Internal Signal Sampling of High-Speed Integrated Circuits using Dynamic Electrostatic Force and Force-Gradient Microscopy

by

Terry (Zhinong) Weng

A thesis
submitted to the Faculty of Graduate Studies
in partial fulfillment of the requirement
for the degree of

Doctor of Philosophy

Department of Electrical and Computer Engineering
University of Manitoba,
Winnipeg, Manitoba
Canada

UNIVERSITY OF MANITOBA

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

Internal Signal Sampling of High-Speed Integrated Circuits using

Dynamic Electrostatic Force and Force-Gradient Microscopy

BY

Terry (Zhinong) Weng

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University

of Manitoba in partial fulfillment of the requirements of the degree

of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES
FINAL ORAL EXAMINATION OF THE PHD THESIS

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a PhD thesis entitled:

"Internal Signal Sampling of High-Speed Integrated Circuits using Dynamic Electrostatic Force and Force-Gradient Microscopy"

BY

Terry (Zhinong) Weng

In Partial fulfillment of the requirements for the PhD Degree

.....
Dr. G. Bridges, Advisor

.....
External Examiner
Dr. Geoff Steeves
Dept. of Physics and Astronomy
University of Victoria
Victoria, B.C.

.....
Dr. C. Shafai

.....
Dr. R. M. Roshko

Date of Oral Examination:..... April 18, 2005 .....

The Student has satisfactorily completed and passed the PhD Oral Examination.

.....
Dr. G. Bridges, Advisor

.....
External Examiner
Dr. Geoff Steeves
Dept. of Physics and Astronomy
University of Victoria
Victoria, B.C.

.....
Dr. C. Shafai

.....
Dr. R. M. Roshko

.....
Dr. J. Baltes , Chair of PhD Oral

(The signature of the Chair does not necessarily signify that the Chair has read the thesis.)

Abstract

The ability to measure the internal signals of integrated circuits (ICs) is critical for the circuit design and failure analysis. Conventional contact probes are often inadequate for testing of current IC technologies which are featured by smaller sizes and faster speeds. Electrostatic force microscopy (EFM) has been proved to be an alternative solution.

In EFM, circuit voltages can be measured by monitoring the mechanical deflection of a noncontacting microfabricated probe (with a sharp tip at the end) as it responds to the localized electrostatic interaction. The measurement bandwidth of the EFM using direct methods is limited to the mechanical frequency response of the probe. To overcome the bandwidth limitation, dynamic-mode EFMs (DEFMs) have been developed in 1990s by employing the nonlinear dependence of the electrostatic interaction on the difference between the probe and the circuit voltages.

This is an extended research on the dynamic mode electrostatic force microscopy. The goal is mainly to improve the spatial and temporal resolution in its application for the time-resolved measurement of the internal voltages of high-speed ICs. The instrument developed in this research measures the high frequency circuit signals by driving the circuit with a synchronized repetitive pattern and the probe with an amplitude modulated sampling pulse train. The detected deflection signal at the modulation frequency is proportional to the circuit test point voltage, due to the signal mixing between the probe and the circuit voltages. By adjusting the phase difference between the two voltages, the instrument provides the wave shape of the test point signal, yielding information such as rise and fall times, ringing and timing.

To obtain quantitative absolute voltage information, a null-force technique is adapted in this research to enable the instrument for quantitative voltage measurements in conjunction with very high bandwidth measurements.

Using the instrument, measurements are performed in ambient conditions, interfacing with existing test equipment and minimal sample preparation requirements. From the measurements, system performance, such as the bandwidth, timing accuracy, voltage sensitivity and probe invasiveness are evaluated in comparison with theoretical characterization. The bandwidth of the instrument is improved by using a high-speed pulse generation system to generate sampling pulses with pulse width as small as 37 ps.

Repetitive sampling measurements for ICs operating up to 10 Gbit/s are presented and the instrument demonstrates a bandwidth of 8 GHz and a timing accuracy of 10 ps. Quantitative voltage measurements for ICs operating up to 250 Mbit/s are presented and the instrument demonstrates a voltage accuracy of < 30 mV over a dynamic range 2.5 V.

The spatial resolution of the instrument using a standard EFM probe has been investigated experimentally and theoretically. Studies show that the probe is unable to achieve sub-micron spatial resolution in practical probing of an operating IC, due to the large tip-to-sample distance (typically >100 nm) required when operated on a probe station in ambient conditions or when a circuit passivation layer is present. This could result in strong coupling of interfering signals to the probe tip sidewall and the cantilever and causes incorrect waveform and quantitative voltage measurements. To improve the probe spatial resolution, two dynamic electrostatic force-gradient microscopy approaches are proposed in this research. One uses a numerical force-gradient detection approach, in which the force gradient is extracted numerically from directly sensing the force induced deflections for two tip-to-sample distances. The second approach enables direct detection of high-frequency force-gradients by employing a high-frequency implementation of a mechanoelectric cross-modulation scheme. It involves using an external vibration oscillator to mechanically modulate the tip-to-sample distance. Measurements on three parallel interconnects of an IC operating up to 500 Mbit/s demonstrate a submicron spatial resolution and a significant reduction of interference from adjacent interconnect signals by using the force-gradient detection approaches.

Acknowledgements

I would like to acknowledge and sincerely thank my supervisor, Dr. Greg Bridges, for his support and guidance. Without his advice and patient supervision, I would not have been able to finish this work. As a supervisor, Greg is truly exceptional. He is dedicated to his work and sets an excellent example for me. I am grateful to have been a part of his lab. Thank you for your time on draft editing as well.

I would like to thank Dr. Doug Thomson and Dr. Cyrus Shafai for their valuable support.

I would like to thank Sunny Cheung for his orientation of the research and Loren Card for his participation in the initial development of the LabView programs for automating the measurements. Special thanks to Chris Falkingham for his valuable discussions in solving the problems related with quantitative voltage measurement and Tom Kaminski for his participation in the numeric simulation of the probe spatial resolution and draft checking as well. Thanks to all members of SPM/MEMS group and the department technical staff at University of Manitoba.

I would like to thank my wife, Mina and my son, Guangzhe. Thank you for your love and your support. I also thank my parents for their love, support and encouragement.

Finally, thanks God in heaven for bringing me to Canada and giving me a peaceful life.

The research reported in this thesis was supported by the Natural Sciences and Engineering Research Council (NSERC), Micronet, and Canadian Microelectronics Corporation (CMC). Nortel Networks provided some test circuits.

Contents

Leaf left blank to correct page numbering

Leaf left blank to correct page numbering

List of Figures

List of Tables

List of Nomenclature

ac, dc	Alternating Current, Direct Current
ADC	Analog-to-Digital Converter
AFM	Atomic Force Microscopy
A. U.	Arbitrary Unit
CPW	Coplanar Waveguide
DEFM	Dynamic Electrostatic Force Microscopy
DGS	Data Generator System
DSO	Digital Sampling Oscilloscope
DUT	Device under Test
EFM	Electrostatic Force Microscopy
FET	Field Effect Transistor
FWHM	Full Width at Half Maximum
GPIB	General Purpose Interface Bus (IEEE-488.2)
HF	High Frequency
IC	Integrated Circuit
MEMS	Microelectronic Mechanical System
MMIC	Monolithic Microwave Integrated Circuit
RMS	Root Mean Square
SNR	Signal to Noise Ratio
SFM	Scanning Force Microscopy
SPM	Scanning Probe Microscopy
TDR	Time-Domain Reflectometer
TTL	Transistor-Transistor Logic
VI	Virtual Instrument, LabVIEW program

Chapter 1

Introduction

The rapid advance of integrated circuit (IC) technologies is characterized by substantial decreases in the size and spacing of structures, and increases in the chip complexity, size and speed of operation [1]. Modern IC designers often find that their circuits do not perform up to specifications, especially at high frequencies, no matter how sophisticated the computer aided design tools they used. Errors due to timing, cross talk, power supply bounce, power grid fluctuation, and ringing may cause the circuit to have a functional failure. Circuit diagnosis and failure analysis require internal probing of the circuitry of the packaged ICs to analyze signals, which are usually not accessible through the ICs' external I/O connections (as shown in Fig. 1.1) without adding extra circuitry in the design.

To perform internal probing of an IC, which can operate at over 1 Gbit/s with feature sizes less than 1 μm, an ideal diagnostic instrument must be able to provide noninvasive measurements with both high temporal and high spatial resolution and to be operated in a normal environment (without the need for complex calibration).

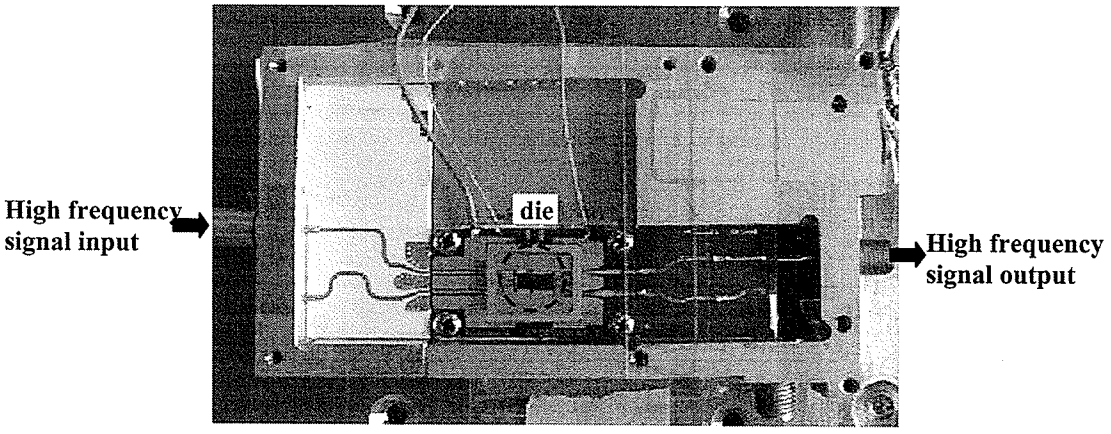

Fig. 1.1. A packaged IC with a die and I/O connections.

1.1 Internal IC probing techniques

At present, internal IC probing is carried out by contact or non-contact probing techniques. Contact probing uses a small metal (needle or wire) probe, which extracts an internal electrical signal by establishing a direct electrical contact with a circuit test point. Such probes are often in various physical sizes and can be organized into three categories: low impedance, medium impedance and high impedance [2].

Low impedance probes, such as coaxial, microstrip, and coplanar waveguide probes, are designed for matched impedance (50 Ω) measurements, such as for monolithic microwave integrated circuit measurements (MMICs) [3]. Low impedance probes usually have very high bandwidths; however, they are usually too invasive for the internal probing due to the 50 Ω loading these probes present.

Medium impedance probes, such as resistive divider probes, are less invasive in comparison with low impedance probes. A tradeoff between loading and voltage measurement sensitivity exists for these probes.

High impedance probes can be categorized into two classes, active and passive. A passive probe uses a large series resistor in parallel with a compensating capacitor to achieve high bandwidth. An active probe is commonly used for a digital sampling oscilloscope. This type of probe utilizes an active buffer near the probe tip to separate the 50 Ω load of the oscilloscope from the circuit under test. The bandwidth of an active probe is typically limited by the bandwidth of the active buffer. At high frequency, probe loading is usually determined by its shunt capacitance, which is the combination of the gate capacitance and the surrounding parasitics of the probe. Figure 1.2 is a picture of an active probe from Picoprobe, which has a tungsten probing wire with a very sharp end [4]. The fastest commercial active probe has been reported to have rise/fall times of 14 ps (a frequency response of DC to 26 GHz), and a nominal loading input impedance of 1.25 MΩ shunted by 0.05 pf (Picoprobe Model 35) [4].

The physical size of a contact probe determines its spatial limitation. A wire probe may have a sharp end radius as small as 100 nm [4]. Smaller contact probes have been developed using techniques based on the contacting Atomic Force Microscope (AFM) probes [5]. The typical AFM probe tip radius is 10 to 40 nm [6].

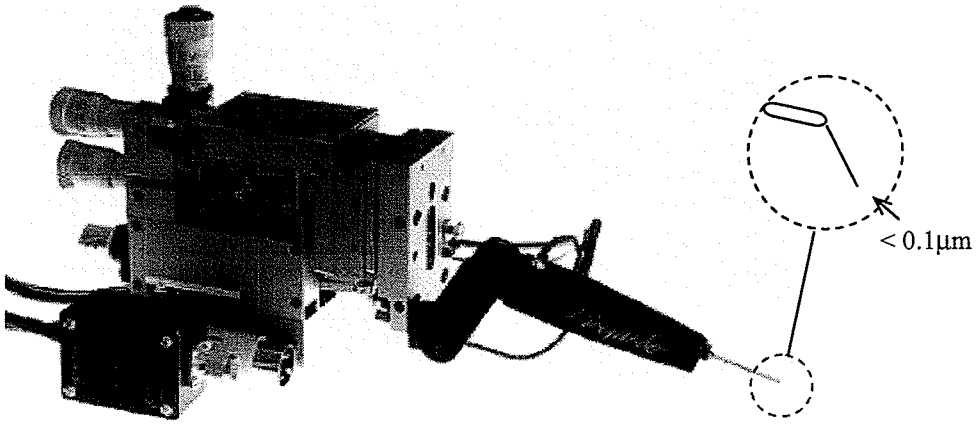

Fig. 1.2. An active high impedance probe with a tungsten probing wire [4].

A contact probe can not be used for buried circuit traces without using a time-consuming and potentially destructive tool such as a focused ion beam (FIB) to "deprocess" the circuit. As alternatives, non-contact internal probing techniques, such as electro-optic probing [7], electron-beam probing [8], near infrared optical probing [9,10], and scanning probe microscopy (SPM) based techniques such as magnetic force microscopy [11] and electrostatic force microscopy [12,13,14], ultrafast scanning tunneling microscopy [15] and photoconductive sampling probing [16], have been developed in the last decades. In this section, all of the mentioned probing techniques will be reviewed.

1.1.1 Electro-optic probing

Electro-optic probing techniques are based on the well known electro-optic effect, the Pockels effect, i.e. the refractive index of an electro-optic active material is changed when an electrical field is applied across the medium [17]. As a result, the polarization of the light passing through the material is changed in proportion to the electrical field. Electro-optic probing techniques have been extensively used in internal probing of high speed digital integrated circuits [18] and field mapping of monolithic microwave integrated circuits [19].

There are two types of electro-optic probing systems, internal (direct) and external. Direct electro-optic probing requires the device under test to be fabricated on a substrate that is an electro-optic active material, such as GaAs [7]. External electro-optic probing utilizes an electro-optic crystal as a probe situated at a close distance above the circuit

test point and focuses a pulsed laser beam is on the front side of the circuit through the crystal as shown in Fig. 1.3 [7]. A change in polarization of the laser beam will be imposed by the electrical field across the substrate in the internal probing system or by the electro-optic crystal in the external probing system. The change in polarization, $\Delta\varphi$, is proportional to the local potential, v_c. A synthesizer was used to phase-lock the pulsed laser train with the circuit stimulus so that measurements of the potential amplitude and phase can be performed [20].

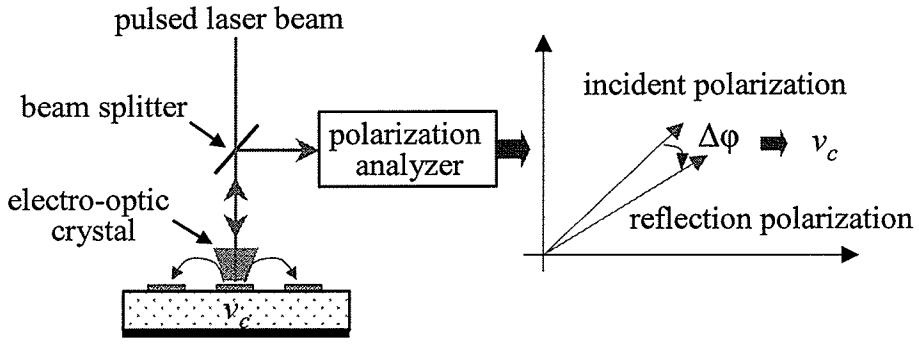

Fig. 1.3. A diagram of indirect electro-optic probing over a circuit test point..

Electro-optic probing techniques can offer extremely high temporal resolution, allowing the characterization of signals with rise times as fast as 1 ps when using sub-picosecond duration laser pulses [7]. The minimum detectable voltage for an electro-optic probing system is about 1 mV and a sensitivity is about 40 mV/$\sqrt{\text{Hz}}$ [20]. The spatial resolution of this method is optically limited to a few micrometers of the laser spot size and is sensitive to cross-talk effects [21]. The capacitive loading depends on the size of the electro-optic crystal and its distance to the circuit in the external electro-optic probing system. Electro-optic probing techniques require careful calibration and complicated circuit setup procedures for absolute value measurements since they are based on indirect measurement of local potential.

1.1.2 Electron-beam probing

Electron-beam probing uses a focused beam of electrons as a probe for the potential measurements at internal nodes of an IC [8]. The technique has been well developed and widely utilized by the microelectronics industry. It provides reliable, real-time data acquisition in measuring precise timing events.

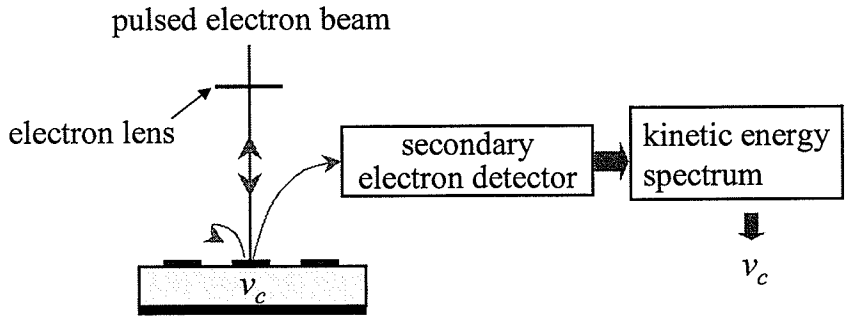

Fig. 1.4. A diagram of electron-beam probing.

Figure 1.4 shows a diagram of electron-beam probing (waveform acquisition mode). An electron gun is used to supply a pulsed primary electron beam which is focused onto a circuit test point. The primary electron beam causes secondary electrons to be emitted from the circuit. The kinetic energy of these secondary electrons, collected by a detector, is analyzed by a retarding grid spectrometer. The kinetic energy of the received secondary electrons depends linearly on the time-varying test point potential [8].

Commercial electron-beam probing system can provide up to 9 GHz bandwidth with +/- 40 ps accuracy and 2.5 ps timing resolution [22]. The spatial resolution of the system is usually limited by the pulsed electron beam spot size. By using a high precision thermal field emission electron beam and sophisticated beam management software, the e-beam probing system can provide sub-0.18 μm spatial resolution [22].

Electron-beam probing system is also complicated and very expensive. It requires a specialized measurement environment, such as a vacuum. The electron beam energy used for IC testing is usually less than 1 keV to avoid disturbing the operation of the circuit under test.

1.1.3 Near infrared optical probing

The flip-chip packaging method is commonly used for modern CMOS ICs [23]. Only the substrate side ("back side") of the package is physically accessible for studies of a flip-chip-packaged IC. Therefore, near infrared optical probing techniques have been developed for backside probing of flip-chip-packaged ICs. Two of the most widely used techniques for localizing faults and defects in flip-chip-packaged ICs are laser voltage probing (LVP) [9] and picosecond imaging circuit analysis (PICA) [10].

LVP technique uses fast laser pulses with 1.064 μm wavelength (1.17 eV photon energy) to probe individual transistors or diodes in CMOS ICs [9]. Since lightly doped silicon is transparent to light at wavelengths longer than 1 μm, it enables photons at these wavelengths to penetrate hundreds of microns into the silicon substrate. As shown in Fig. 1.5, the pulsed laser is focused on a working chip (a CMOS inverter) through the substrate. The reflected laser beam intensity modulation describes the time variation of the voltage on the test point [24].

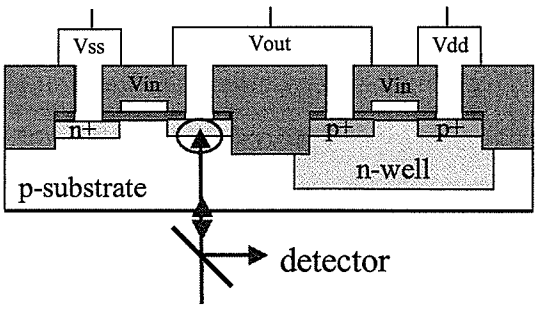

Fig. 1.5. The LVP probing a drain/substrate p-n junction from the substrate side [24].

The modulation of the reflected beam due to the circuit activity can arise from either the voltages across a junction in the gate, or by changes in the charge density with the electronic state of the gate [25]. Both effects produce intensity modulation in the reflected laser beam. By detecting the intensity modulation, the technique can be used to determine changes in voltage across the junction.

The reflected laser beam modulation from electrical activity is very weak ($1{:}10^6$) and difficult to detect [24]. Increasing the intensity of the incident laser beam can inject enough carriers into a device to perturb device performance [24]. The temporal resolution of the LVP is limited by the method for synching a mode-locked laser to the operating frequency of an IC to 20-50 ps [9]. It is also inadequate for probing device size < 0.5 μm due to the micron size of laser spot.

