However, the pulse generators used in the measurements to characterize pulse settling performance of the probe were capable of sustaining constant static levels. The waveform final value was therefore approximated as the mean value of the data record resulting from measuring the static level corresponding to the pulse's settled state.

The estimated settling uncertainties were determined from an ensemble of measurements made with Probe # 305 using four different reference step generators each having different step amplitude and/or polarity of transition and used with and without a wideband 6 dB attenuator. These included a Picosecond Pulse Labs model 6110 generator that produced a negative transition pulse with an amplitude of 0.5 V; two Yokogawa-Hewlett Packard generators producing positive and negative pulses respectively, each having an amplitude of about 0.6 V; and a Tektronix generator producing a positive pulse having an amplitude of 0.25 V. For each of the generators, the voltage level following the step transition was nominally zero, and the generators were designed to give fast, accurate settling to that level.

The resulting uncertainties are estimates of the maximum settling errors that could be considered common to all of the measurements and, therefore due to the probe itself rather than the individual step generators. These settling errors were computed from a mean waveform consisting of the average of the four waveforms (after amplitude normalization and time alignment) from the four measured pulse generators.

Figures 30 and 31 show the computed pulse settling error for waveform epochs of 20 ns and 200 ns, respectively. Each plot shows the mean waveform with and without application of the thermal correction method described in Section 4 and with both attenuated and unattenuated signal levels. At 1 ns, the thermal correction scheme improves probe settling considerably from 0.7 % to 0.27 %. At 2 ns and 3 ns, the thermal error correction scheme actually degrades probe settling slightly. From 4 ns and beyond, thermally corrected and uncorrected settling are the same because the probe's enable time was set to 4 ns, and model-generated corrections are zero beyond the enable time. The good agreement between the attenuated and unattenuated settling responses illustrates

good dynamic linearity of the probe, an insight into performance normally described in the frequency domain in terms of distortion parameters, but seen here in the time domain. Also, given in each of the graphs for reference purposes is the expanded uncertainty curve for pulse settling error using the ASIC Probe.