The PICA technique utilizes the naturally occurring light emission from switching field-effect transistors (FETs) to passively perform timing analysis as seen in Fig. 1.6 [10]. When switching the logic state, the FETs are briefly in saturation and a weak picosecond light pulse is emitted (due to intraband hot carrier emission), which can be detected from the substrate side.

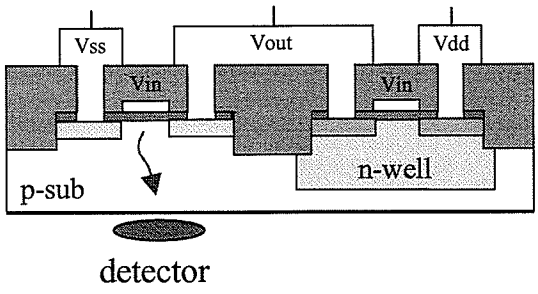

Fig. 1.6. The PICA probing from the substrate side [10].

The intensity of the light emission is very weak. In general, for currently available IC technologies and detectors, only one photon is detected for every 10^5 to 10^6 switching events. Measurements could last minutes to hours depending on duty cycle [10]. Time resolution in PICA is obtained from the time correlated photon counting technique. The current measuring technique provides adequate signal-noise to identify time shifts of the emission peaks on the 10 ps scale (Commercial PICA features at 2.1 GHz bandwidth and 10 ps timing repeatability, +/-15 ps timing accuracy [26]). FETs separated by less than the diffraction limit of conventional optics (microns) can often be separated because the emission occurs at distinct, non-overlapping times [10].

LVP and PICA benefit from advances in laser and electro-optic technology especially with regards to improvements in sensitivity and temporal resolution. Both are very expensive probing methods due to their complexities.

1.1.4 Scanning probe microscopy and magnetic force microscopy

Since 1982, the scanning probe microscope (SPM) has been widely used for extracting the topographical, thermal, optic, mechanical, electrical, and magnetic properties of microscopic materials [27]. The key component of a SPM is a microfabricated cantilever with an extremely sharp tip that ends in only a few hundred atoms wide. SPM is capable of measuring in atomic scale which was first achieved with the scanning tunneling microscope (STM) [28].

In the last decade, several novel instruments based on noncontact scanning force probes have been developed for diagnosis and characterization of microelectronics [29]. Noncontact probing minimizes the deformation of surfaces and the disturbance of the normal operation of circuits. In noncontact probing, the SPM probe is held above the

sample with a typical tip-sample distance, r, of 10 nm – 1000 nm. At that distance, forces such as electrostatic and magnetic forces ($\sim 1/r^2$) are dominant (the Van Der Waals force dominates for the tip-sample distance < 10 nm) [30].

One of these instruments is the magnetic force microscope (MFM) [11] which is capable of measuring static or low frequency magnetic surface phenomena with very high spatial resolution. MFM consists of a magnetic tip (usually coated with ferromagnetic material such as nickel or cobalt for their strong magnetic dipole moments), which can be simply considered as a magnetic point dipole, $\boldsymbol{m}$, as shown in Fig. 1.7 [31]. The force on the cantilever is $\boldsymbol{F}_m = \mu_0(\boldsymbol{m} \cdot \nabla)\boldsymbol{H}$, where $\boldsymbol{H}$ is the magnetic field of the sample under test.

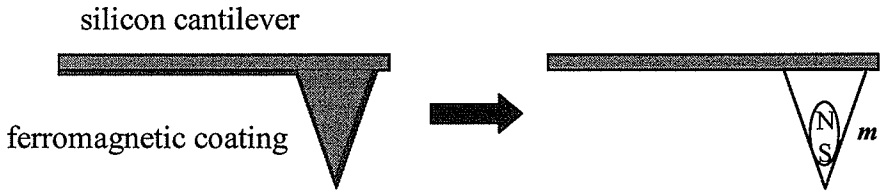

Fig. 1.7. A magnetic force probe with ferromagnetic coating [31].

MFM detects the magnetic force on the tip in order to map the field distribution within the sample. It is used for imaging magnetic domains in a wide range of materials, and is most suitable for imaging magnetic recording media which produce strong stray magnetic fields. MFM is also used for probing the internal current of ICs, especially for the detection of leakage currents [31].

Commercial MFM microscopes have spatial resolution of about 10 nm and a minimal detectable force better than 1 pN [32]. The state-of-the-art two dimensional current contrast imaging was performed on 100 nm wide conducting lines with a current as small as 1.7 μA (ac, with frequency lower or at the resonant frequency of the probe) [31]. MFM is usually operated in a vacuum to improve the measurement sensitivity.

Whereas the conventional MFM is bandwidth limited by the frequency response of the mechanical cantilever, high frequency MFMs (HFMFMs) have been used for imaging of the high frequency magnetic field on magnetic write-heads. Mainly there are two approaches for HFMFMs. One utilizes the direct phase detection method [33] and the other employs modulation techniques [34]. For HFMFM, a simple model of the magnetic behavior of the tip could be that of a dipole containing remanent and susceptible terms

$\boldsymbol{m}(\boldsymbol{H})=\boldsymbol{m}_0+\chi_t\boldsymbol{H}$, where $\boldsymbol{H}$ is the field from the sample and χ_t is the susceptibility of the tip moment [34]. HFMFMs have been used for evaluating the write-head field strength at frequency up to 1 GHz [34]. The prerequisite for the HFMFMs is that the tip magnetic moment must follow the external field switching speed synchronously.

1.1.5 Ultrafast scanning tunneling microscopy

Conventional STM applications have largely been restricted to the study of static surface properties (topography and electrical surface phenomena). The STM tip-metal sample tunneling current-voltage (*I-V*) characteristics are similar to those of a planar metal-insulator-metal junction. The nonlinearity of STM tunnel junction can be used for mixing ultrafast electrical transients to perform ultrafast time-resolved measurement. This technique is referred to as junction-mixing STM [15, 35].

To enable picosecond sampling, junction-mixing STM uses photoconductive switches to create ultrafast electrical pulses ($V_1(t)$ and $V_2(t-\tau)$) as seen in Fig. 1.8 (measurements were performed on a transmission line). A photoconductive switch is simply a slit between two metallic electrodes fabricated on a doped semiconductor material such as GaAs, which has a small energy-band gap to allow a fast electron-hole recombination time. When excited by short laser pulses, the extra carriers are created in the semiconductor, thus causing temporary conduction between the two electrodes and a short-duration electrical short circuit (a current is generated). Junction-mixing STM has demonstrated a picosecond temporal resolution combined with nanometer spatial resolution [35]. It is one of the promising tools for nanoelectronic applications.

STM relies on the electron tunneling between a sharp tip and the samples. The STM tip-to-sample distance is usually on the order of 1 nm. In the practical probing of an operating IC, possibly on a probe station, the tip-to-sample distance is usually quite large and the circuit under test could be covered by a passivation (or oxidation) layer. Complex system setup and calibration procedures are required for voltage measurement using STMs.

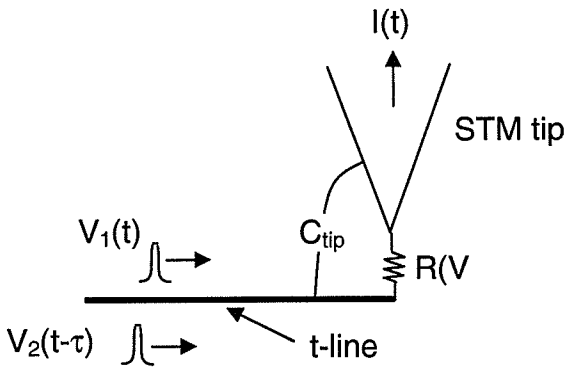

Fig. 1.8. Ultrafast STM probing a transmission line [15].

1.1.6 Photoconductive sampling probing

Photoconductive sampling probing technique uses a conductive probe with a micromachined interdigitated photoconductive switch near the probe tip to perform circuit measurement [16]. The photoconductive switch creates high speed electrical pulses when using an ultrashort pulse laser as an optical source. The probe makes direct contact with the test point to measure circuit signals. The technique demonstrated a 2.3 ps temporal resolution and 8 μm spatial resolution in [16].

The photoconductive sampling probe has similar limitations as with the previous ultrafast STM.

1.2 Motivation

In the past decade, electrostatic force microscopes (EFMs) have been introduced in performing microelectronics diagnostics [27][29]. EFMs measure the localized surface potential by monitoring the mechanical deflection of a small SPM probe as it responds to the Coulomb force interaction between the probe and the sample. EFM has been used for dc voltage measurements [36] and has achieved nanometer spatial resolutions [37].

The measurement bandwidth of the conventional EFMs, however, is limited by the mechanical frequency response of the probe (typically <100 kHz), which is far below the operating speed of current ICs. To overcome the bandwidth limitation, dynamic mode electrostatic force microscopes (DEFMs) [12,13,14,38,39] have been developed by employing the nonlinear dependence of the coulomb force interaction on the probe-

circuit voltage difference. This thesis concerns dynamic mode electrostatic force microscopy. The goal is to improve the spatial and temporal resolution in its application for the time-resolved measurement of the internal voltages of high-speed ICs. An instrument based on DEFM is developed and characterized. The instrument performance is demonstrated by the measurement of various high-speed ICs.

1.3 Outline of the Thesis

The following chapter (Chapter 2) describes the operation of the various components in EFM along with the principle of EFM and DEFM. Particularly, various DEFM pulse sampling techniques are summarized and briefly compared with their advantages and disadvantages.

Chapter 3 describes details of a DEFM pulse sampling instrument providing improved bandwidth. The bandwidth of the instrument is improved by using the high-speed pulse generation system to generate sampling pulses with pulse width as small as 37 ps. Pulse sampling measurements of signals up to 10 Gbit/s (improved from 3.2 Gbit/s as presented by previous researchers) on a matched transmission line and a Nortel wideband distributed amplifier are presented.

Chapter 4 characterizes the DEFM pulse sampling instrument performance, such as bandwidth, temporal resolution, phase shift linearity, noise, voltage sensitivity and capacitive loading. It describes a noniterative deconvolution process which could be used to compensate the waveform distortion caused by the nonideality of the sampling pulse.

Chapter 5 describes details of a null-force technique, which is successfully adapted in this research to enable the instrument to achieve quantitative voltage measurements in conjunction with very high bandwidth measurements. A method of performing precise voltage measurements, which are insensitive to dc offset effects, is presented as well. Quantitative voltage measurements of signals of 250 Mbit/s (1-2 Mbit/s as presented by other researchers) on a matched transmission line and a Nortel wideband distributed amplifier are presented.

In Chapter 6, the spatial resolution of the instrument using a standard scanning force microscopy probe is investigated experimentally and theoretically. Studies show that the probe is unable to achieve sub-micron spatial resolution in practical probing of an

operating IC, due to the large tip-to-sample distance (typically >100 nm) required when operating on a probe station in ambient conditions or when a circuit passivation layer is present. This could result in strong coupling of interfering signals to the probe tip sidewall and the cantilever and causes incorrect waveform and quantitative voltage measurements.

In Chapter 7, two dynamic electrostatic force-gradient microscopy approaches are proposed. One uses a numerical force-gradient detection approach, in which the force gradient is extracted numerically from directly sensing the force induced deflections for two tip-to-sample distances. The second approach enables direct detection of high-frequency force-gradients by employing a high-frequency implementation of a mechanoelectric cross-modulation scheme. Measurements on three parallel interconnects of an IC operating up to 500 Mbit/s demonstrate a submicron spatial resolution and a significant reduction of interference from adjacent interconnect signals by using the force-gradient detection approaches.

Chapter 8 contains the conclusions and future considerations.

Chapter 2

Electrostatic Force Microscopy

This chapter describes the operation of the various components in EFM along with the principle of EFM, DEFM, and in particular the pulse sampled DEFM for high frequency voltage measurements.

2.1 The operation principle of EFM

A schematic of voltage measurement using EFM is shown in Fig. 2.1. A conducting probe cantilever, which has a very sharp tip at its end, is held in close proximity to the point on the test circuit surface at which the voltage $v_c(x,y,t)$ is to be measured. To extract the voltage information, a specific voltage $v_p(t)$ is applied to the probe tip to produce a potential difference between the probe tip and the circuit test point, where a small localized capacitor is formed. The instrument operates by sensing the mechanical deflection of the probe due to the Coulomb force between the tip and the test point. This force can be simply derived from the energy required to charge up the capacitor (i.e. $E = \frac{1}{2} CV^2$ and $F_z = \frac{\partial E}{\partial z}$) and can be expressed as [12]

$$F_z = \frac{1}{2}\frac{\partial}{\partial z} C(x,y,z)[v_p(t) - v_c(x,y,t) + \Delta\Phi]^2, \tag{2.1}$$

where $C(x,y,z)$ is the localized capacitance between the probe tip and the circuit test point and the term $\Delta\Phi$ represents dc offset effects, such as that due to material work function differences [37].

By sensing the cantilever deflection Δz, induced by the electrostatic force, information about the test point signal $v_c(x,y,t)$ can be obtained. As seen in Fig. 2.1, a laser beam-bounce detection system [40] is used for sensing the small cantilever deflection.

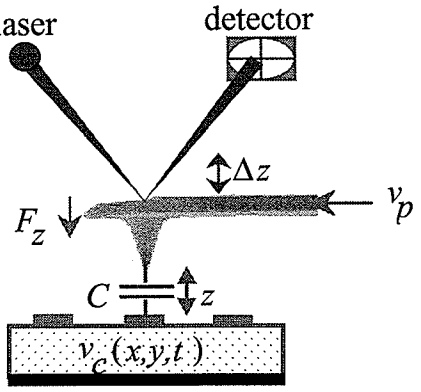

Fig. 2.1. A schematic of the voltage measurement in EFM.

The magnitude of the cantilever deflection depends on the mechanical properties of the cantilever and the force acting on the cantilever. Moreover, the force in (2.1) is proportional to the derivative of $C(x,y,z)$ which includes coupling capacitance between the tip and the circuit test point as well as other coupling capacitances from the contribution of the tip sidewall and the large cantilever. Thus, in order to determine the instrument sensitivity and spatial resolution, it is necessary to examine the mechanical probe. The next section contains a detailed study of the mechanical properties of the probe. The analysis of the instrument sensitivity and the spatial resolution of the probe are postponed to Chapter 4 and Chapter 6.

2.2 The mechanical properties of the probe

A suitable cantilever for internal IC probing should have a high resonance frequency and quality factor, with a small spring constant. A cantilever with higher resonant frequency is less susceptible to low frequency noise. A cantilever with smaller spring constant results in a cantilever deflection system that is able to detect a smaller force. Moreover, the smaller cantilever with a taller tip and a smaller tip apex improves the spatial resolution and the parasitic loading of the probe. These are the primary reasons for choosing small microfabricated cantilevers.

The cantilever (micromachined using electrochemical etching) employed in this thesis was a commercial silicon cantilever from NT-MDT, i.e. CSC12 [41]. The rectangular cantilever is 350 μm long, 30 μm wide and 1.2 μm thick with a 10 μm high conical tip at its end. The cone angle of the tip is approximately 40°. The radius of the tip

curvature is less than 35 nm. The scanning electron microscopy image of the tip is shown in Fig. 2.2. The entire cantilever is coated with a thin W_2C layer (~25 nm) for its conductivity and hardness.

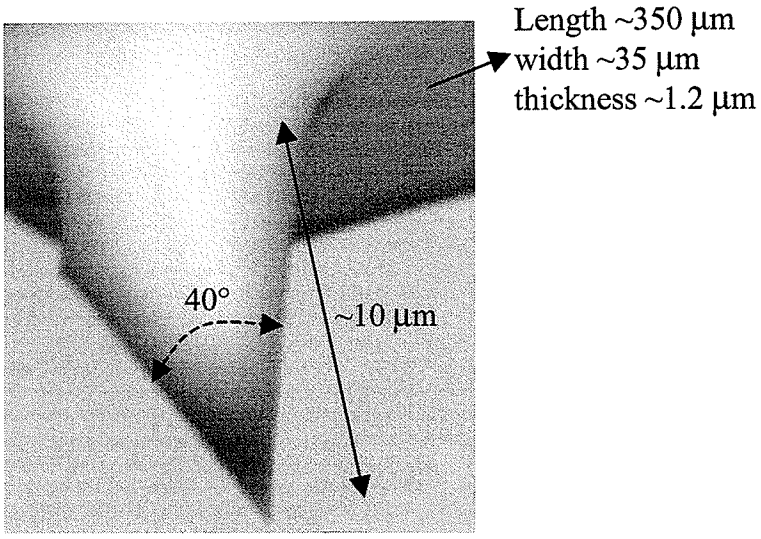

Fig. 2.2. A scanning electron microscopy image of the silicon cantilever tip.

2.2.1 The spring constant

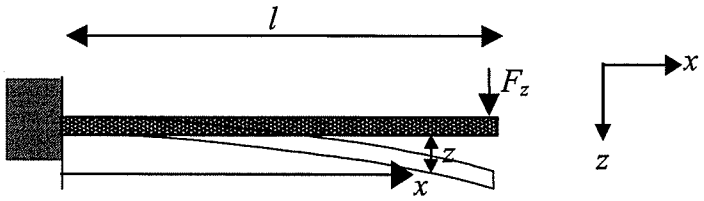

Fig. 2.3. A simple mode for the cantilever under a constant applied force.

A rectangular cantilever with one end fixed and the other end free vibrates like a spring as shown in Fig. 2.3. Assuming the cantilever has a uniform cross section, the spring constant, k, defines the rigidity of its vibration, and is given by [27]

$$k = \frac{Ewt^3}{4l^3}, \text{ N/m} \tag{2.2}$$

where l, w, t are the length, width, and thickness of the cantilever and E is the modulus of elasticity. For the silicon cantilever used in this work, E=1.79×10^{11} N/m^2, l=350 μm,

w=35 μm and t~1 μm. The spring constant k is calculated to be approximately 0.04 N/m. k is typically 0.03 N/m from its manufacturer's specifications [41].

The deflection z for each point $x \le l$ along the cantilever due to a constant force F_z applied at the end of the cantilever, is given by [27]

$$z = \frac{F_z}{2kl^3} x^2 (3l - x), \qquad x \le l \tag{2.3}$$

when a simple static mode is used for the cantilever (neglecting the tip since it is much smaller than the cantilever).

2.2.2 Mechanical frequency response

The mechanical frequency response of the cantilever is essential for describing its vibration behavior under an external driving force.

2.2.2.1 Free vibration

For the cantilever in Fig. 2.3, the homogeneous equation of free motion with flexural small vibrations is a differential equation of the fourth order [42]:

$$EI \frac{\partial^4 z}{\partial x^4} + \rho A \frac{\partial^2 z}{\partial t^2} = 0, \tag{2.4}$$

where E is the modulus of elasticity, $I = wt^3/12$ for the rectangular cross section is the area moment of inertia (μm^4), ρ is the mass density (kg/m^3), and $A = wt$ is the cross sectional area (μm^2) of the cantilever.

Applying boundary conditions $\{z(0)=0$, $(d/dx)z(0)=0$, $(d^2/dx^2)z(l)=0$ and $(d^3/dx^3)z(l)=0$ [42]$\}$ to Eq. (2.4), it can be solved only if the following characteristic equation

$$\cos(\kappa_n l)\cosh(\kappa_n l) + 1 = 0, \tag{2.5}$$

is satisfied, where κ_n is the wave number of the vibration mode n. The solutions of this equation are $\kappa_n l \approx \{1.875, 4.694, 7.855, ...\}$ [43]. The natural resonance frequencies of the cantilever can be calculated from κ_n:

$$\omega_n = \sqrt{\kappa_n^4 EI / \rho A}. \tag{2.6}$$

For the rectangle cantilever of Fig. 2.2 with ρ = 2330 kg/m^3 for the silicon, the fundamental mode ω_1 is calculated to be 11.56 kHz. The 2nd resonant frequency $\omega_2 = 6.26\omega_1$ = 72.36 kHz.

The cantilever's first two vibration modes (eigenfunctions of the Eq. (2.4)) are shown in Fig. 2.4, where the position is normalized to the length of the cantilever (l) and A_n represent the normalizing amplitudes [42]. Only the first vibration mode of the cantilever is used in this work.

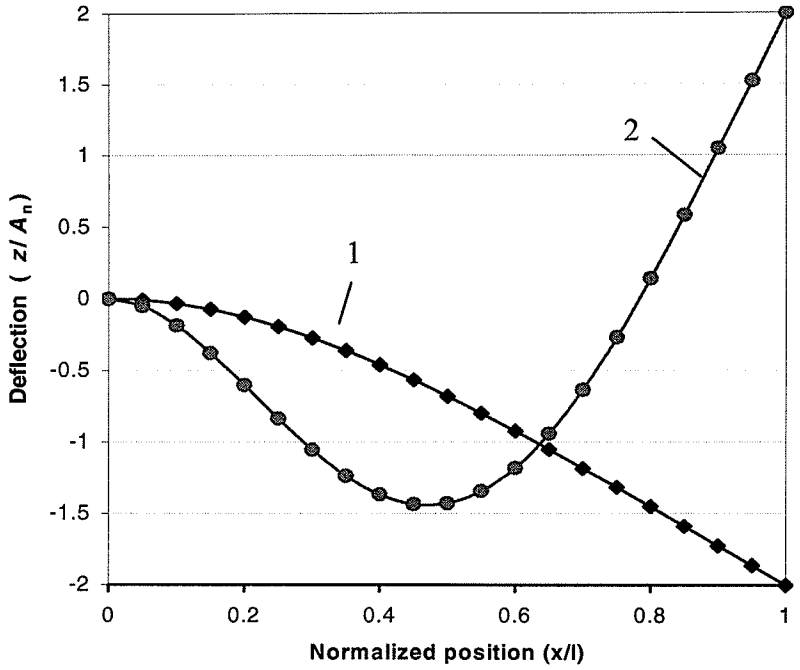

Fig. 2.4. The first two vibration modes of the cantilever.

Using a point-mass model, the cantilever can be replaced by a point mass m^* attached to a spring with spring constant k. m^* is determined such that the resonance frequency $\sqrt{k/m^*}$ equals to the lowest resonance frequency calculated from Eq. (2.6). For a rectangle cantilever, the effective mass m^* is ~ $0.24m$, where $m=\rho Al$ [42]. Using m^*, the fundamental resonance frequency of the cantilever with an attached tip (its mass: m_{tip}) is

$$\omega_1 = \sqrt{\frac{k}{m^* + m_{tip}}} = \sqrt{\frac{k}{0.24m + m_{tip}}}. \tag{2.7}$$

2.2.2.2 Forced vibration

Usually the cantilever is excited near or at one of the resonant frequencies. The forces acting on the cantilever are the distributed weight of the cantilever, the weight of the tip, the externally applied force (electrostatic force in EFM), the spring force of the cantilever which is opposing the motion, and a retarding force which is proportional to the speed of

the cantilever due to the air friction and the internal friction. It can be modeled as an underdamped mechanical oscillator with a single resonant frequency ω_1. The equation of motion of the cantilever with damping is

$$m_T \frac{d^2 z}{dt^2} + \gamma \frac{dz}{dt} + kz = W_T + F_z(t), \tag{2.8}$$

where $m_T = m^* + m_{tip}$, W_T is the contribution due to the weight of the cantilever at the tip plus the weight of the tip, and γ is the damping factor. The cantilever's equilibrium position, z_o, is determined by the total weight W_T ($z_o = W_T/k$, ~0.6 nm for the silicon cantilever).

Using the Fourier transformation, the steady state solution of (2.8) in the frequency domain is

$$-m_T \omega^2 Z(\omega) + j\omega\gamma Z(\omega) + kZ(\omega) = \boldsymbol{F}_z(\omega), \tag{2.9}$$

where j is the complex $\sqrt{-1}$, $Z(\omega)$ and $\boldsymbol{F}_z(\omega)$ are the Fourier transforms of z and $F_z(t)$, respectively.

The cantilever deflection, $Z(\omega)$, is determined by the applied force, $\boldsymbol{F}_z(\omega)$, and the mechanical frequency response of the cantilever, $H(\omega)$, as seen in Fig. 2.5.

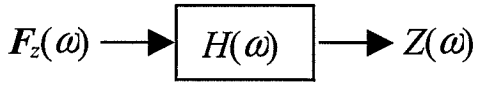

Fig. 2.5. The mechanical frequency response of the cantilever.

The mechanical frequency response of the cantilever, $H(\omega)$, can be derived from Eq. (2.9) as

$$H(\omega) = \frac{Z(\omega)}{\boldsymbol{F}_z(\omega)} = \frac{1}{(k - m_T\omega^2) + j\omega\gamma} = |H(\omega)|e^{j\theta}, \tag{2.10}$$

where

$$|H(\omega)| = \frac{1}{\sqrt{(k - m_T\omega^2)^2 + (\omega\gamma)^2}}, \quad \theta(\omega) = -\tan^{-1}\frac{\omega\gamma / m_T}{\omega_1^2 - \omega^2}. \tag{2.11}$$

$|H(\omega)|$ is the cantilever deflection amplitude normalized to the applied force amplitude and $\theta(\omega)$ is the vibration phase relative to that of the force.

$|H(\omega)|$ is maximum for a particular value of frequency, ω_r, which is called the resonant frequency of the cantilever under the influence of external and damping forces. This occurs at $d|H(\omega)|/d\omega = 0$ and is

$$\omega_r = \sqrt{\omega_1^2 - \frac{\gamma^2}{2m_T^2}}. \tag{2.12}$$

When the external applied force $F_z(t)$ is zero, the cantilever vibration function is

$$z(t) = z_0 e^{-(\gamma/2m_T)t} \cos(\omega_{r0} t + \phi_0), \tag{2.13}$$

where z_0, ϕ_0 are the initial vibration amplitude and phase, and $\omega_{r0} = \sqrt{\omega_1^2 - \frac{\gamma^2}{4m_T^2}}$.

Figure 2.6 shows a sample of $z(t)$ with $\gamma/2m = \omega_0/100$.

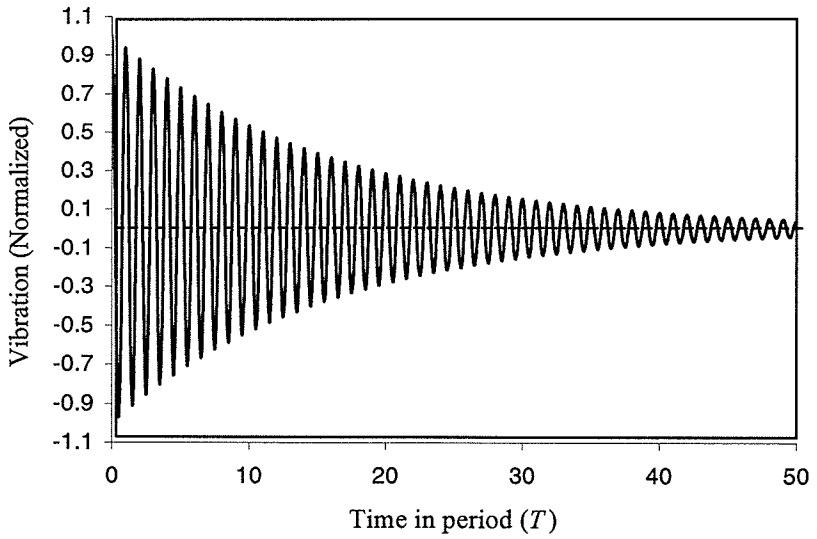

Fig. 2.6. The normalized vibration of the cantilever as a function of time.

The vibration energy loss rate is usually used to define the quality factor of the cantilever, which is approximately the number of cycles needed for the vibration amplitude decreased to $1/e$ of the initial amplitude in the Fig. 2.6. It can be expressed in terms of the damping factor γ as [44]

$$Q \sim \frac{m\omega_{r0}}{\gamma}. \tag{2.14}$$

Since Q is usually large (Q~100/2π=16 in Fig. 2.6), $\omega_0 \approx \omega_r \approx \omega_1$. The frequency response of the cantilever Eq. (2.11) can then be rewritten to be

$$|H(\omega)| = \frac{Q/k}{\sqrt{1+Q^2(1-\frac{\omega^2}{\omega_r^2})^2}}, \quad \theta(\omega) = -\tan^{-1}\frac{\omega/Q\omega_r}{(1-\frac{\omega^2}{\omega_r^2})}. \tag{2.15}$$

At the resonance frequency $\omega=\omega_r$, $|H(\omega)|$ reaches its maximum as $|H(\omega_r)|=Q/k$ and the vibration phase $\theta(\omega_r) = -90°$. The vibration amplitude is enhanced Q times in comparison with the static case

$$A_z(\omega_r) = \frac{Q}{k}|F_z(\omega_r)|. \tag{2.16}$$

A typical frequency response for the cantilever in Fig. 2.2 with Q = 16 is shown in Fig. 2.7 in comparison with the measured cantilever frequency response (normalized frequency ω/ω_r). The ratio between the resonant frequency and 3 dB full bandwidth (at which $|H(\omega)|/|H(\omega_r)|$ = 0.707 as shown in Fig. 2.7), $\omega_r/\Delta\omega$, is approximately equal to the quality factor of the cantilever Q. From Eq. (2.14), Q depends upon mechanical characteristics of the cantilever and the damping properties of the media.

The measured frequency response of the cantilever in Fig. 2.7 was obtained by driving the probe with a sinusoidal voltage signal. The frequency of the driving signal was varied and the deflection signal due to the electrostatic force was detected by the deflection detector and recorded by a scope. The measured resonant frequency is approximately 11.6 kHz. The quality factor of the cantilever Q is approximately 16.

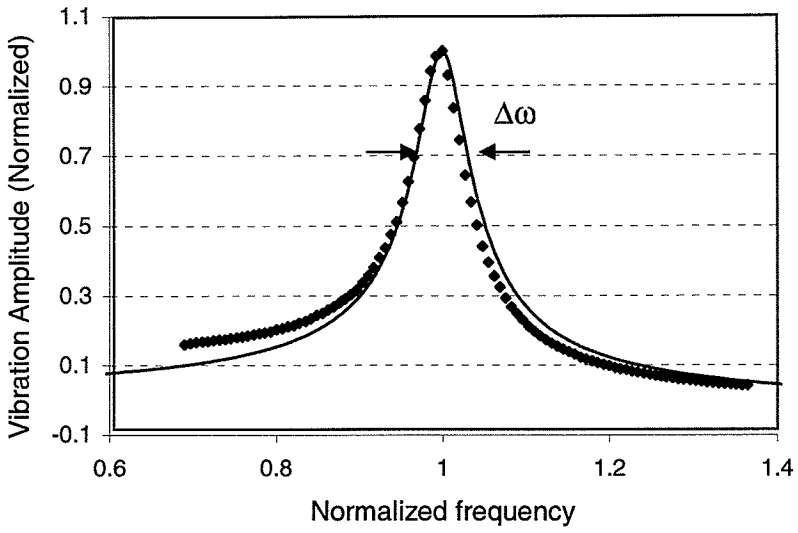

Fig. 2.7. The theoretical frequency response (the thin curve) is calculated with $Q = 16$ in comparison with the measured frequency response (dots) of the cantilever.

2.3 Optical deflection detection

The typical magnitude of cantilever deflection in electrostatic force probing is very small. Assuming 1 pN force (a typical value when probing a circuit as discussed in Chapter 6 and 7) applied to the silicon cantilever near the tip, the corresponding deflection is less than 1 nm using Eq. (2.16) (for $k = 0.05$ N/m and Q=16). To detect such a small deflection, an optical beam deflection technique [40] is used in this work.

The optical beam deflection technique measures the cantilever vibration by detecting the deflection of a laser beam reflected from the cantilever surface. As shown in Fig. 2.8, a laser source directs a focused laser beam onto the surface of the cantilever free end. The reflected laser beam from the cantilever is captured by a carefully positioned bi-cell photodiode.

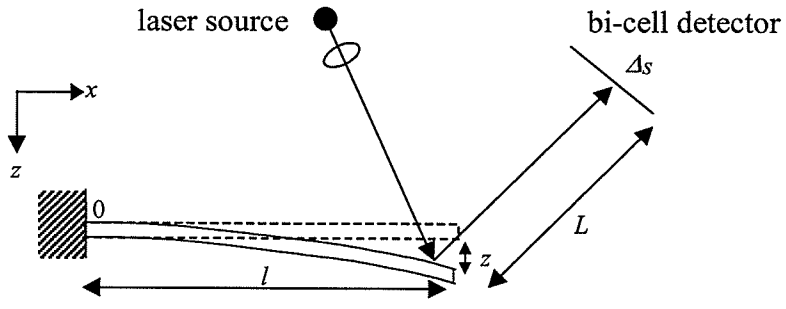

Fig. 2.8. An optical beam deflection system for detecting the cantilever vibration.

In the cantilever free vibration mode, the bi-cell photo diodes are placed such that the reflected laser beam is incident on both photo diodes with equal optical power (original laser spot). An applied force pulls or pushes the cantilever away from its equilibrium position. The deflection Δz causes the reflected laser beam to be shifted by Δs at the bi-cell detector as [45]

$$\Delta s \approx \frac{3L}{2l}\Delta z, \tag{2.17}$$

where L is the distance between the cantilever and the center of the bi-cell photodiode. More optical signal is accepted by one cell of the photo diode than the other cell as shown in Fig. 2.9 (the dotted circle is original laser spot).

The deviation in optical signals absorbed between the two photo diodes is proportional to the shift Δs and the corresponding deflection Δz [40]. The optical beam detection system is usually a linear system and its output signal is amplified by a low noise operation amplifier with a flat frequency response up to 1 MHz, which is much higher than the mechanical resonance frequency of the cantilever. This method enhances the signal-to-noise ratio (SNR), resulting in increased sensitivity of the vibration detection. The sensitivity of this method has been reported to be 7.9×10^{-6} nm/$\sqrt{\text{Hz}}$ [46].

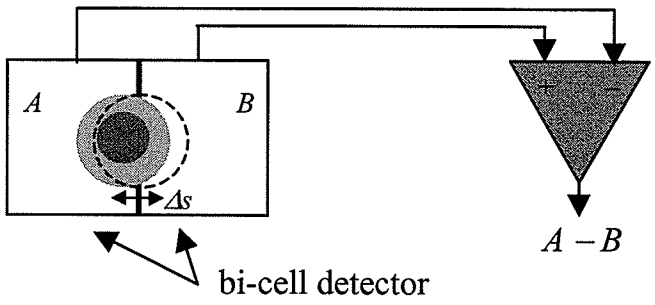

Fig. 2.9.The operation concept of the bi-cell detector for detecting the deflection signals.

2.4 Mechanical probing structures

The probe and its holder should be physically compact and firmly integrated into a mechanical structure with the optical deflection detection system in order to achieve

internal access to the test structures. These test structures are often mounted inside a deep and small cavity of a test fixture or other IC package.

The probing structure compatible for use on a commercial probe station was designed by Micron-Force Instruments (MFI) [47] and has incorporated automatic alignment of the optical beam-bounce system as shown in Fig. 2.10. The aluminum probing structure contains a probing head, a 3D piezoelectric positioner and some parts of the optical deflection detection system (including a laser diode, a prism as beam splitter and a bi-cell detector). This structure uses a magnetic positioner to be mounted on a probing station which is supported by a vibration isolation table to isolate low frequency ground noises. The probe station includes an optical microscope and a 3D manual circuit positioner for positioning the probe over the device-under-test (DUT).

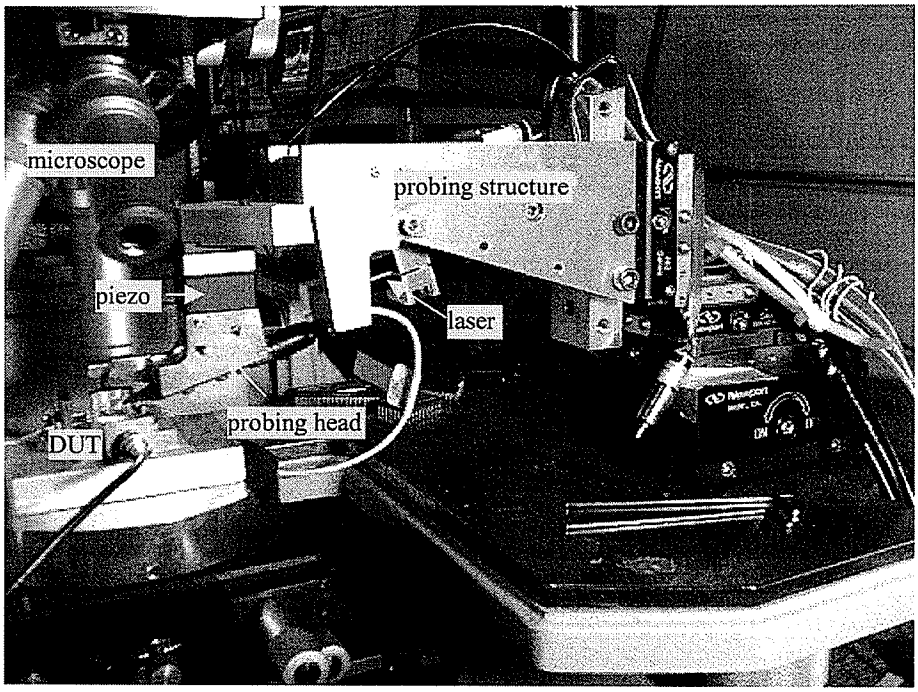

Fig. 2.10. The first probing structure holds the cantilever, the optical beam-bounce system, the probe positioning system and the laser spot positioning system.

The optical paths of the laser beam deflection system are shown in Fig. 2.11. The laser diode on the probing structure focuses its beam through a prism which acts as a light beam splitter onto the mirror which directs the beam directly onto the cantilever. The cantilever is usually tilted 15° to provide separation of the cantilever supporting block from the DUT and to allow optical access from above for the microscope as well. The

reflected beam off the end of the cantilever travels back to the prism which redirects the beam to the bi-cell photodiode.

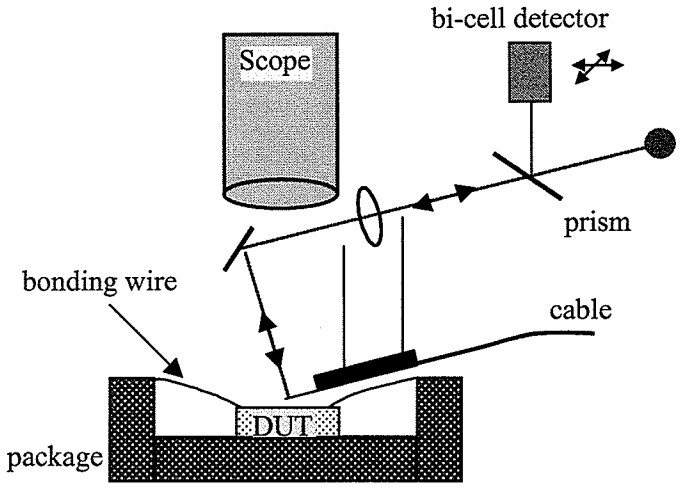

Fig. 2.11. The optical paths of the probing system. The probing head is mounted on a 3-D piezoelectric positioner for positioning the cantilever.

The probing head, which holds the lens, the mirror and the printed circuit board, is mounted on a 3D piezoelectric positioner for finely adjusting the cantilever position or for scanning the circuit. The printed circuit board is usually a 50 Ω transmission line with the cantilever supporting block mounted on one end of the transmission line while the other end of the transmission line is connected to a flexible coaxial cable for the propagation of probe-driven high frequency electrical signals to the cantilever. The printed circuit board is screwed onto the bottom of the probing head.

The probing structure has limited accessibility for probing inside a deep circuit package cavity. As seen in Fig. 2.11, the printed circuit board interferes with the sidewall of the package and the bonding wires. To solve this problem, the probing head was redesigned with a much larger tilted angle (60°) for the cantilever supporting board as shown in Fig. 2.12. The flexible coaxial cable is fixed at the bottom of the board. It is also directly connected to the cantilever supporting block (which is tilted with an angle of 15°) by using two-component epoxy, which is electrically conductive because of a high density of silver particles.

A computer with software provided by MFI was used to receive and display the signals from the bi-cell photodiode. The original position of the laser beam spot on the bi-

cell photodiode (free vibration) can be achieved by adjusting the laser and the bi-cell photodiode with manual knobs and screws.

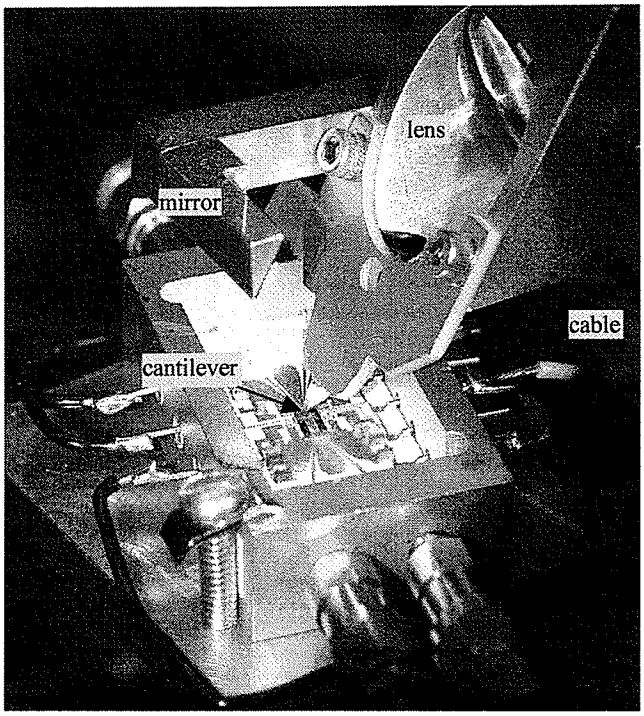

Fig. 2.12. A redesigned probing head for deep cavity access.