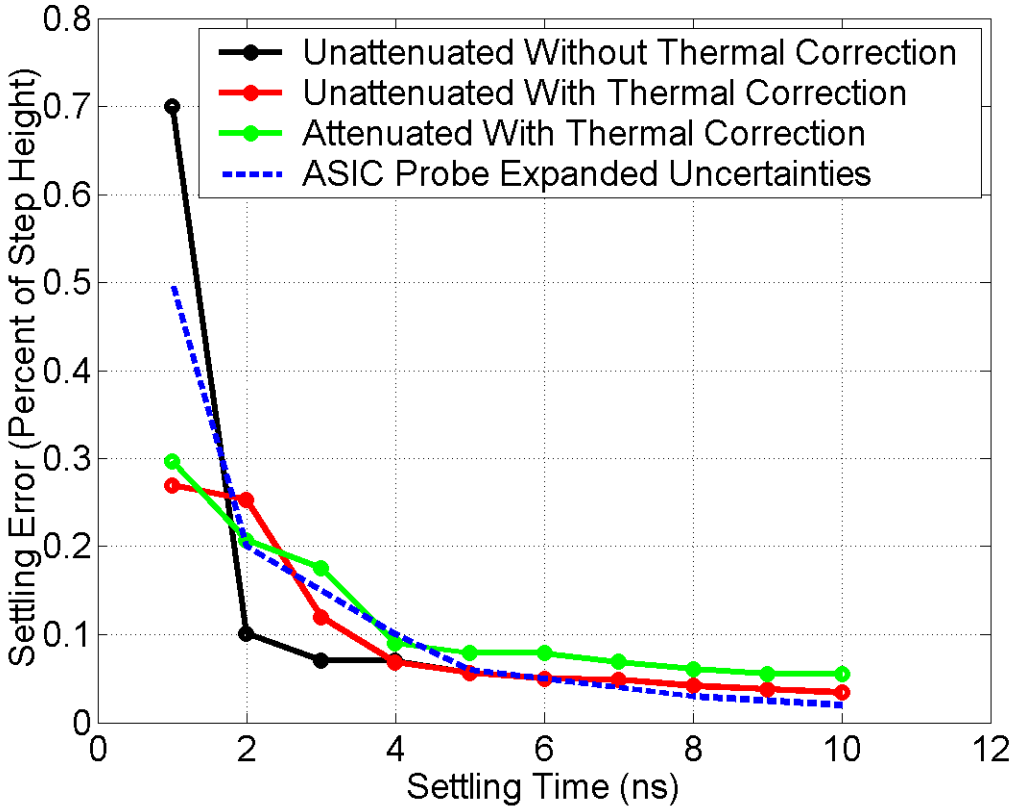


Fig. 30. NGSC Probe settling error determined from mean waveform of pulse generator ensemble measurements. Waveform epoch equals 20 ns.

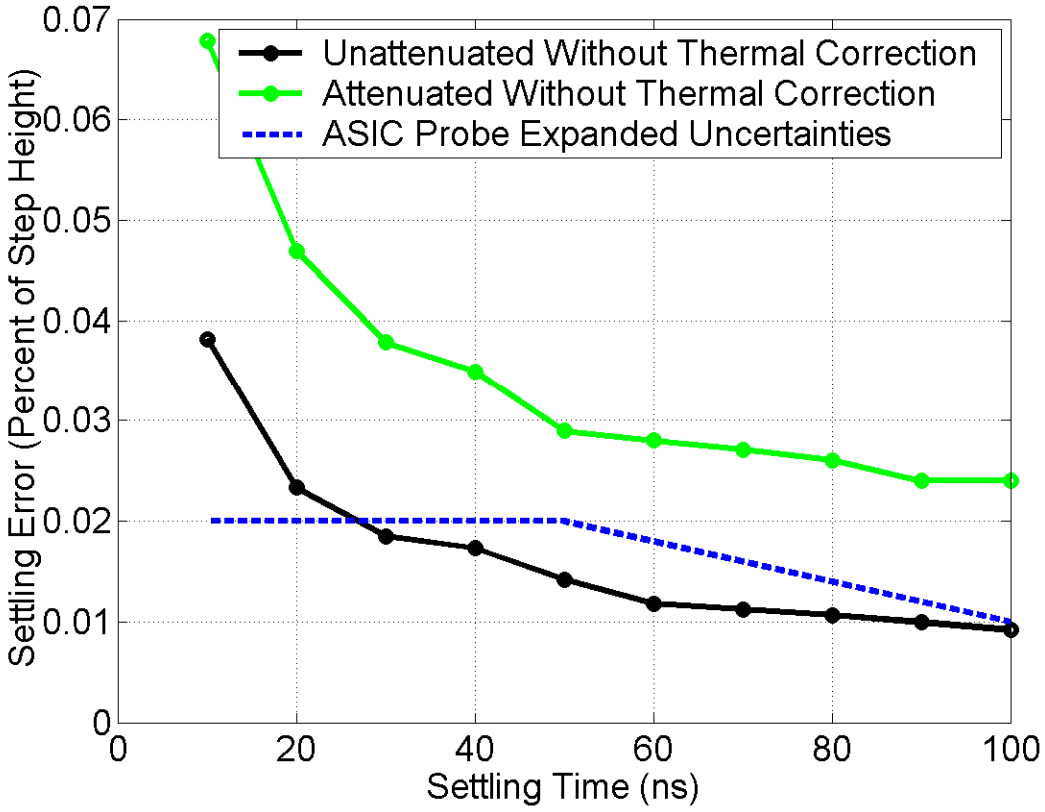


Fig. 31. NGSC Probe settling error determined from mean waveform of pulse generator ensemble measurements. Waveform epoch equals 200 ns.

From the results summarized in Table III and the above discussion, we see that the NGSC Probe generally outperforms the ASIC Probe in bandwidth, distortion, and noise. However, in terms of pulse settling performance the NGSC probe produces mixed results. Thermal error correction whether for large or small signals generally reduces the settling error. Clearly below 2 ns the settling error is much improved by employing correction. It is unfortunate that the enabling technique is by itself insufficient in eliminating all of the settling error of the NGSC Probe.

6. Acknowledgement

The authors gratefully acknowledge the efforts of Robert Palm who performed all of the circuit layouts and assembly of the probes. He also tirelessly assembled the many assorted boards and components of the NIST Sampling Waveform Analyzer.

7. References

[1] T. M. Souders, et. al., "A Wideband Sampling Voltmeter," IEEE Trans. Instrum. Meas., Vol. 46, No. 4, August 1997, pp 947-953.

[2] O. B. Laug, T. M. Souders, D. R. Flach, "A Custom Integrated Circuit Comparator for High-Performance Sampling Applications," IEEE Trans. Instrum. Meas., vol 41, Dec. 1992, pp 850-855.

[3] D. I. Bergman, B. C. Waltrip, "A Low Noise Latching Comparator Probe for Waveform Sampling Applications," IEEE Trans. Instrum. Meas., vol. 52, No 4, Aug. 2003, pp 1107-1113.

[4] D.I. Bergman, "NIST Sampling System for Accurate AC Waveform Parameter Measurements," IEEE Trans. Instrum. Meas., vol. 56, No 2, April. 2007, pp 244-248.

[5] S. P. McCabe III, "A sampling voltage tracker for analyzing high speed waveforms," M.S. thesis, Univer. of California, Los Angeles, 1975.

[6] B. Peetz, B. D. Hamilton, J. Kang, "An 8-bit 250 Megasample per Second Analog-to-Digital Converter: Operation Without a Sample and Hold." IEEE Journal Solid State Circuits, vol 21, No. 6, Dec. 1986, pp, 997-1002.

[7] C. Gyles, "Repetitive Waveform High Frequency, High Precision Digitzer," IEEE Trans. Instrum. Meas., vol 38, No. 4, Aug. 1989, p 917

[8] Owen B. Laug, T. Michael Souders, "Improved Voltage Comparator with Reduced Settling Time," United States Patent Number 5,347,175, Sept. 13, 1994.

[9] Maxim Integrated Products, "High-Frequency ASIC Development Handbook, Second Edition," Maxim Integrated Products is publisher