2.5 DC voltage probing

The electrostatic force microscopy technique can be employed to measure DC voltages on a circuit, $V_c(x,y)$. From Eq. (2.1), the force on the cantilever due to this DC voltage is

$$F_z = \frac{1}{2}\frac{\partial}{\partial z}C(x,y,z)[v_p(t) - V_c(x,y) + \Delta\Phi]^2. \tag{2.18}$$

Directly extracting V_c by sensing the electrostatic force on the probe requires the knowledge of the mutual capacitance gradient between the tip and the circuit, which requires precise and bothersome cantilever positioning. The concept of force nulling by applying an ac signal to the tip was conceived for overcoming this problem [36]. Since static force measurement using dc force nulling does not work well due to the large low frequency noise, a modulation technique was typically employed to avoid force

measurement at dc. This allows for the use of a lock-in amplifier to improve overall SNR and is accomplished by driving the cantilever with an ac signal [36]

$$v_p(t) = A + K\cos(\omega t). \tag{2.19}$$

where A is a dc bias, K is the ac amplitude and both are adjustable. The frequency ω is usually chosen at the fundamental resonance of the cantilever, ω_r to enhance deflection amplitude by a factor of Q according to Eq. (2.16). The resulting force has a dc component and a frequency component at ω_r, and $2\omega_r$. The static deflection caused by the dc force is not required. The deflection caused by the force component at ω_r is [36]

$$\Delta z(t)\big|_{\omega_r} = \frac{\partial C}{\partial z}\frac{Q}{k}[A - V_c(x,y) + \Delta\Phi]K\sin\omega_r t. \tag{2.20}$$

From this equation, the circuit voltage is known as $(A+\Delta\Phi)$ when the force induced deflection is null. $\Delta\Phi$ is unknown and usually small (<100 mV [37]). The measured dc voltage is invariant to the value dC/dz, Q or k, thus eliminating the need for complex calibration and accurate probe positioning above the circuit test point.

The small cantilever deflection signal can be measured by using a lock-in amplifier, which is commonly used for recovering a signal of interest with a well-defined frequency, in the presence of an overwhelming "noise" background. A lock-in amplifier uses a reference signal as seen in Fig. 2.13 to select the weak steady signal of interest (in our case the deflection of the cantilever at the frequency ω_r) in the input signal and down-convert to baseband (at dc). The baseband signal is then processed through a low pass filter followed by a dc amplifier (gain: G_{dc}) [48]. The output signal includes the information of the amplitude (V_{in} in Volts rms) and the phase (ϕ_1 in Degree) of the input signal of interest. This inherent tracking ability allows extremely small bandwidths to be defined for the purpose of signal-to-noise improvement.

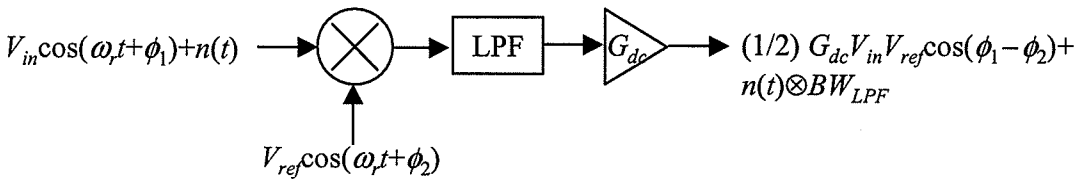

Fig. 2.13. A lock-in amplifier block diagram.

2.6 Vector voltage probing

The objective of vector voltage measurement is to acquire the $V_c \angle \phi_c$ of a high frequency sinusoidal signal at a certain position in a circuit:

$$v_c(x,y,t) = V_c \cos(\omega_c t + \phi_c). \tag{2.21}$$

where ω_c is usually much higher than the resonant frequency of the cantilever.

The heterodyne mixing technique [13] is used to generate electric force harmonics that fall within the cantilever mechanical response. To do that, the non-linear Coulomb force interaction (depending on the square of the probe-circuit potential difference) is employed. Frequency down conversion is realised by driving the probe with an amplitude modulated high-frequency voltage signal as [13]

$$v_p(t) = (A + K\cos\omega_r t)\cos(\omega_c t + \phi_p), \tag{2.22}$$

where A, K, and ϕ_p are adjustable parameters. This signal mixes with the circuit signal (according to Eq. (2.1)) right at the probe tip. The intermediate frequency signal (at ω_r) can be measured by the cantilever. If ω_r is chosen to be the cantilever resonance frequency, the cantilever deflection will be [13]

$$\Delta z(t)\big|_{\omega_r} = \frac{1}{2}\frac{\partial C}{\partial z}\frac{Q}{k}[A - V_c\cos(\phi_p - \phi_c)]K\sin\omega_r t. \tag{2.23}$$

In order to measure the magnitude and phase of $v_c(x,y,t)$, the force nulling technique should be once again applied to avoid complex calibration and accurate probe positioning for unknown values of dC/dz, Q or k. The technique involves cleverly adjusting the parameters A and ϕ_p to null the cantilever deflection induced by the electrostatic force frequency component at ω_r in order to identify V_c and ϕ_c. The nulling algorithm details are described in [49].

Figure 2.14 shows a block diagram of the vector voltage measurement for high frequency circuit voltage signal amplitude and phase measurement [50]. A high frequency signal source excites the circuit-under-test from its input port and triggers the probe signal $v_p(t)$ with an adjustable phase ϕ_p. The probe signal $v_p(t)$ is also amplitude modulated at the resonance frequency of the cantilever. The vibrations of the cantilever induced are detected by an optical detection system. The lock-in amplifier acts as a

narrow band pass filter at the cantilever resonance frequency ω_r to allow reading of the magnitude A_z of the cantilever deflection $\Delta z(t)$ at ω_r.

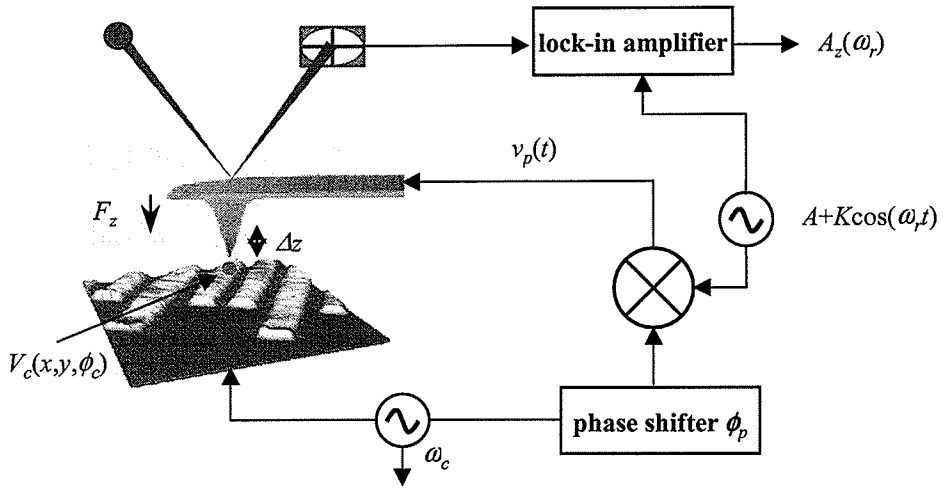

Fig. 2.14. A vector voltage measurement block diagram for the high frequency circuit voltage signal amplitude and phase measurement.

The vector voltage measurement technique is independent of the dc offset since $\Delta\Phi(x,y)$ is absent from Eq. (2.22). As an example, the EFM vector voltage measurement was applied to a coplanar waveguide transmission line and the circuit was operating at 25.6 GHz. The measured result is shown in Fig. 2.15, which demonstrates the linear dependence of the measured voltage amplitude to the circuit voltage amplitude [51].

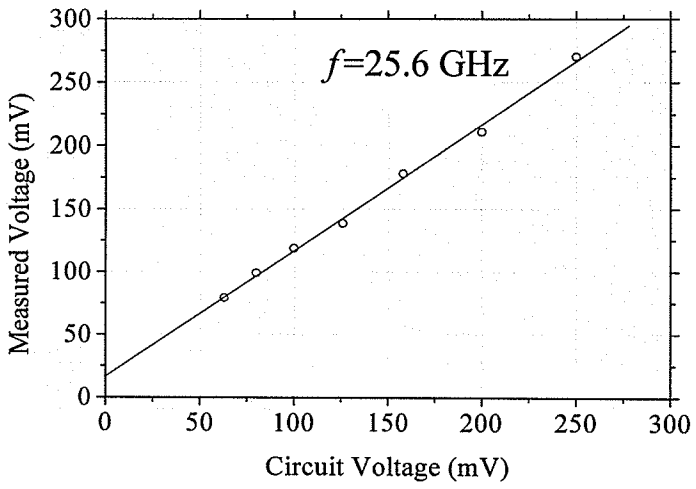

Fig. 2.15. The linear dependence of the measured voltage amplitude to the circuit voltage amplitude.

The vector voltage measurement technique is useful for the testing of analogue microwave monolithic integrated circuits (MMICs), helping resolve issues such as harmonic distortion, cross-talk, dispersion, transmission loss and spurious reflections [50]. However, this technique is unable to measure arbitrary circuit signals, which requires time-resolved measurement techniques. This is the main emphasis of the remaining work presented in this chapter.

2.7 Repetitive time sampling

For time-resolved waveform measurements, the waveshape, voltage transitions, and timing delay information of the signals are often the major interest in microelectronics design and failure analysis. Repetitive time sampling techniques are commonly used for time-resolved waveform measurements in many high-speed test apparatus, such as digital sampling oscilloscopes. Compared to real time oscilloscopes, digital sampling oscilloscopes have no need for high-speed data sampling and large data storage. High frequency repetitive signals are sampled by short pulses and characterized through the analysis of lower frequency signals, i.e. samples are collected from sequential occurrences of the high frequency signals and reassembled at a much slower rate [52].

Repetitive time sampling techniques have also been used in EFM to overcome the low pass filter behavior of the cantilever mechanical frequency response and to extend the dynamic range of the EFM probe [39][53]. High-speed sampling pulses are applied to the probe and mixed with circuit signals right at the tip due to the nonlinear electrostatic force interaction at that point. The cantilever behaves like a 'mixer' with the output of such a 'mixer' being the cantilever deflection.

There are two main approaches for the use of repetitive time sampling in EFM: a heterodyne equivalent-time sampling technique [39] and homodyne equivalent-time sampling technique, which is used in this work.

2.7.1 Heterodyne equivalent-time sampling

In the heterodyne equivalent-time sampling technique, a high-speed sampling pulse waveform is repetitively applied to the probe tip at a rate slightly different from the rate of the circuit signal using phase-locked loop devices as shown in Fig. 2.16 [39]. The

difference frequency, or the intermediate frequency (IF), is $\Delta f = f_c - f_p$, which is chosen to be well below the resonance frequency of the cantilever.

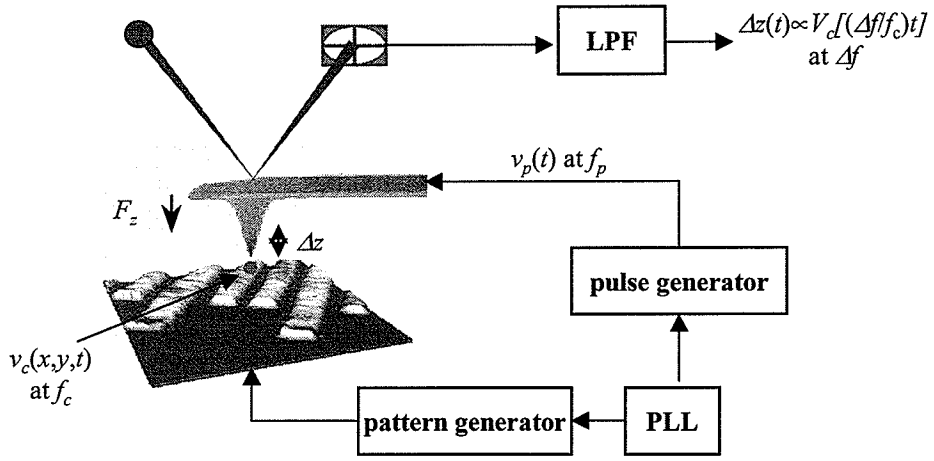

Fig. 2.16. An implementation of the heterodyne equivalent-time sampling technique.

In the heterodyne equivalent-time sampling technique, usually only one sampling point is measured for each waveform period. At the following waveform period, the sampling pulse arrives at a slightly later phase (time delay increment Δt, when the sampling rate is $1/\Delta t$). Over N periods ($N=1/\Delta f \Delta t$, sample number), the sampling pulse has gradually walked through the entire waveform, sampling successively delayed points. As a result, the averaged output intensity changes in proportion to the signal waveform, but at the much slower rate of Δf, as derived next.

If the sampling pulse waveform is $V_p G_\delta(t)$, where V_p is the amplitude of the pulse, and $G_\delta(t)$ is a unit amplitude pulse with pulse width δ. T_p ($=1/f_p$) is the pulse period. Due to the nonlinear dependence of the electrostatic force on the voltage difference, Eq. (2.1), the multiplication of the probe signal and the circuit signal causes the deflection of the cantilever to be

$$\Delta z(t) = \frac{1}{k}\frac{\partial C}{\partial z} V_p \int_{-\infty}^{+\infty} v_c(t_1) G_\delta(t - t_1) dt. \tag{2.24}$$

Assuming the pulse width $\delta << T_p$ and T_c, the deflection will be

$$\Delta z(t) = \frac{1}{k}\frac{\partial C}{\partial z} V_p \frac{\delta}{T_p} v_c\left(\frac{\Delta f}{f_c} t\right), \tag{2.25}$$

which is a low frequency representation of the circuit signal $v_c(t)$ [39].

The difference frequency (or IF) Δf has to be low enough for all of its relevant harmonics to lie within the flat region of the cantilever mechanical frequency response in order to produce non-distorted output signals. This means that the raw IF bandwidth of the system $N\Delta f$ must be less than f_r, the resonance frequency of the cantilever. At the same time, Δf must be higher than the $1/f$ noise. Other low frequency noise also may dominate the fundamental thermal noise. This means that the technique is not optimal for measuring long pattern waveforms, since a large number of sampling points (due to Nyquist criterion) are required. There is no such problem for the homodyne equivalent-time sampling techniques presented next.

2.7.2 Homodyne equivalent-time sampling

The homodyne equivalent-time sampling, which is used in this work, employs a modulated sampling pulse signal applied to the probe at a repetition rate similar to that of the circuit signal. The sampling pulse has to be manually walked through the entire waveform. The advantage of this approach is that only a single intermediate frequency component is measured at one time. In the homodyne approach, pulse modulation has been implemented using pulse amplitude modulation (PAM) [53] and pulse width modulation (PWM) [54][55]. This work focuses on the pulse amplitude modulation method.

2.7.2.1 Pulse amplitude modulation

A block diagram of the pulse amplitude modulation based frequency down conversion technique is shown in Fig. 2.17. The frequency down conversion is realized at the probe tip as a result of the nonlinear interaction between the probe and circuit signals according to Eq. (2.1). As illustrated in Fig. 2.17, this is accomplished by synchronising the periodic signal applied to the probe, $v_p(t)$, with the circuit test point signal, $v_c(x,y,t)$. As in a conventional IC test system, the circuit-under-test is driven externally by a digital pattern generator which provides N-bit length test vectors with a period $T=N/f_b$, where f_b is the system clock rate. One example of $v_c(x,y,t)$ is shown in Fig. 2.18 (a). The probe is driven by an amplitude modulated high bandwidth sampling pulse $v_s(t)=G_\delta(t-\tau)$ as shown in Fig. 2.18 (b) with a period of T_p which usually equals T ($T_p=T$). δ is the pulse width and τ is an adjustable time delay.

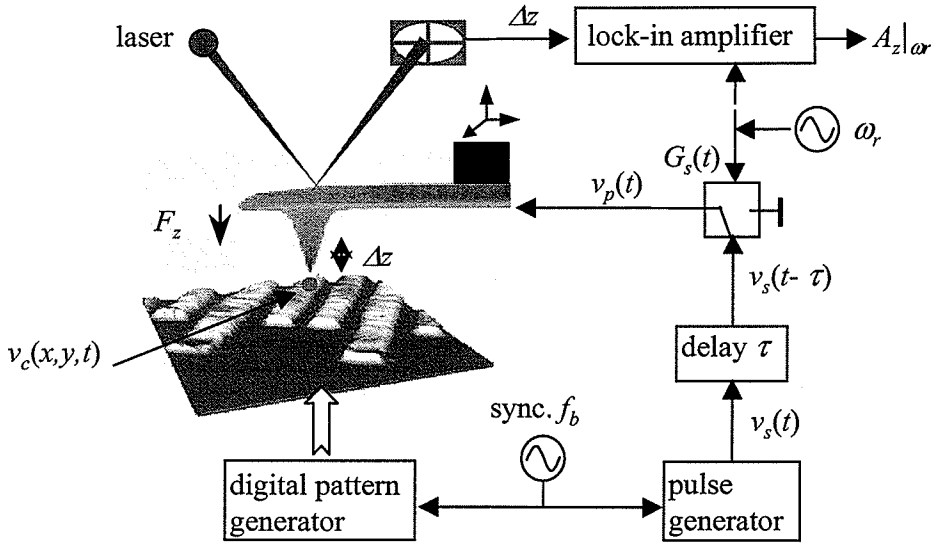

Fig. 2.17. A block diagram of the pulse amplitude modulation sampling setup in dynamic electrostatic force microscopy.

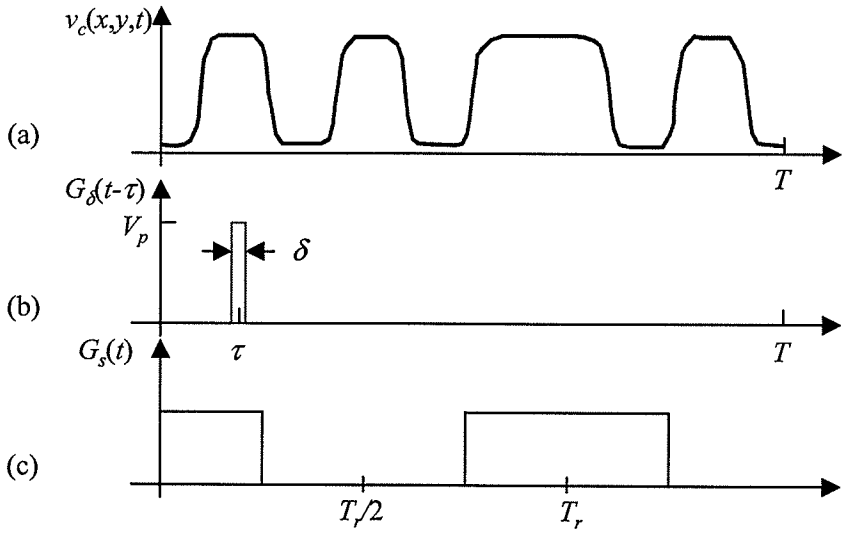

Fig. 2.18. An example of (a) a periodic circuit wave form $v_c(x,y,t)$ to be measured and (b) an ideal rectangular sampling pulse $G_\delta(t-\tau)$. (c) is the switch control signal $G_s(t)$. $T_r >> T$.

The pulse amplitude modulation can be easily realized by using a high speed switch, turned on and off at the frequency ω_p. ω_p is usually chosen at the mechanical resonance of the cantilever, ω_r, to enhance the cantilever deflection Q times, as illustrated by Eq. (2.16). The control signal of the switch $G_s(t)$ is shown in Fig. 2.18 (c).

The amplitude-modulated sampling pulse at the probe, $v_p(t)$, can be expressed mathematically as

$$v_p(t) = v_s(t)G_s(t), \tag{2.26}$$

where the control signal $G_s(t)$ is a square wave as

$$G_s(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1,odd}^{\infty}\frac{1}{n}\cos(n\omega_r t). \tag{2.27}$$

The electrostatic force component at the cantilever resonance frequency ω_r can then be expressed as

$$F_z(t)\big|_{\omega_r} = \frac{1}{\pi}\frac{\partial C}{\partial z}(\langle v_s(t), v_s(t)\rangle - 2\langle v_s(t), v_c(x,y,t)\rangle + 2\Delta\Phi(x,y)\langle v_s(t)\rangle)\cos(\omega_r t), \tag{2.28}$$

where

$$\langle a(t), b(t)\rangle = \frac{1}{T}\int_0^T a(t)b(t)dt, \tag{2.29}$$

$$\langle a(t)\rangle = \frac{1}{T}\int_0^T a(t)dt. \tag{2.30}$$

In Eq. (2.28), the first and third terms are independent of the desired circuit voltage $v_c(x,y,t)$. For extracting wave shapes of a digital signal, these terms contribute an offset force at ω_r. This force induced cantilever deflection, according to Eq. (2.16) is

$$\Delta z(t)\big|_{\omega_r} = [-\frac{2}{\pi}\frac{\partial C}{\partial z}\frac{Q}{k}\langle v_s(t), v_c(x,y,t)\rangle + D_z]\sin(\omega_r t), \tag{2.31}$$

where D_z is the constant deflection offset due to the offset force in Eq. (2.28).

If the sampling pulse is an ideal pulse, $v_s(t)=V_p G_\delta(t-\tau)$, ($V_p$ is the pulse amplitude) and assuming its bandwidth is much larger than the circuit signal bandwidth ($\delta \ll t_r$, t_f, which is the circuit signal rise and fall time respectively) with $\delta/T \ll 1$, the probe deflection monitored using a lock-in amplifier locked at ω_r can be approximated as

$$A_z(\tau)\big|_{\omega_r} = -\frac{2}{\pi}\frac{\partial C}{\partial z}\frac{Q}{k}\frac{\delta}{T}V_p v_c(x,y,t=\tau) + D_z. \tag{2.32}$$

The measured deflection is proportional to the circuit voltage at $t = \tau$, which is also illustrated in Fig. 2.19.

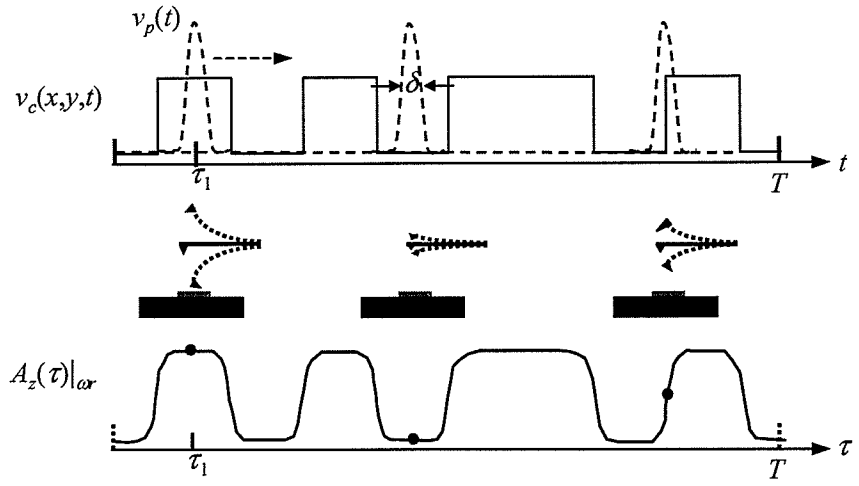

Fig. 2.19. The measured probe deflection is proportional to the circuit signal. The sampling pulse walks through the period T for the entire waveform (three pulse positions in corresponding to three different voltage levels are demonstrated).

The entire wave form $v_c(x,y,\tau)$ can be captured by shifting the delay τ. From the detected wave form, circuit signal information such as the timing, the rise and the fall times, and the ringing etc., scaled by $\partial C/\partial z$, Q and k, can be extracted.

To extract the circuit waveform with the proper polarity, it is important to realize the several factors that affect the polarity of the measurements, including the polarity of the sampling pulse and the phase of the lock-in amplifier reference signal. In practice, this problem can be solved by measuring a known circuit pattern.

The pulse amplitude could also be modulated with a sinusoidal signal. In this case the probe signal is simply

$$v_p(t) = v_s(t)\cos(\omega_r t), \tag{2.33}$$

and the deflection at the resonance frequency will be

$$A_z(\tau)\big|_{\omega_r} = -\frac{1}{2}\frac{\partial C}{\partial z}\frac{Q}{k}\frac{\delta}{T}V_p v_c(x,y,t=\tau) + DC_z. \tag{2.34}$$

The difference is that a constant ½ replaces $2/\pi$ in square wave modulation, as in the Eq. (2.32). Also, DC_z only comes from the circuit dc offset $\Delta\Phi$.

For a sinusoidal modulated sampling pulse, if the pulse is ac-coupled, i.e. $v_s(t)=V_p[G_\delta(t-\tau)-\langle G_\delta(t-\tau)\rangle]$ ($\langle G_\delta(t-\tau)\rangle$ is the dc component of the pulse), the cantilever deflection at ω_r will be

$$A_z(\tau)\big|_{\omega_r} = -\frac{1}{2}\frac{\partial C}{\partial z}\frac{Q}{k}\frac{\delta}{T}V_p[v_c(x,y,t=\tau)-v_c(x,y,t)]. \tag{2.35}$$

In this case the sampled waveform would also be ac-coupled. The dc offset $\Delta\Phi$ is thus not present in the measured wave form. An ac-coupled sampling pulse is useful for testing high-speed circuits which can be dc sensitive.

2.7.2.2 Pulse width modulation

A block diagram of the pulse width modulation approach used in EFM is depicted in Fig. 2.20 [54]. A step generator is used for generating fast rise time leading edges. The step generator is set by $v_s(t-\tau-\delta_1)$, where δ_1 is continuously changed at a rate close to the cantilever resonance frequency ω_r and reset by $v_s(t-\tau-\delta_0)$, where δ_0 is a constant delay. τ is an adjustable delay, which is used for pulse positioning relative to the circuit waveform as used for the pulse amplitude modulation approach.

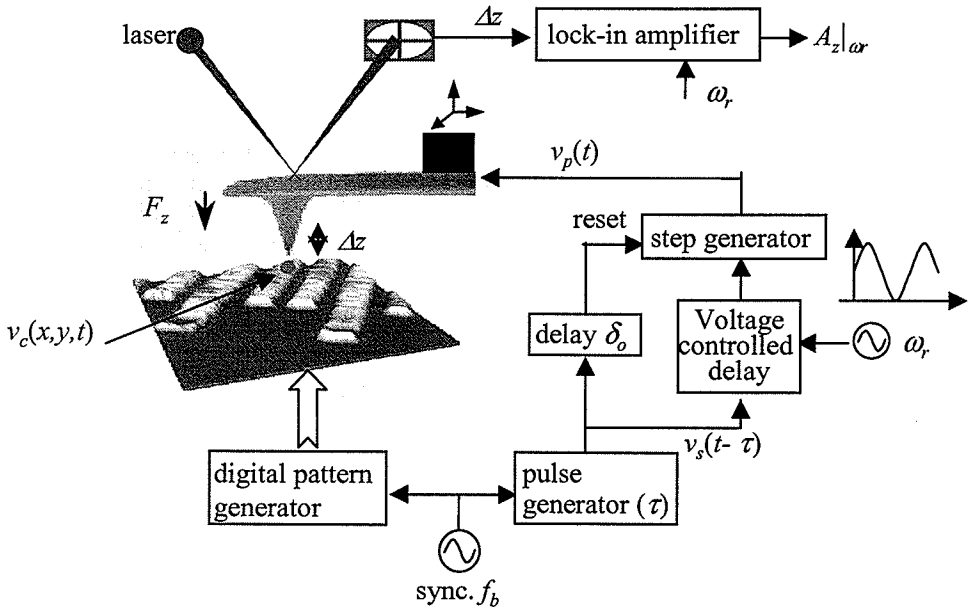

Fig. 2.20. A block diagram of the pulse width modulation sampling setup in dynamic electrostatic force microscopy.

In this case, the resulting pulse signal applied to the probe can be written as

$$v_p(t) = V_p G_{\delta(t)}(t-\tau), \tag{2.36}$$

and the pulse duration $\delta(t)$ can be expressed as

$$\delta(t) = \delta_o - \delta_1 = \delta_o - \frac{1}{2}\delta_m(1+\cos(\omega_r t)), \tag{2.37}$$

where δ_m is the modulation depth as seen in Fig. 2.21. Using this probe signal, the induced probe deflection at frequency ω_r (set at the cantilever resonance) is approximately [54]

$$\Delta z(t)\big|_{\omega_r} = [-\frac{1}{2}\frac{\partial C}{\partial z}\frac{Q}{k}\frac{\delta_m}{T}V_p v_c(x,y,t=\tau)+D_z]\sin(\omega_r t), \tag{2.38}$$

assuming $t_r << \delta_m << t_c$. t_r is the rise time of the probe signal, which is determined by the step generator and t_c is the rise time of the circuit signal. The pulse width modulation approach provides measurements with temporal resolution that relies mainly on the modulation depth δ_m as illustrated in Eq. (2.37).

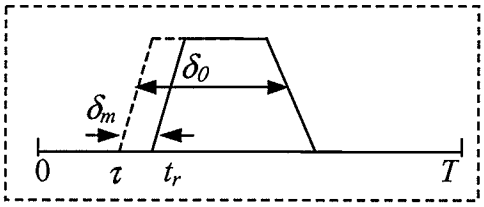

Fig. 2.21. The parameters associated with a pulse width modulation sampling pulse.

If the delay δ_1 is square-wave modulated at a rate of ω_r, the pulse duration $\delta(t)$ can be expressed as

$$\delta(t) = \delta_o - \delta_1 = \delta_o - \delta_m[\frac{1}{2}+\frac{2}{\pi}\sum_{n=1,odd}^{\infty}\frac{1}{n}\cos(n\omega_r t)]. \tag{2.39}$$

This results in a probe deflection at the cantilever resonant frequency similar to that expressed in Eq. (2.37), except for the constant $2/\pi$ in place of ½ [55].

The pulse width modulation approach requires a fast rise time leading edge (or falling edge) rather than a narrow pulse used in pulse amplitude modulation.

2.7.3 Summary

In comparison with the heterodyne equivalent-time sampling technique, the homodyne pulse sampling techniques result in a force component at a signal frequency (the cantilever resonance) that is proportional to the circuit voltage. This enables high resolution measurements by using a lock-in amplifier. The homodyne pulse sampling techniques should provide better measurement signal-noise-ratio since the low frequency end noise will not affect the measurements.

In comparison with the pulse amplitude modulation approach, the pulse width modulation approach only requires a high speed leading edge, which can be provided by a high-speed step generator. Most available signal generator circuits can be optimized to have a much faster rise time than the fall time, but it is harder for them to optimize both the rise and fall time. The pulse width modulation approach could be easier due to the hardware. Further discussion will be given in the next chapter.

Chapter 3

DEFM Experimental Setup and Measurements

Chapter 3 describes the experimental setup for performing DEFM pulse sampling measurements. Pulse sampling measurements of signals up to 10 Gbit/s on a matched transmission line and a Nortel wideband distributed amplifier (V-2) are presented.

3.1 Pulse generation

Using repetitive time sampling techniques, the bandwidth of the DEFM is no longer restricted by the frequency response of the cantilever but only by the sharpness and non-idealities in the sampling pulse. This is because the measured waveform is a convolution between the circuit signal and the sampling pulse as seen in Eq. (2.31).

One of the challenges of the pulse sampling EFMs is to generate high-speed electrical pulses to drive the probe. The idealities of the effective sampling pulses, and in turn the bandwidth of the measurement system, seriously depend on the performance of the pulse generation system.

Commercial high speed vector generators are capable of producing bit rates in the range of 40+ Gbit/s. These sources can generate pulses with widths of about a few tens of picoseconds. The vector generator used in this thesis was a multi-channel 1 GHz HP80000 data generator from Hewlett-Packward, which is capable of producing 1 ns width pulses with 10-90% rise and fall times of less than 120 ps [56]. This is satisfactory for sampling Mbit/s signals, but is not good enough for sampling Gbit/s signals. Vector generators that produce 40 Gbit/s pulses are commercially available today [52], but they are usually too expensive for academic researchers. In this work, the HP80000 data generator is used as the synchronizing source for the circuit excitations and the sampling pulses. Custom-built high-speed logic gates are used for ultra-short pulse generation.

3.1.1 High-speed logic gates

High-speed sampling pulse generation using lower speed pulses can simply be achieved by using a single high-speed AND/NAND logic circuit. Figure 3.1 shows a sample of the generation process. The two input pulses, $v_1(t)$ with the width δ_1 and $v_2(t)$ with the width δ_2, are synchronized with the same time period. The delay Δ between v_1 and v_2 is finely adjusted so that the time during which both inputs are high, is very short, i.e. the gate turn-on time is very short. Assuming the two input pulses have infinitesimal rise and fall time, the width of the generated ideal output pulses v_o is equal to the very short gate turn-on time δ ($=\delta_1-\Delta$). The pulse widths of the two inputs are trivial. The actual rise and fall time of the output pulses are governed by the transition speed of the gate control circuitry. As same as that of the two input pulses, the output sampling pulse has a period of T_p.

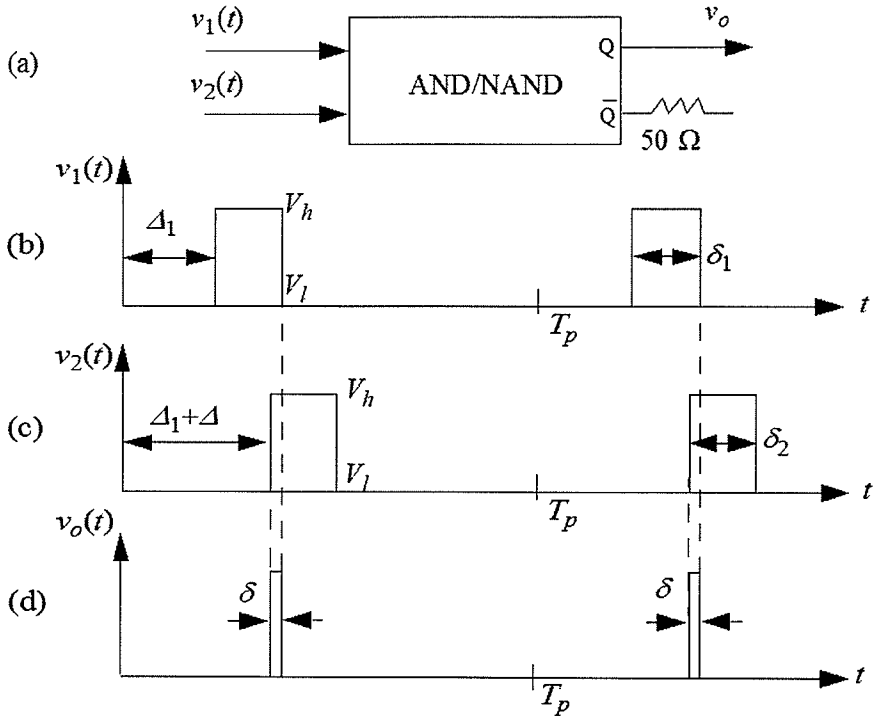

Fig. 3.1. A sample process of high-speed pulse generation. (a) An AND/NAND logic gate with two input pulses; (b) Input pulse v_1 with width δ_1; (c) Input pulse v_2 with width δ_2; and (d) Output pulse v_o with a width δ. δ equals $\delta_1-\Delta$, and Δ is the delay between the two pulses.

Figure 3.2 shows an example of the experimental setup for generating a high-speed sampling pulse using a single logic gate. The HP80000 data generator 1 ns pulse was

used as an input, v_i, for a broadband (dc – 18 GHz) resistive power divider (from Picosecond) to produce two identical pulses, $v_1(t)$ and $v_2(t)$. The delay between the two pulses was adjusted using two cables with a designed length difference. A wideband phase trimmer could be added in for fine adjustment. The logic gate is a commercial ultra-high speed 2 input AND/NAND (NLG4119), which operates up to 10 Gbit/s.

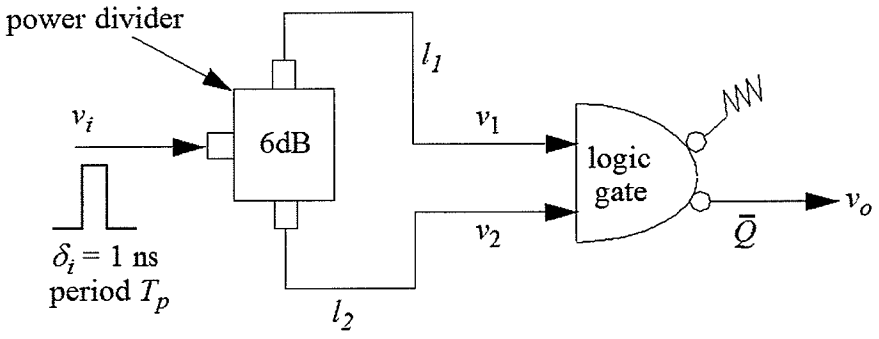

Fig. 3.2. A high-speed sampling pulse generation system using a single NAND/AND gate. The cable lengths l_1 and l_2 are used for a specified delay Δ between the two signal v_1 and v_2.

The NLG4119 was designed with Low-power Source Coupled FET technology and was fabricated with 0.15-μm GaAs MESFET process [57]. The input impedance matching was very good owing to built-in 50 Ω termination resistors between the input pins and the ground. The NLG4119 output rise and fall time are 30 ps and 25 ps (20 to 80%), respectively. Its logic levels are 0 for high and -0.9V for low. The reference level is set to -0.45 V, which determines the switching threshold of the gate. Figure 3.3 shows a photograph of the NLG4119 AND/NAND circuit and its functional diagram.

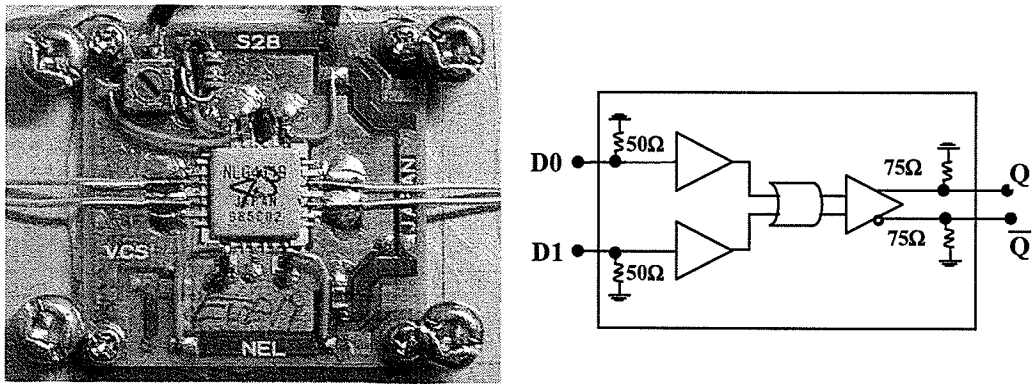

Fig. 3.3. A photograph of the NLG4119 AND/NAND circuit with 2 inputs and 2 outputs and its functional diagram.

The HP80000 1ns pulses have typical 90 ps rise time and 120 ps fall time as seen in Fig. 3.4. Using the setup in Fig. 3.2, a serious problem was found for generating an output pulse < 50 ps. The amplitude of the output pulses had up to 30% fluctuation during the phase shifting of the input signal $v_i(t)$. It was found that the overlapping range with both signals above the gate threshold level was very small for creating an output pulse < 50 ps and the gate was actually operating in its linear region. The output was very sensitive to any small change to the overlapping range. The change might come from the change of the input pulse width or the rise/fall time as seen in the Fig. 3.1.

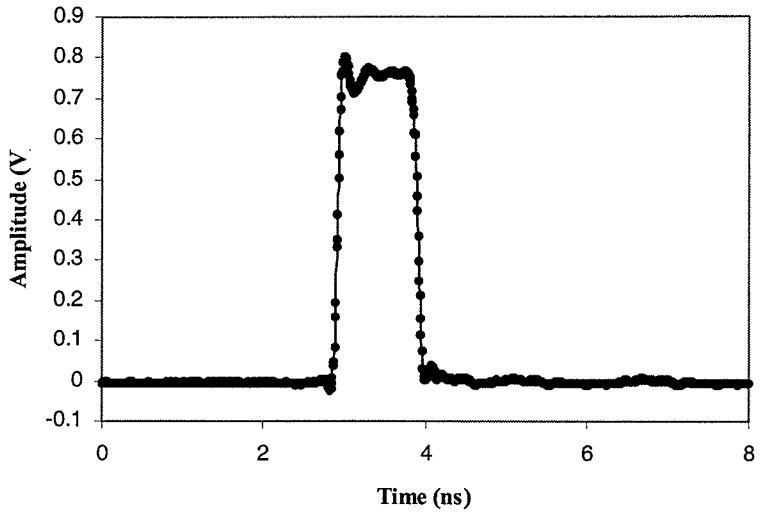

Fig. 3.4. A sample of 1ns pulse from the HP80000 data generator. The 10-90% rise time is ~ 90 ps.

It was found that the HP80000 1ns pulse had a fixed rise/fall time but its width changed during phase shifting. The pulse width to its delay (from 0 to 1000 ps) measured by a 20 GHz Tektronix sampling oscilloscope is shown in Fig. 3.5. The 1ns pulse width changes from 977 ps to 1002 ps. The change is significant to the overlapping range for generating a high-speed pulse with a pulse width < 50 ps. This problem is possibly caused by the asymmetric loading effect of the HP80000 data generator's phase shift circuitry to the rising edge and the falling edge.

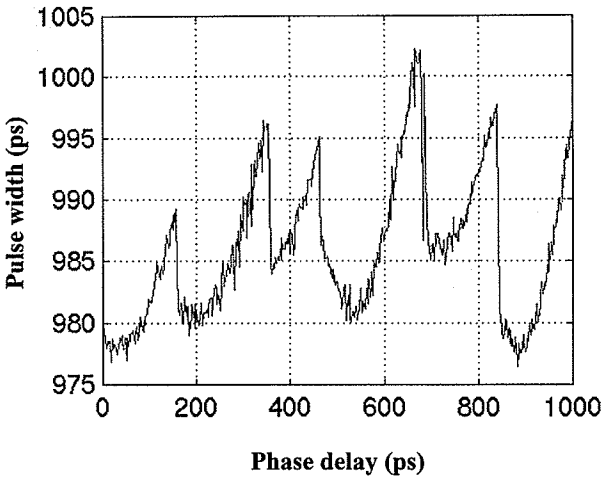

Fig. 3.5. The HP80000 1 ns pulse width changes with the delay.

The solution to this problem is to use only one edge of the HP 1ns pulse to generate high-speed sampling pulses. The two logic gate implementation is used for the purpose. We used the two differential outputs (Q_1, $\overline{Q_1}$) of an 8 Gbit/s logic gate (NL4519 [57]), photographed in Fig. 3.6, as the two input pulses $v_1(t)$ and $v_2(t)$ for the logic gate NLG4119 since the NL4519-2 has the same logic levels as NLG4119. The diagram of the experimental setup is shown in Fig. 3.7. The pulse generation process is described in Fig. 3.8.