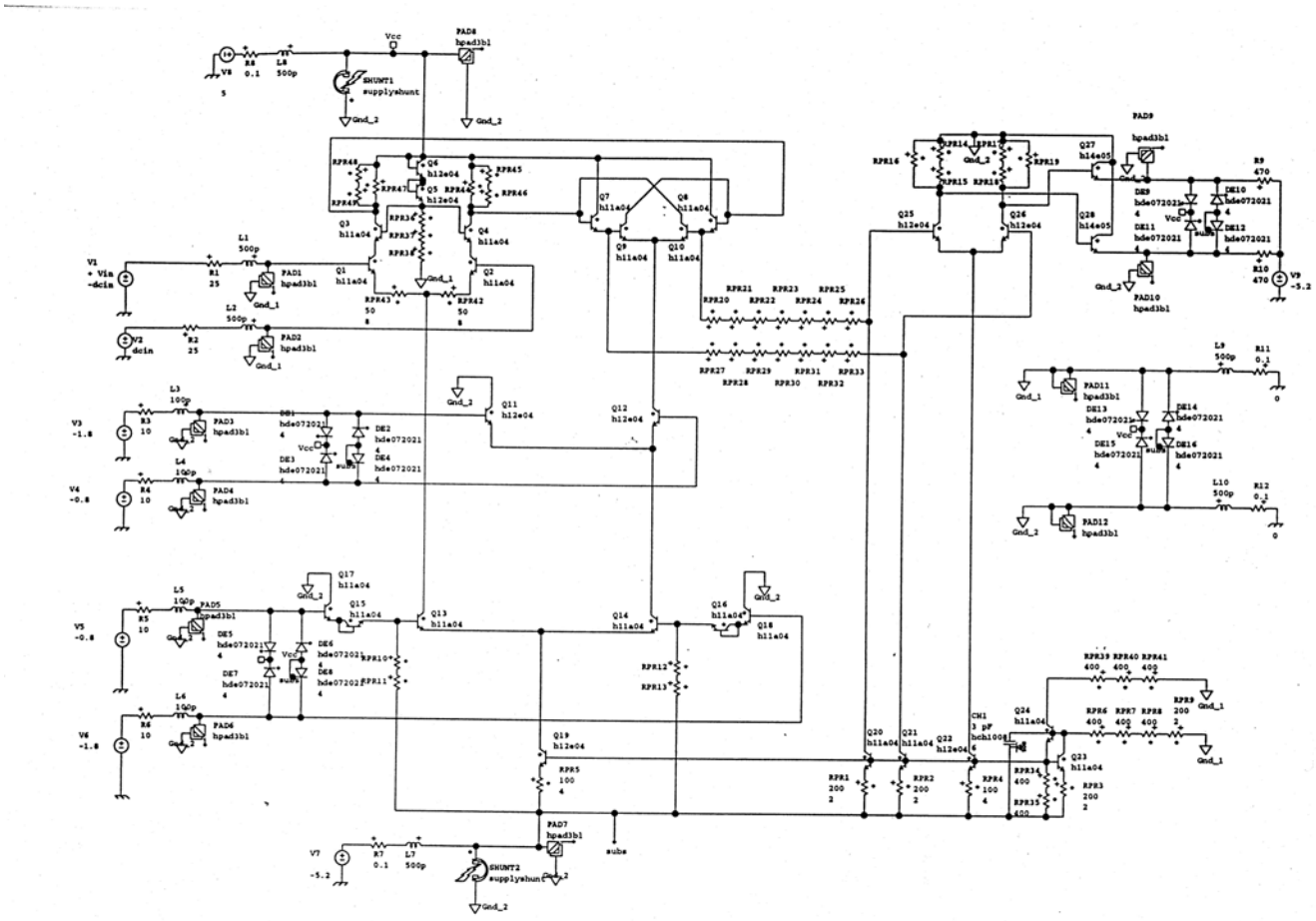
[10] M. Busani, et. al., "Dynamic Thermal Characterization and Modeling of Packaged AlGaAs/GaAs HBTs" IEEE Trans. on Components and Packaging Technologies, Vol. 23, No. 2, June, 2000, pp 352-359.

[11] R. C. Joy, "Thermal Properties of Very Fast Transistors," IEEE Trans. on Electron Devices, vol. ED-17., No. 8, August 1970, pp 586-594.

[12] IEEE Std 181™-2003, "IEEE Standard on Transitions, Pulses, and Related Waveforms," IEEE Instrumentation and Measurement (I&M) Society, March 2003.