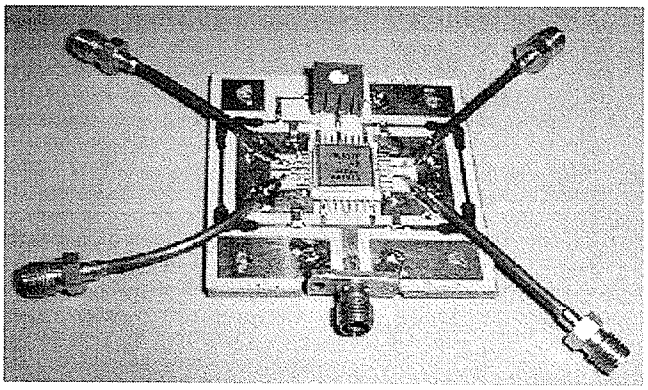

Fig. 3.6. A photograph of the logic gate NL4519-2.

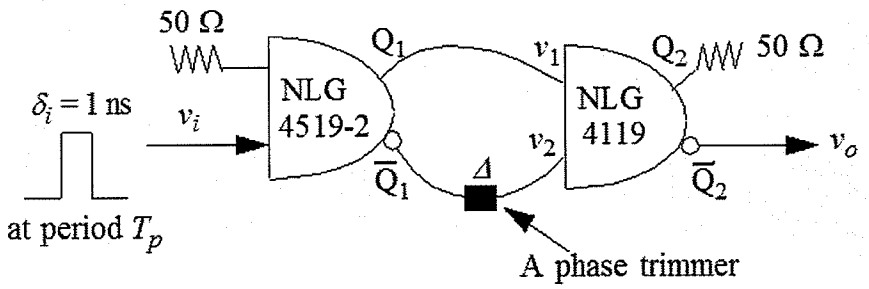

Fig. 3.7. The two differential output signals of the NL4519-2 were used as the two inputs for the NLG4119 for high-speed pulse generation.

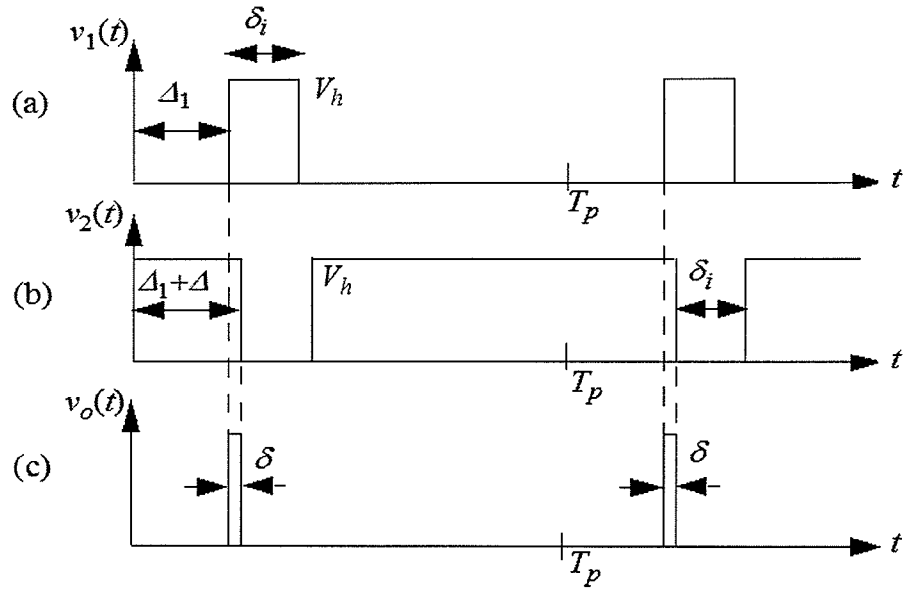

Fig. 3.8. The pulse generation process with two differential input signals. Only the rising edge of the HP80000 1 ns pulse was used. (a) Input pulse v_1; (b) Input pulse v_2 with a delay Δ; and (c) Output pulse v_o with a width δ.

Since the rising edge of the signal $v_1(t)$, and the falling edge of the signal $v_2(t)$ came from the same rising edge of the HP 1ns pulse v_i, the overlapping region as an example shown in Fig. 3.9, would not change with the phase shifting. The port Q_2 of NLG4119 was used for narrow pulse output while port $\overline{Q}_2$ was matched to avoid possible reflections. With this setup, the measured pulse amplitude fluctuation during the phase shifting was much better (less than 5%).

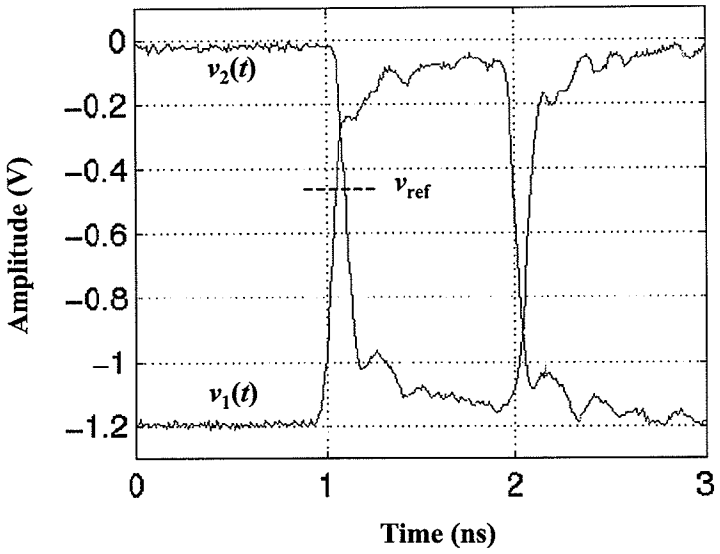

Fig. 3.9. A sample of a fixed overlapping range for the logic gate NLG4119 to be turned on by using two differential input signals.

A Tektronix TDS 8000 Digital Sampling Oscilloscope with a 20 GHz electrical sampling module was used to capture the high-speed pulses. A sample pulse generated by the NGL4119 is shown in Fig. 3.10. The pulse had an amplitude of 0.7 V and a 47 ps pulse width (full width at half maximum or FWHM).

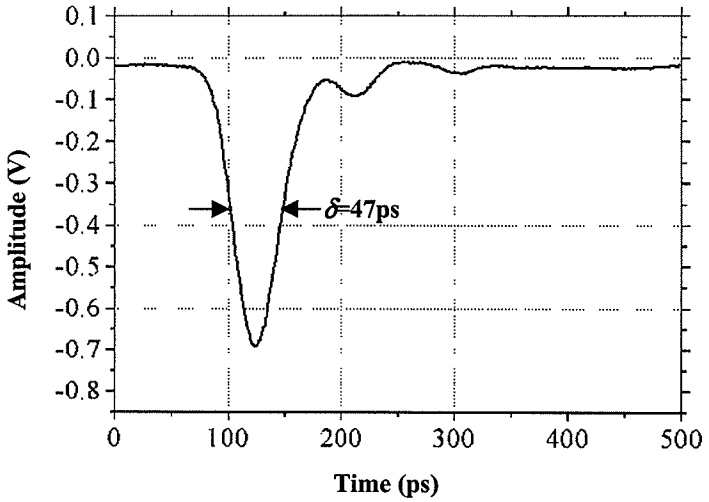

Fig. 3.10. A sample of 47 ps sampling pulse generated by the logic gate NLG4119 using two differential input signals.

To further improve the system bandwidth, a NEL 20 Gb/s 2 input AND/NAND module (MOF19A) was used to replace the NLG4119. The module has a similar function to the NLG4119 but has a faster output rise/fall time (20-80%), typically 20 ps [57]. A 37

ps pulse was generated and measured by the 20 GHz Tektronix oscilloscope, as shown in Fig. 3.11. It has an amplitude of about 0.8 V.

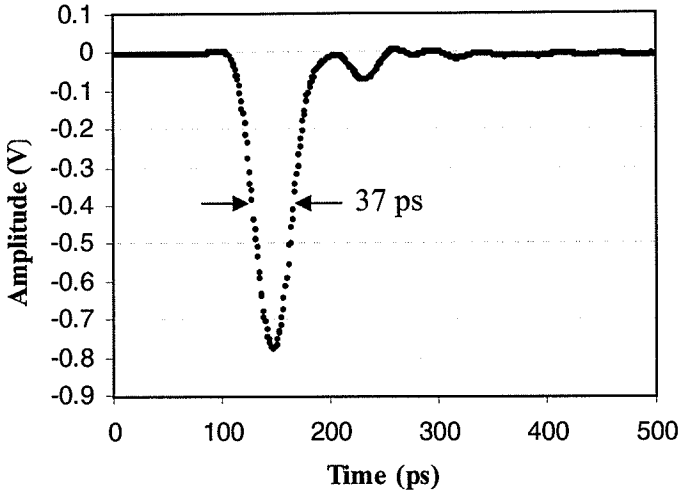

Fig. 3.11. A sample of 37 ps pulse generated by the logic gate MOF19A.

Pulse width can be adjusted by using the phase trimmer to shorten the logic gate's tune-on duration. Started at one point (at which the 37 ps pulse was generated), the logic gate's output amplitude dropped rapidly. Further decreasing the pulse width would significantly sacrifice the pulse amplitude and the SNR of the pulse sampling measurement.

3.1.2 High-speed voltage step generators

High-speed voltage step generators can be used for the EFM pulse width modulation approach as described in last chapter. They can also be used for generating virtual sampling pulses. In this section, the process for generating virtual pulses by using voltage steps will be described. The advantage of this process also is that it only requires a fast leading or falling edge rather than a narrow pulse.

3.1.2.1 High-speed voltage step generators

High-speed step generators include a non-linear transmission line (NLTL) [58], a step recovery diode [59] and an edge compressor [60]. The NLTL is an ordinary transmission line with the addition of a non-linear element. It has been used for generating sub-picosecond voltage steps [58]. A step recovery diode (SRD) is also a non-linear component. Edge compressors are usually based on the NLTL technology.

A nonlinear transmission line is an ordinary transmission line with the addition of a non-linear element to purposely distort the input signal of the line. By adding varactor diodes periodically in shunt with a transmission line, a nonlinear transmission line is realized. A single section of an NLTL is shown in Fig. 3.12.

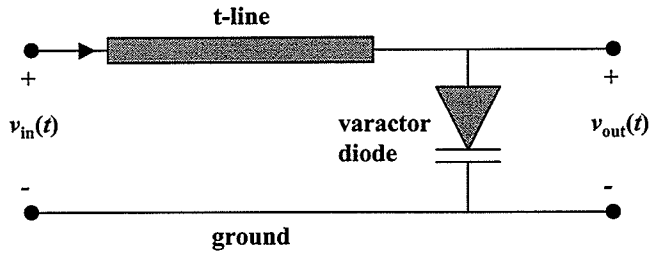

Fig. 3.12. A single section of a nonlinear transmission line.

Since the capacitance of a varactor diode rises and drops in concert with its input voltages, these diodes in a nonlinear transmission line cause the capacitance of the original transmission line to be a function of the input voltage and the propagation speed to vary with the input voltage. The higher instantaneous voltages travel faster than the lower instantaneous voltages down the nonlinear transmission line. For example, a sinusoidal input signal can distorted and compressed to output a high-speed voltage step.

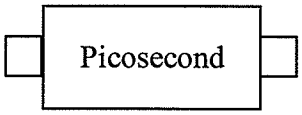

Fig. 3.13. A Picosecond state-of-the-art edge compressor [60].

Based on nonlinear transmission line technique, a state-of-the-art edge compressor as shown in Fig. 3.13 can drive 50 Ω with a transition time as short as 4 ps, compressing the input pulse edges by as much as 80 ps and providing peak output amplitudes greater than 10 V with a maximum repetition rate larger than 20 GHz [60]. Due to its small size, it could be installed close to the probe tip to minimize the cable loss and to improve the DEFM measurement bandwidth.

3.1.2.2 Virtual pulse generation

The DEFM and its modulation approach can be employed in a special way to generate virtual sampling pulses [45] by using high speed voltage step generators. A block diagram of the virtual sampling pulse generation approach is shown in Fig. 3.14.

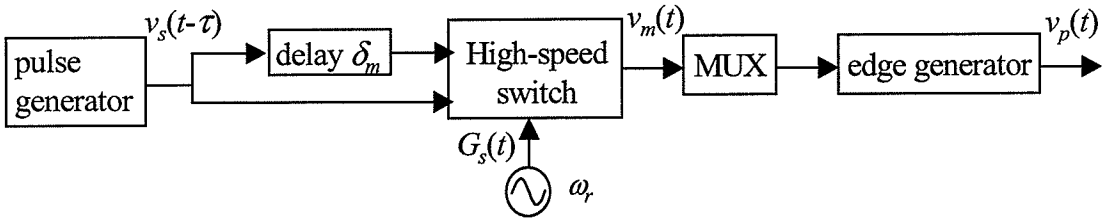

Fig. 3.14. A block diagram of the virtual pulse generation approach.

Two identical voltage pulses, $v_s(t-\tau)$ and $v_s(t-\tau-\delta_m)$ (δ_m is a small fixed delay), are toggled by a switch at a rate close to the cantilever resonant frequency ω_r. Due to the square wave modulation of the two pulses, the resulting switch output signal has a frequency component of ω_r as

$$v_m(t)\big|_{\omega_r} = \frac{2}{\pi}[v_s(t-\tau) - v_s(t-\tau-\delta_m)]\cos(\omega_r t). \tag{3.1}$$

Figure 3.15 shows that $v_s(t-\tau)$-$v_s(t-\tau-\delta_m)$ results in two virtual pulses, one positive and one negative. Assuming the edge generator generates fast falling edges, the rising edges of the input pulses can be gated out by using a MUX to avoid the positive pulse. Then the probe signal at the frequency ω_r is equivalent to a virtual narrow pulse as

$$v_p(t)\big|_{\omega_r} = v_m(t)\big|_{\omega_r} \cdot G_u(t)G_e(t), \tag{3.2}$$

where the signal $G_u(t)$ represents the MUX's operation and $G_e(t)$ represents the edge generator's operation (to sharpen the falling edges).

The virtual pulse approach is actually a pulse width modulation approach (as discussed in Chapter 2). The probe deflection at the cantilever resonant frequency is similar to that expressed in Eq. (2.37), except for the constant $2/\pi$ in place of ½.

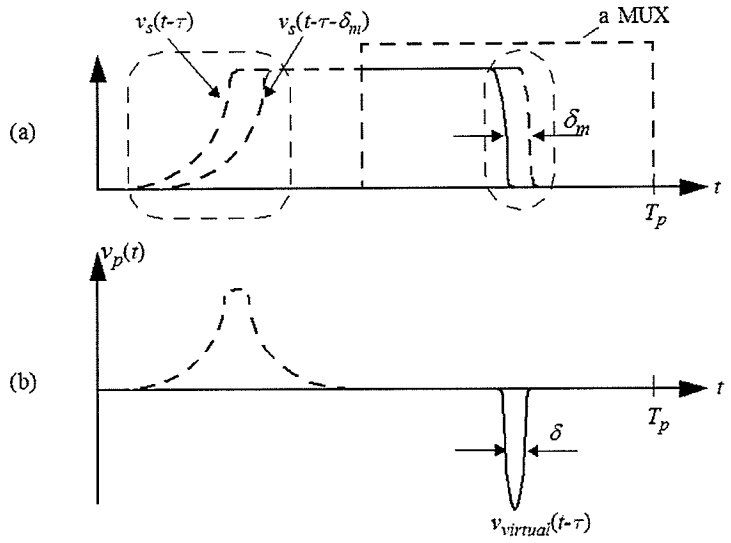

Fig. 3.15. A virtual sampling pulse generation process. The rising edges of the input pulses are gated out by using a MUX and the falling edges are sharpened and used for generating a virtual sampling pulse.

The virtual pulse width δ_p depends on the step generator's transition time t_r and the modulation depth δ_m. Two identical Gaussian steps (0 to 1V) were used to estimate the width and amplitude of a virtual pulse for the different modulation depth δ_m, assuming the step's rise time t_r = 32 ps. Table 1 shows that $\delta_p \sim \delta_m$ if $\delta_m << t_r$. However, when $\delta_m << t_r$, $\delta_p \sim \delta_r$ and the virtual pulse amplitude is much smaller than the maximum when $\delta_m << t_r$ (i.e. the measurement SNR will be sacrificed). This means that the temporal resolution of the virtual pulse approach is limited by the step generator's transition time. This also applies to the other pulse width modulation approaches.

Table 1. The width and the amplitude of a virtual pulse in corresponding to the delay between the two 1 V Gaussian steps with a 32 ps transition time.

Delay δ_m (ps)	Width δ_p (ps)	Amplitude (V)
50	51.3	0.958
40	43.2	0.896
30	37.1	0.777
20	32.2	0.584
10	29.6	0.315

Experimentally, the virtual pulse approach was implemented by using the HP80000 data generator's 1ns pulses as the input signals, $v_s(t-\tau)$ and $v_s(t-\tau-\delta_m)$, and by using the logic gate NL4519-2 as an edge generator. EFM pulse sampling measurements using the virtual pulse approach with delay δ_m = 50 ps were performed on a transmission line to prove the concept [45].

3.2 Pulse propagation

3.2.1 The reflection problem

The reflection problem exists in the pulse propagation to the tip of the cantilever. As seen in next chapter, the coupling capacitance between the probe and the circuit is usually very small, which means that the probe can be simply modeled as an open circuit. The reflection from the tip may damage the pulse generator or generate multiple reflections by any possible mismatch on the path from the generator to the tip. A wideband amplifier with series attenuators at its output as seen in Fig. 3.16 (the pulse output path) could be used as a pulse output matching network to solve the problem. The amplifier improves the pulse amplitude and the measurement signal-to-noise-ratio. The attenuators adsorb the reflection from the cantilever to protect the gate and also to avoid multiple reflections.

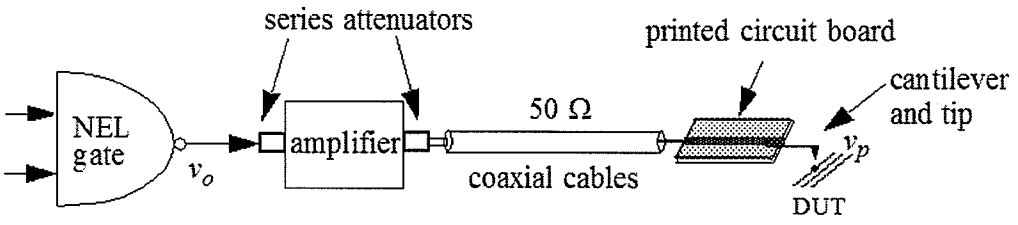

Fig. 3.16. The block diagram of the pulse output path.

One of the wideband amplifiers used in this work is shown in Fig. 3.17. It has a nominal 26 dB gain and a bandwidth from 30 kHz up to 15 GHz. When it was used to amplify the NLG4119 47 ps pulse, the resulting amplifier's output was measured and shown in Fig. 3.18. The output pulse has a FWHM of about 57 ps. This means that the amplifier's pulse response is not very good. The amplifier also adds some ringing to its output pulse which will be discussed later.

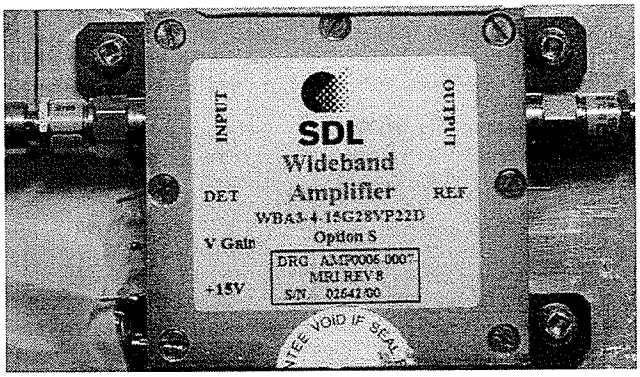

Fig. 3.17. A photograph of the wideband amplifier with series attenuators on its input/output port to protect the pulse generator.

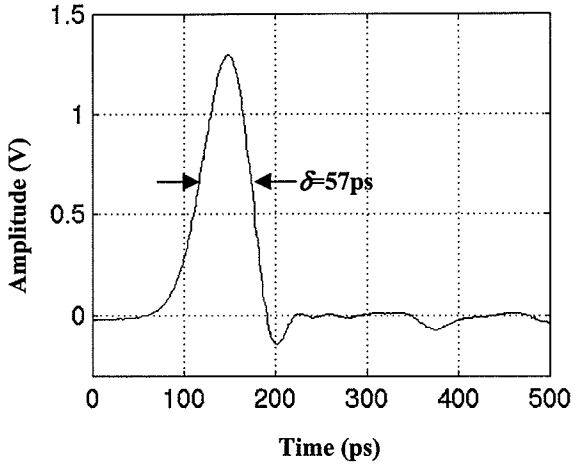

Fig. 3.18. The amplifier's output pulse with a FWHM 57 ps.

Since the logic gate MOF19A's output port is well matched with a reflection coefficient S_{22} < -10 dB, the gate output pulse can be directly applied to the probe after carefully adjusting the cable length to avoid the reflected pulse overlapping with the output pulse at the port. Using this setup, the reflected pulse and multiple reflections were captured by a 20 GHz digital sampling oscilloscope through a broadband (0~18 GHz) power splitter. The results are shown in Fig. 3.19 for a period of 8 ns. As seen in Fig. 3.19, the multiple reflections were quite small, which means the reflected pulse from the cantilever was mostly absorbed at the gate output port. This setup was used for 10 Gbit/s circuit measurements.