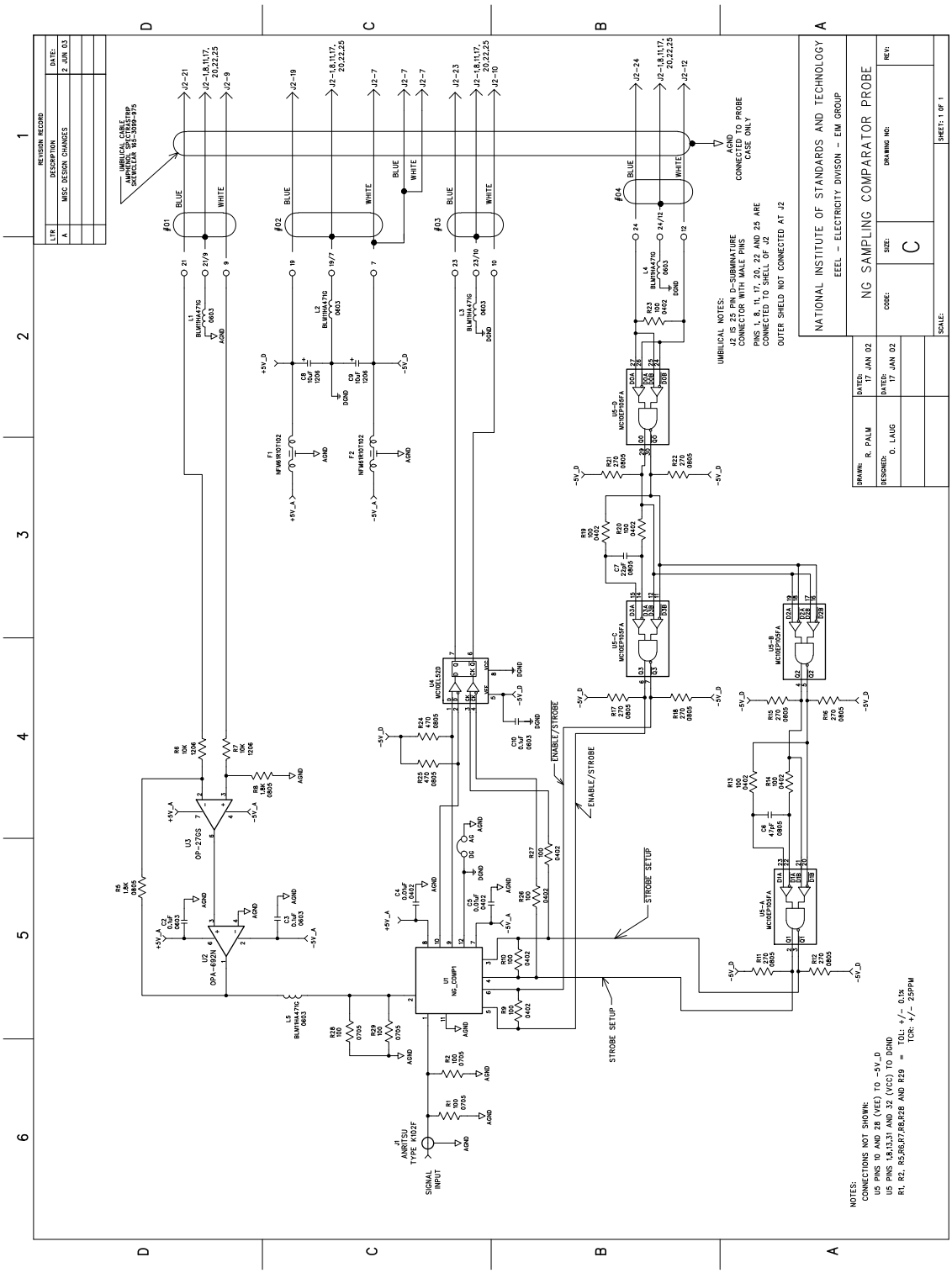
Appendix A

Schematic Diagram of the Next Generation Sampling Integrated Circuit Comparator



Appendix B

Schematic diagram of NGSC Probe



Appendix C
Notes on the Computational Implementation

For a given m-point waveform record, the discrete-time version of (1) proceeds as follows:

1. Express t_e and τ in units of sampling period, t_p, giving N_e = intPart (t_e/t_p) and t_τ = intPart (τ/t_p) where intPart stands for "the integer part of".
2. Depending on the signal type (repetitive, e.g., sinusoids, or non-repetitive, e.g., steps) append an appropriate extension of N_e samples on the front of the waveform to minimize end effects, i.e., to deal with the enabling window preceding the first points in the record. For repetitive signals, N_e samples from the end of the record are appended to the beginning. For non-repetitive signals, the value of the first point is repeated for N_e samples appended to the beginning.
3. Acquire each of the m intervals of N_e points in turn, starting from the beginning of the extended waveform. (The 1st interval will end with the 1st data point of the original waveform, and so on). Call the last point of each interval the ref. level for that interval. Use this to select the heating curve to use (as in fig. 5). Compute an array A of N_e differential power levels that correspond to the N_e sample values (input signal levels) in the interval. Calculate another array B of N_e points of exponential decay values, for each relative time point in the interval. (Since all of the intervals are the same for a given t_e, τ, and t_p, this array need only be calculated once.) The k[th] element is given by

$$e^{-\frac{(N_e-k+1)}{N_\tau}} \qquad (2)$$

4. Compute the scalar product (A^T B) of these two arrays to get the effective differential heat at the sampling instant (end of interval), and then multiply this by constant $C = 1.4 \times 10^8 \times t_p$ to get the effective thermal offset voltage for the sample. (Factor t_p appears in C to reintroduce the unit of time that was removed in step 1.)

Note:

An HP-VEE program was written to execute the computations described above. Significant time was spent in trying to optimize the program to minimize execution time; nevertheless, the program still takes longer to execute than we would like: Running on a 600 MHz Pentium II, the execution time is approximately 12 ms per sample, i.e., 12 s for a 1000 point record.

Appendix D
Notes on the Constant C

The constant, C, was used in two different ways during the course of this work. In the beginning, the functional dependence of the differential power dissipation, $\Delta P(t)$, on the input and reference voltages had not been determined. In the earlier work based on a linear model, the constant C_L was determined experimentally. As (5) indicates, it can be represented as the product of the constant C_{NL} used later in the nonlinear work, and the zero-crossing slope, S_{ramp}, of the linear portion of the differential power curves. Later, the relationship for $\Delta P(t)$ (and hence S_{ramp}) was determined and used in the nonlinear modeling work. At that point, constant C_{NL} was preset using (5) and the then known values for C_L and S_{ramp}. The relationship between C_L and C_{NL} is derived below, starting with (1) from the main text.

$$\Delta V(t_s) = C_{NL} \int_{t_s - t_e}^{t_s} \Delta P(t)\, e^{-(t_s - t)/\tau} dt \tag{1}$$

Where C_{NL} is the constant that is used in the nonlinear case, and $\Delta P(t)$ is the differential power dissipation at time t that occurs when the input voltage is $V_{in}(t)$ and the reference voltage is V_{ref}. For the linear case, the relationship between $\Delta P(t)$ and the difference voltage $V_{in}(t) - V_{ref}$ is approximated by the central ramp segment of Fig. 16. Using this relationship, we can express the differential power as:

$$\Delta P(t) = (V_{in}(t) - V_{ref}) \times S_{ramp} \tag{2}$$

Where S_{ramp} is the slope of the central ramp segment, with a value of 0.0408 W/V.

Therefore, for the linear case, we have

$$\Delta V(t_s) = C_{NL} \int_{t_s - t_e}^{t_s} (V_{in}(t) - V_{ref}) \times S_{ramp}\, e^{-(t_s - t)/\tau} dt \tag{3}$$

$$= C_L \int_{t_s - t_e}^{t_s} (V_{in}(t) - V_{ref})\, e^{-(t_s - t)/\tau} dt \tag{4}$$

Where C_L is the constant that is used in the linear case and is given by

$$C_L = C_{NL} \times S_{ramp} \tag{5}$$

In the initial work using a linear model, the value of C_L was found to be approximately 5.6×10^6 s^{-1}. Working backwards from this value and the calculated value for S_{ramp} given above, we get a value for C_{NL} of 1.37×10^8 V/W·s.