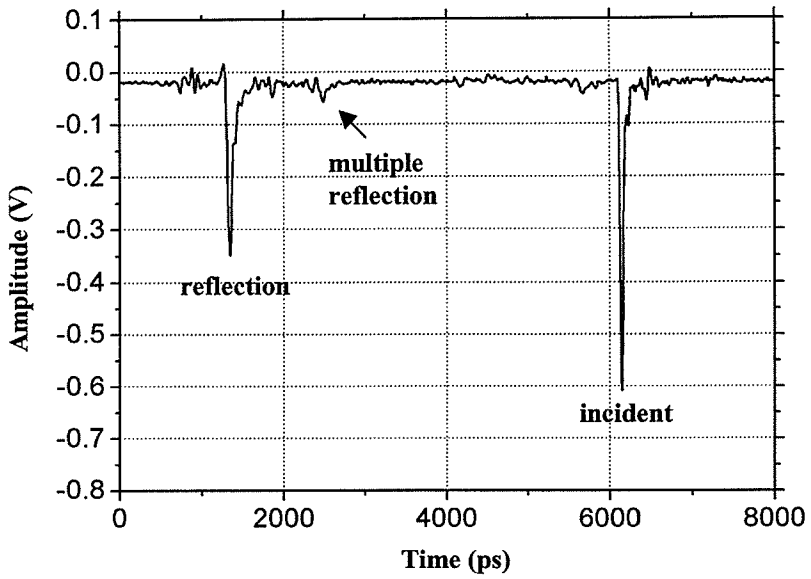

Fig. 3.19. The logic gate MOF19A output pulse and its reflection and multiple reflections.

3.2.2 The loss and pulse dispersion

As seen in Fig. 3.19, the reflection from the cantilever is about half the amplitude of the incidence. The calculated input pulse energy is about three times as much as the reflected pulse energy. This means that the loss along the pulse propagation path was quite large with the cable loss being the main contributor. The cable connecting the pulse generator to the printed circuit board has to be flexible for adjusting and isolating the mechanical noise as well. However, a flexible cable usually has a bigger loss due to its thin dielectric layer. To suppress this problem, the pulse generator should be placed as close as possible to the cantilever.

If the frequency response of the pulse output circuitry (including the probe/circuit coupling etc.) is defined as $H_g(\omega)$, then the actual pulse at the tip $v_p(t)$ is $F^1[H_g(\omega)V_o(\omega)]$ rather then the gate output pulse $v_o(t)$, where $V_o(\omega)$ is $F[v_o(t)]$. F is the Fourier transform operation.

The frequency response $H_g(\omega)$ can be analyzed using a static equivalent circuit as shown in Fig. 3.20 [61]. In the model, the lumped elements C_p, R_p and L_p represent the probe shunt capacitance, resistance and inductance, respectively. C_{pc} is the coupling capacitance between the probe and the circuit traces, which depends on the probe/circuit geometry. These elements, and that of the cable and the transmission line which supports

the cantilever, contribute to the $H_g(\omega)$. It is quite difficult to mathematically solve for $H_g(\omega)$.

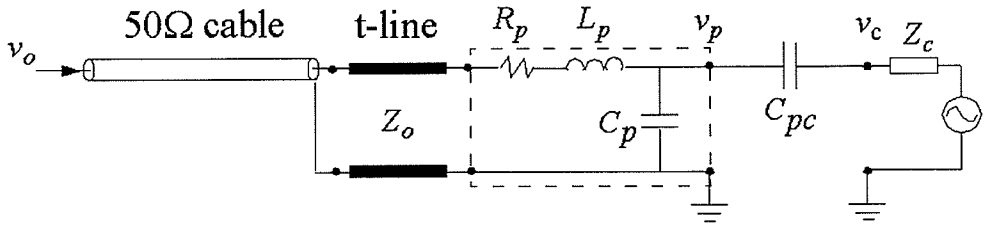

Fig. 3.20. An equivalent circuit coupling model for the pulse output circuitry and the probe/circuit transmission line.

Experimentally characterizing the performance of $H_g(\omega)$ could be done by using a time domain reflectometer (TDR) with a high-speed pulse/step to measure the reflected pulse/step from the output circuitry. The incident pulse/step should have much higher bandwidth than the frequency response of the circuitry. The frequency response $H_g(\omega)$ can then be estimated from the reflected signal by using a deconvolution process.

In this work, the TDR measurement was performed using the TDS 8000 Tektronix 20 GHz digital sampling oscilloscope. Its output is a step transient with a typical rise time of 28 ps and an amplitude of 0.25 V [62] as seen in Fig. 3.21 (the $v_i(t)$). This voltage step was used as the incidence for the pulse output circuitry and the measured reflected signal from the probe was also shown in Fig. 3.21 (the $v_r(t-\tau)$, which is phase-shifted for the comparison). The $v_r(t-\tau)$ was significantly dispersed. The dispersion was caused by the cable loss, the transmission line and the parasitics at the probe end.

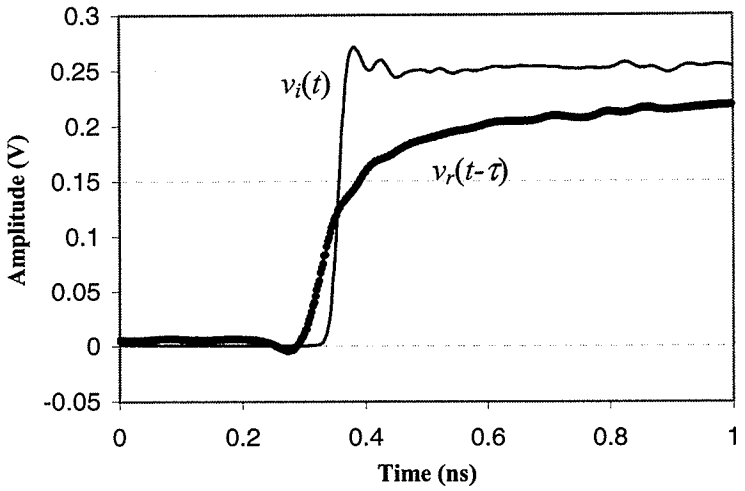

Fig. 3.21. The TDR measurement of the pulse output circuitry.

To simulate the pulse response of the pulse output circuitry, a 40 ps delay was added to the incidence. The difference between the two incidences, $[v_i(t)-v_i(t+40)]$, is shown in Fig. 3.22 (the input pulse) and has a FWHM pulse width of 50 ps. The difference between the two reflected signals, $[v_r(t)-v_r(t+40)]$, is also shown in Fig. 3.22 (the reflected pulse). The signal energy loss along the pulse output circuitry was calculated to be about 40%. This result is close to the real pulse multiple reflection measurement as shown in Fig. 3.19.

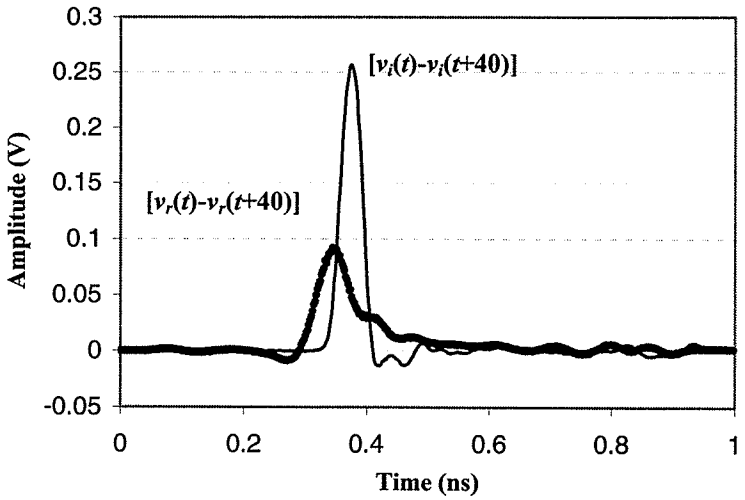

Fig. 3.22. The pulse response of the pulse output circuitry simulated from the TDR measurements with a 40 ps delay added to the incidence.

3.2.3 A coplanar waveguide transmission-line cantilever

The parasitics of the cantilever and the cantilever supporting chip limit the bandwidth of the sampling pulse that can propagate to the tip. To achieve greater measurement bandwidth, a custom-designed cantilever, as shown in Fig. 3.23(a) [63], could be used in the future. The cantilever uses a coplanar waveguide (CPW) transmission line with matched impedance all the way to the tip and is designed for a measurement bandwidth up to 40 GHz [63]. When stressed, the CPW cantilever provides access for EFM probing of deep trenches or high step heights (up to 100 μm) as shown in Fig. 3.24 (C_t is the coupling capacitance between the tip and the circuit) [64]. The fabrication of the CPW cantilever requires usage of the MEMS technique and follows a standard microfabrication procedure [64]. Figure 3.23(b) is one example of the fabricated cantilever before releasing. The investigation on the CPW cantilever is still undergoing.

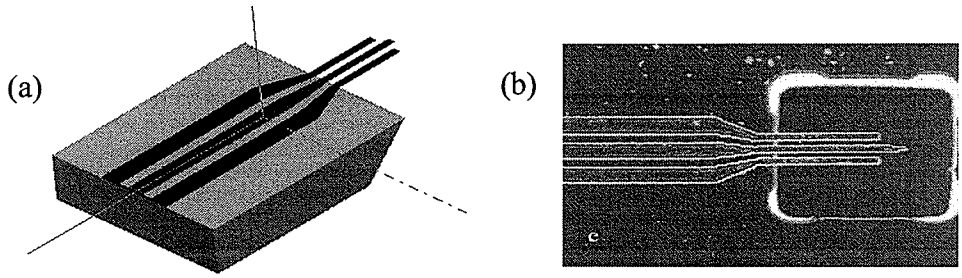

Fig. 3.23. A custom-designed CPW cantilever [63].

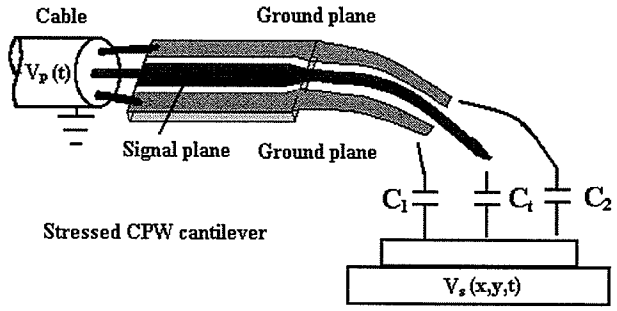

Fig. 3.24. EFM probing with a stressed CPW cantilever [64].

3.3 Measurement system in general

An EFM high-speed pulse sampling system was constructed and is shown in the block diagram in Fig. 3.25. The HP80000 data generator system (DGS) has multiple channels for down to a 1 ns pulse output. One of them was used as the input for the high-speed pulse generation system as described in previous sections. Depending on the applications, another channel may be used as a data pattern generator for the circuit stimulus at, or below 1 Gbit/s. PC LabVIEW [65] programs control the HP outputs (including their delays, their magnitudes and their logic pattern through a General Purpose Interface Bus (GPIB)) and automate the pulse sampling process by shifting the delay of the pulse and collecting sampled data from the lock-in amplifier (LIA) through the PC plug-in data acquisition (DAQ) board.

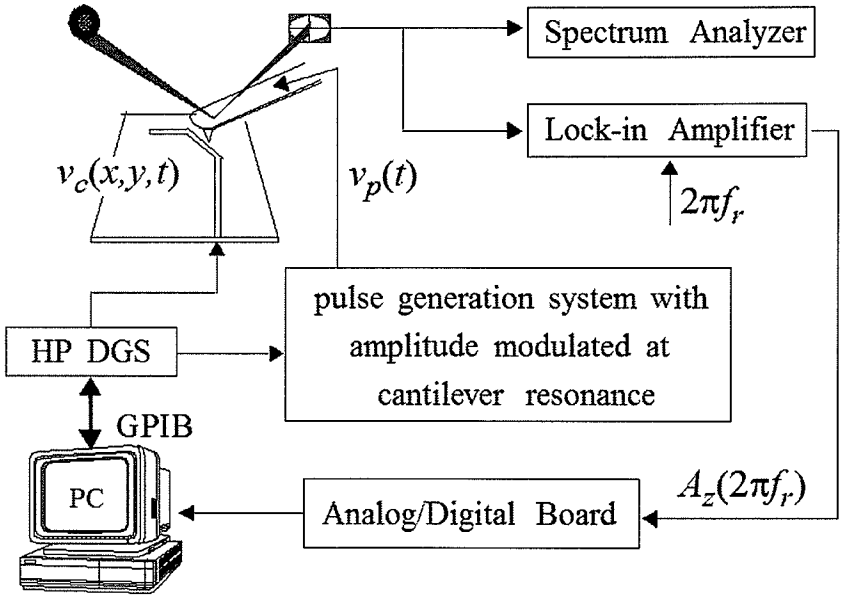

Fig. 3.25. A block diagram of DEFM measurement system.

The measured deflections of the cantilever were monitored by a spectrum analyzer and collected by a lock-in amplifier (from Stanford Research Systems). The lock-in amplifier is used to measure the amplitude and phase of the desired cantilever deflection signal, which is buried in broadband noise. It acts as a narrow bandpass filter to only allow the desired signal, which is set by the chosen reference signal, to pass through.

The DAQ board quantifies the lock-in amplifier analog output with a sample rate which is usually set to 100 samples per second. Averaging time is selected to be one second to match the LIA intermediate frequency (IF) bandwidth.

The LabVIEW programs display the measurement result on the computer screen and store the data in a text file. The advantage of LabVIEW programming over the classical text/script based programming is its graphical interface where the user builds a program by connecting (wiring) built-in component icons, i.e. by drawing the program's algorithm.

3.4 The tip-to-sample distance control

The coupling capacitance between the tip and the circuit depends on the tip-to-sample distance. To control the tip-to-sample distance during the EFM measurement, a feedback system, based on controlling the probe capacitance gradient, was investigated. First, the probe deflection signal transfer function was investigated.

3.4.1 The probe deflection signal transfer function

As discussed in chapter 2, the laser beam shift Δs is proportional to the cantilever deflection Δz and the deviation in optical signals absorbed between the two photo diodes is proportional to the shift Δs.

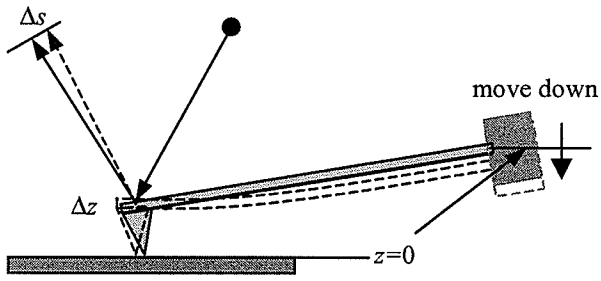

Fig. 3.26. The experimental investigation of the probe deflection signal transfer function.

The experimental investigation of the probe deflection signal transfer function was performed by measuring the cantilever deflection signal from the photo detector as a function of the tip-to-sample distance z (moving the probe down to a sample surface), as shown in Fig. 3.26. The result is shown in Fig. 3.27. The cantilever deflection signal keeps its original voltage as the probe moves closer to the circuit until the tip snaps on to the circuit surface ($z = 0$). Further pushing the cantilever, the cantilever deflection signal

increases rapidly and is saturated afterwards (the laser spot locates only on one photo diode). The negative tip-to-sample distance means that the cantilever is bending upwards at its end (a free cantilever is tilted in 15°). The gain (or transfer function) of the deflection detecting system could be estimated from the slope, $|k_d|$, to be about 1.0 V/μm in this example, i.e. a 1 mV deflection signal approximately equals 1 nm cantilever deflection. In this measurement, the cantilever and the circuit were both grounded to avoid the electrostatic force interaction.

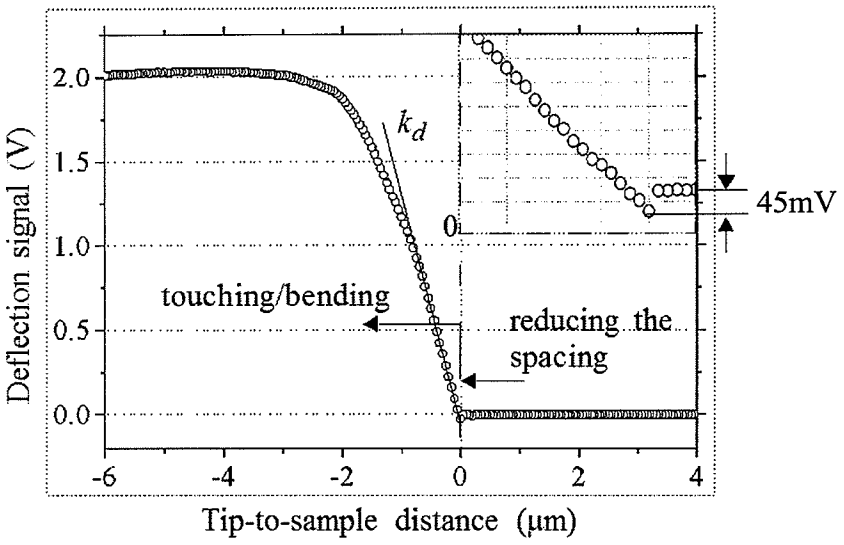

Fig. 3.27. The cantilever reflection signal in a function of the tip-to-sample distance. The zoom-in measurement was done with a 20 nm step.

The minimum tip-to-sample distance can be estimated from the cantilever's snap-down distance. The detected difference signal was about –45 mV for this case when the cantilever was pulled down to the circuit surface, as shown in Fig. 3.27. The cantilever's snap-down distance was estimated to be 45 nm for this specific case. The snap-down distance is mainly determined by the environmental noise and the surface cleanliness. The dominant noise source is thermal noise and the typical RMS noise was measured to be about 10 mV for the probe station which corresponded to a 10 nm cantilever vibration. The measured peak-to-peak noise was about 40 mV, i.e. 40 nm cantilever vibration.

The typical tip-to-sample distance for the EFM measurements in this work is between 100 nm to 2 μm (i.e. in noncontact mode).

3.4.2 The probe height control

A probe position feedback system to maintain the tip-to-sample spacing at small tip-to-sample distances is proposed and investigated. The feedback system involves adding a sinusoidal signal of the frequency f_z to the probe sampling pulse v_p, where f_z is slightly lower than half of the cantilever resonant frequency, f_r, i.e. $2f_z \sim f_r$. The probe driving signal is now

$$v_p^*(t) = v_p(t) + V_z \cos(2\pi f_z t). \tag{3.3}$$

The square-law of the electrostatic force equation (Eq. (2.1)) produces an additional force component at $2f_z$ (comes from $[V_z \cos(2\pi f_z t)]^2$) and the force induced cantilever deflection component is

$$\Delta z(t)\big|_{2f_z} = -\frac{1}{4}\left(\frac{\partial C}{\partial z}\frac{Q}{k}V_z^2\right)\sin(4\pi f_z t). \tag{3.4}$$

The magnitude of this signal can also be monitored by a lock-in amplifier to be

$$A_z\big|_{2f_z} = -\frac{1}{4}\left(\frac{\partial C}{\partial z}\frac{Q}{k}V_z^2\right). \tag{3.5}$$

From this equation, the $2f_z$ signal depends only on $\partial C/\partial z$, the cantilever mechanical properties and the bias voltage V_z. It is independent of the circuit signal $v_c(x,y,t)$ and $v_p(t)$, however. Therefore, it could be used as a control signal to maintain the tip-to-sample distance during the repetitive pulse sampling process.

Experimentally, the sinusoidal wave may be added on to the sampling pulse by using a wideband bias-T as shown in Fig. 3.28(a) [66]. The deflection signal at f_r and $2f_z$ can be monitored by using a spectrum analyzer. Here, the sampling pulse (the HP80000 1 ns pulse train at 1 V) was modulated at 12.625 kHz, the sinusoidal wave signal was at 6 kHz (i.e. $2f_z$=12 kHz), and the circuit (a microstrip line) was energized at 1 V. The f_r and $2f_z$ signals were both located close to the cantilever resonance peak as shown in Fig. 3.28(b) to enhance the deflection by approximately a factor of Q. The frequency difference of the two desired signals was 625 Hz.

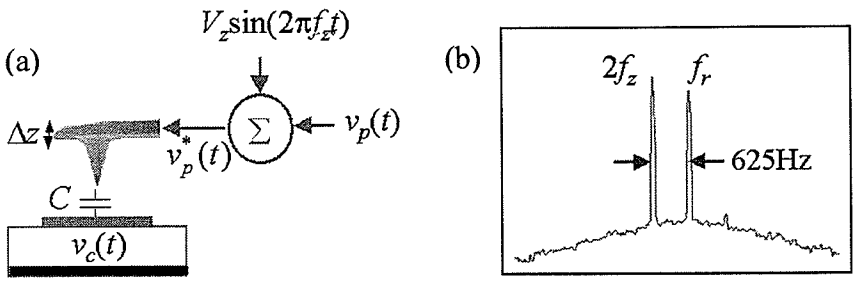

Fig. 3.28. The control signal for maintaining a constant tip-to-sample distance.

For this case, the deflection signals at $2f_z$ and f_r as a function of the tip-to-sample spacing z were measured by using two lock-in amplifiers. In the measurement, the probe tip moved from 4 μm away to the circuit surface in a step size of 50 nm. As seen in Fig. 3.29, the deflection signals at $2f_z$ and f_r increase when the tip-to-sample spacing decreases until the tip touches the surface. The f_r signal increases slightly quicker than the $2f_z$ signal (the ratio of $A_z|_{f_r}/A_z|_{2f_z}$ increases from 0.75 to 1.1 when the tip-to-sample distance changes from 4 μm to 100 nm). This is possibly because the deflection signal $A_z|_{f_r}$ depends on the circuit voltage (and its distribution) but the signal $A_z|_{2f_z}$ does not as seen in Eq. (3.5).

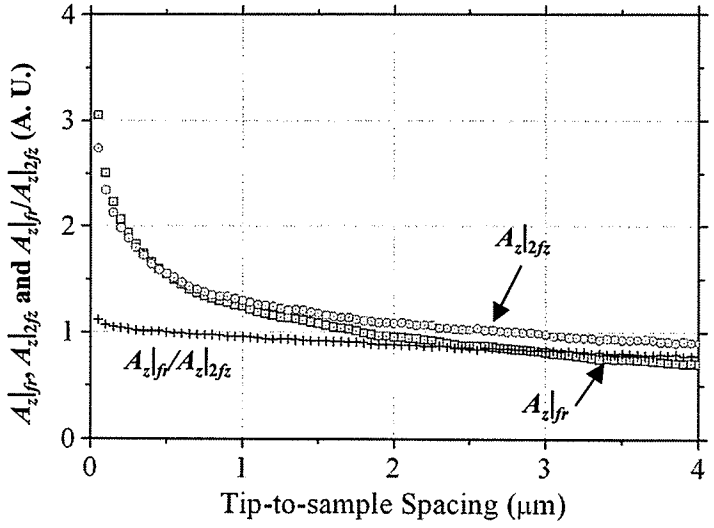

Fig. 3.29. An example of the measured signals: $A_z|_{f_r}$, $A_z|_{2f_z}$, and the ratio $A_z|_{f_r}/A_z|_{2f_z}$ as a function of the tip-to-sample spacing.

Figure 3.30 shows the $\log(A_z|_{f_r})$ and the $\log(A_z|_{2f_z})$ as a function of the $\log(z/z_0)$ (z_0=1 μm). From these curves, $dC/dz \sim 1/z^n$, where $n \sim 0.3$ for this case.

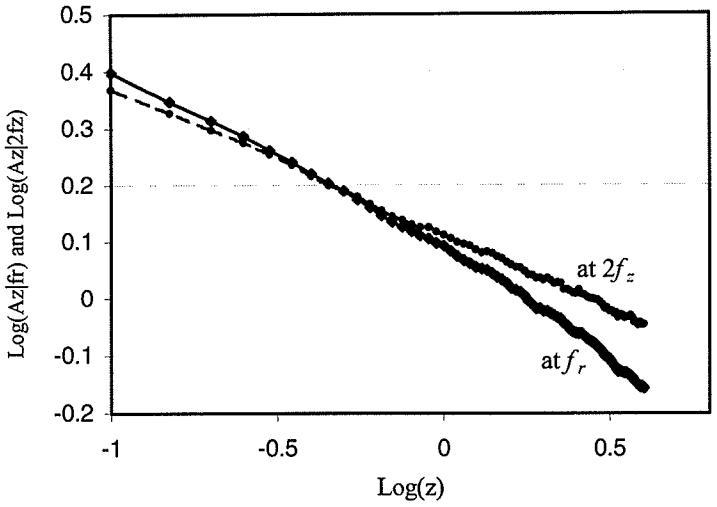

Fig. 3.30. The $\log(A_z|_{f_r})$ and the $\log(A_z|_{2f_z})$ as a function of the log(z).

3.5 High-speed pulse sampling measurements

3.5.1 Pulse sampling up to 1 Gbit/s

An implementation scheme used for the pulse sampling measurements up to 1 Gbit/s is shown in Fig. 3.31. The 1 ns pulse from the HP data generator was used to generate a high-speed sampling pulse through the pulse generation system, as discussed in the previous section. Up to 1 Gbit/s test patterns were also generated by the HP80000 data generator as the test circuit stimulus. The pulse was synchronised with the circuit stimulus and was amplitude-modulated at the cantilever's resonant frequency to apply repetitive time sampling. The amplitude modulation was done by a high-speed switch from Mini-Circuits. The switch control signal came from a function generator, which was also used to supply a reference signal for the lock-in amplifier. The pulse period, T_p, was usually set the same as that of the circuit signal, i.e. T_p=T_c=T.

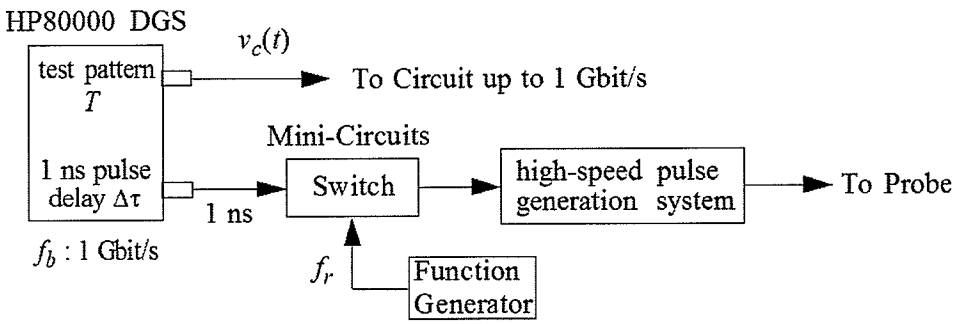

Fig. 3.31. The implementation of up to 1 Gbit/s sampling measurements.

Pulse sampling measurements were initially performed on a 300 μm wide microstrip transmission line mounted on a printed circuit board. The sample had approximately a 5 GHz bandwidth. The circuit input was from the HP80000 data generator at 1 Gbit/s which had about a 90 ps rise time and a 120 ps fall time. The circuit output was connected to the TDS 8000 Digital Sampling Oscilloscope for monitoring purposes. The probe was usually located at the middle of the transmission line with a typical tip-to-sample spacing of about 300 nm. The probe signal used here was the output from the high-speed pulse generation system and amplified by the SDL wideband amplifier with a FWHM of 57 ps (Fig. 3.18).

The sampling rate was 100 GSamples/s, (i.e. 10 ps delay increment) so that there were about 10 sampling points on the circuit signal transitions. A sample measurement result is shown in Fig. 3.32. The measured waveform is compared with the circuit output measured by the oscilloscope after performing a convolution with the 57 ps sampling pulse. For the 8 bit data (T=8 ns), the pulse duty time δ/T was about 0.007.

As seen in Fig. 3.32, the pulse nonidealities, such as its limited rise time, ringing etc., caused distortion in the measured waveform (the curve b). This is verified by comparing with the convolution (the curve a). As seen in this measurement, the actual pulse at the probe tip suffered a significant diffusion, as discussed in last section, which caused rounded corners and additional ripples.

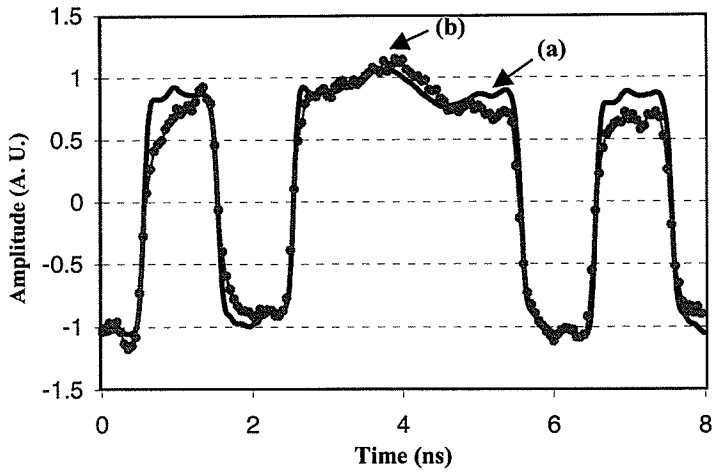

Fig. 3.32. The measurement of 1 Gbit/s signals applied to the microstrip transmission line. Curve (a) shows the circuit output measured by a 20 Gbit/s oscilloscope after performing a convolution with the 57 ps sampling pulse. Curve (b) is the wave form measured by the DEFM probe using the 57 ps sampling pulse.

Pulse sampling measurements were also performed on a 0.5 μm CMOS inverter chain for pattern extraction [67]. A photograph and the equivalent schematic of the integrated circuit are shown in Fig. 3.33 (a) and (b), respectively. The circuit contains three rows of inverter chains. The top two rows have 10×10 μm^2 unpassivated probing pads between every two inverters. The third row has only a passivated 1.2 μm interconnect. The packaged circuit was mounted onto a printed circuit board with a 50 Ω transmission line to transport the high speed digital input signal into the chip.

Here, a 250 Mb/s patter [1010 ...] was applied to the circuit. The measured wave forms are shown in Fig. 3.33 (c). Curve (1a) was measured on the test point 1 using a 0.04 pf 1 GHz active contact probe [5] for comparison. Curve (1b) was measured on the test point 1 by an EFM probe with a sampling pulse of FWHM 130 ps (the output pulse of the logic gate NLG4519-2). Curve (2) was measured on the test point 2 by the EFM probe to show that signals can be obtained directly from passivated interconnects. These measurements demonstrate the performance and the ability of the EFM probe.

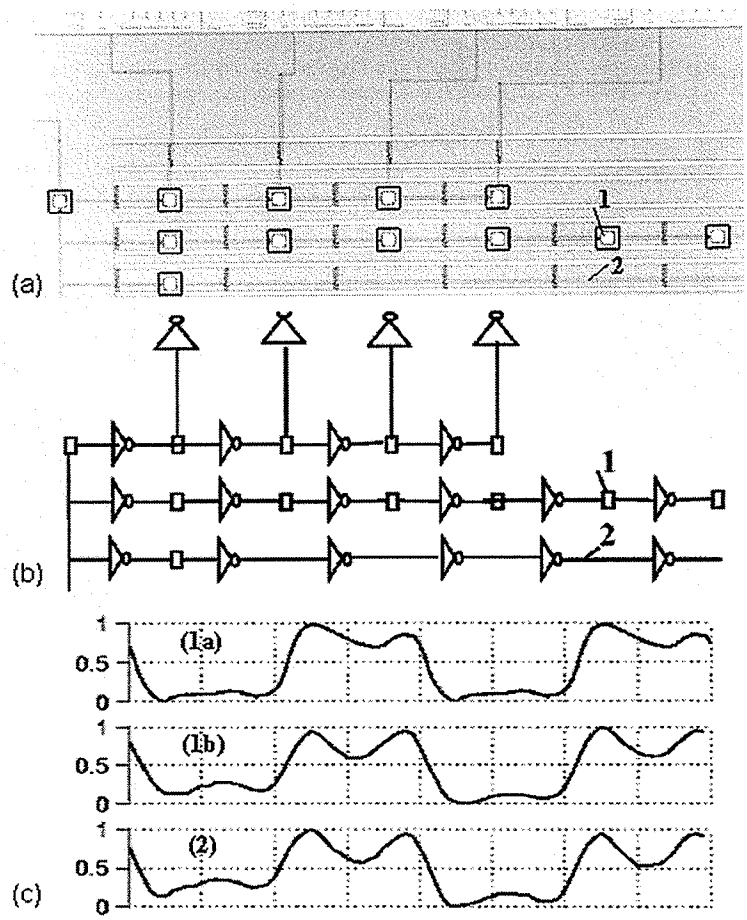

Fig. 3.33. A microphotograph of the CMOS inverter chain (a) and the circuit schematic (b). Measured wave forms on the test point 1 using an active probe (1a) and the EFM probe (1b) and the test point 2 (passivated) using the EFM probe (2) [67].

3.5.2 Pulse sampling up to 10 Gbit/s

The EFM pulse sampling technique is also used for the circuit signal measurements > 1 Gb/s. The measurement setup for up to 10 Gbit/s pulse sampling is shown in Fig. 3.34. The 37 ps sampling pulse as seen in Fig. 3.11 from the MOF19A logic gate was used for the measurements > 1 Gb/s. An HP83595 RF generator was used to generate up to 10 Gbit/s digital signals (through a high speed logic gate) for the circuit stimulus. To synchronise the probe signal with the circuit signal, a sinusoidal signal was also used as an external clock for the internal phase lock loop of the HP data generator, after being stepped down by a frequency divider. In this work, 2.5 Gbit/s and 10 Gbit/s sampling measurements were performed.

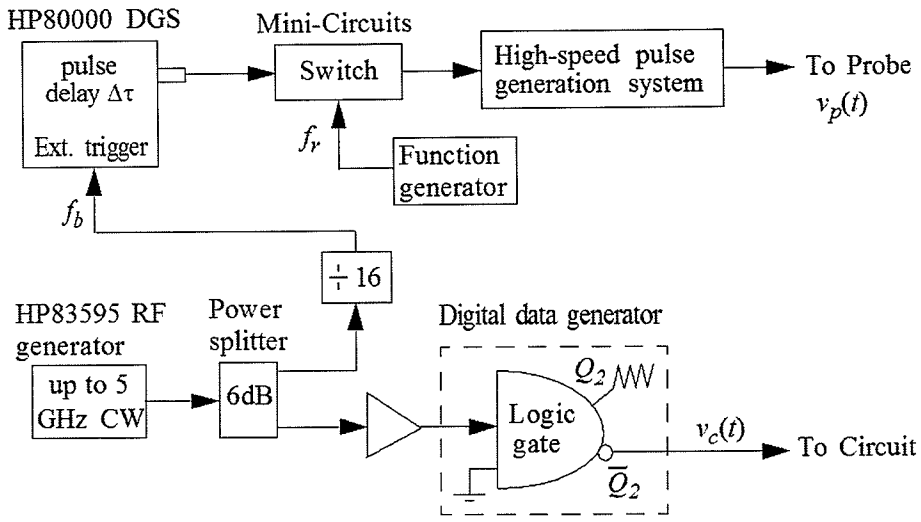

Fig. 3.34. The implementation for up to 10 Gbit/s sampling measurements.

For 2.5 Gbit/s sampling measurements, the RF generator output is a 1.25 GHz continuous wave and the HP phase lock loop external trigger signal f_b is 0.3125 GHz (divided by 8). The sampling pulse period, T_p=1/f_b =3.2 ns, and the circuit signal period T_c=800 ps. The pulse duty time δ/T_p is about 0.012 for the 37 ps sampling pulse.

For 10 Gbit/s sampling measurements, the RF generator output is at 5 GHz and f_b = 0.625 GHz (divided by 16). The sampling pulse period is T_p=1.6 ns and the circuit signal period is T_c=200 ps. The pulse duty time δ/T_p was about 0.023.

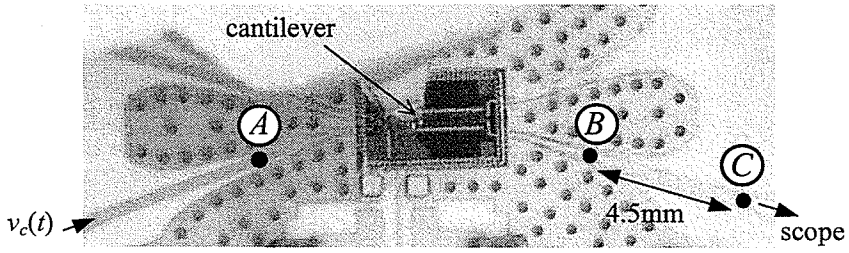

Fig. 3.35. A microphotograph of a GaAs wideband distributed amplifier indicating the test points on its input/ output off-die transmission lines.

Pulse sampling measurements were performed on an operating GaAs wideband distributed amplifier as shown in Fig. 3.35. The circuit was designed for OC-192 application but had a bandwidth of only about 10 GHz. It has a passivation layer approximately 300 nm thick. The circuit stimulus is a 2.5 Gbit/s square-wave signal (with +350 ps/-450 ps duration) from a high-speed logic gate (NLG4119). The probe was

placed about 100 nm above the circuit during the measurements. The EFM measurement on the off-die input transmission line (the test point *A*) is shown in Fig. 3.36 (dots) in comparison with the circuit input signal measured by a digital oscilloscope (the solid curve). The distortion in the measured wave form is caused by pulse nonidealities. The sampling rate is 200 GHz, i.e. 5 ps sampling step size.

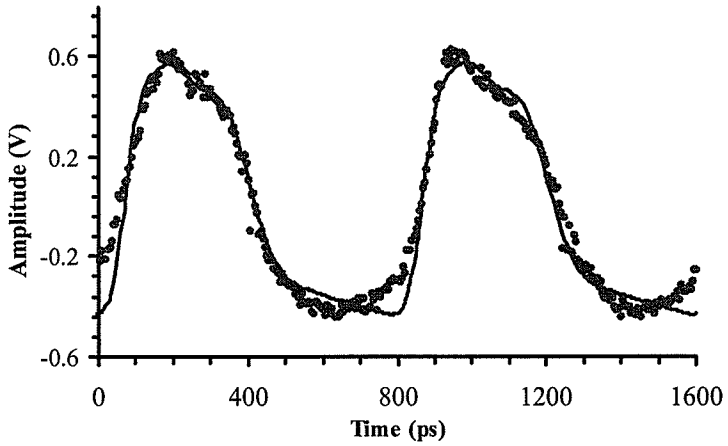

Fig. 3.36. EFM pulse sampled signal (dots) in comparison with the circuit input signal measured by a digital oscilloscope (the solid curve). The EFM measurement was performed on a wideband distributed amplifier's off-die input transmission line.

Figure 3.37 shows the EFM signal measured on the amplifier's off-die output transmission line, i.e. the test point *B* in Fig. 3.35. The signal has been smoothed using a low pass filter (which will be discussed in Chapter 4) and is to compare with the amplifier output signal measured by the scope (the solid curve). The amplifier was operated under saturation.

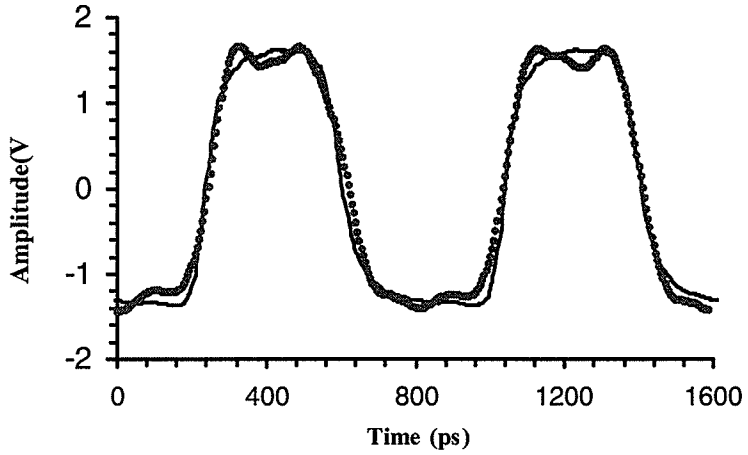

Fig. 3.37. The EFM pulse sampled signal (dots) in compared with the circuit input signal measured by a digital oscilloscope (the solid curve). The EFM measurement was performed on a wideband distributed amplifier's off-die output transmission line.

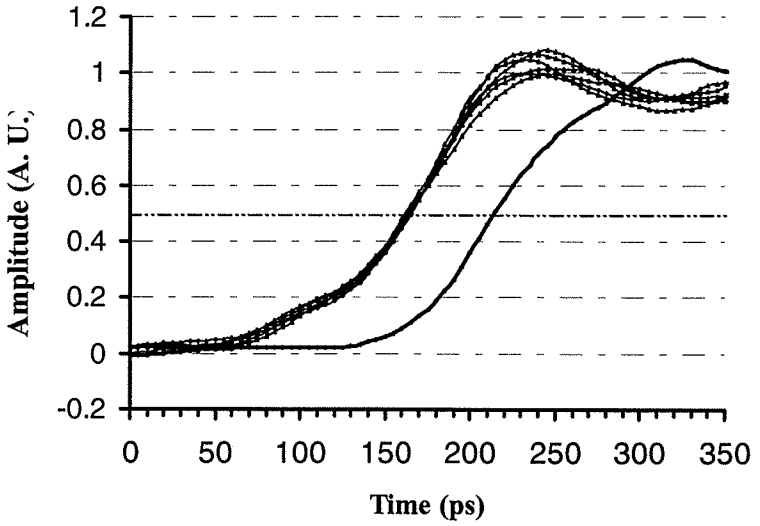

Fig. 3.38. EFM pulse sampled rising edges for delay measurement between the test point *B* and *C*. Six repeating measurements were performed on the test point *B* to indicate the measurement repeatability and phase jitter.

The delay between test points *B* and *C* (shown in Fig. 3.35) was estimated from the measured signals to be 45±5 ps as seen in Fig. 3.38 (with six repeated measurements on test point *B*). The measured distance between *B* and *C* is approximately 4.5 mm. From the 10 GHz vector voltage measurements, the effective dielectric constant for the transmission line is estimated to be about 6.0 and the calculated time delay between *B*

and C for a 2.5 Gbit/s signal is about 37 ps [68]. These measurements show that the EFM probe has the capability to measure internal circuit delay with a resolution of 10 ps.

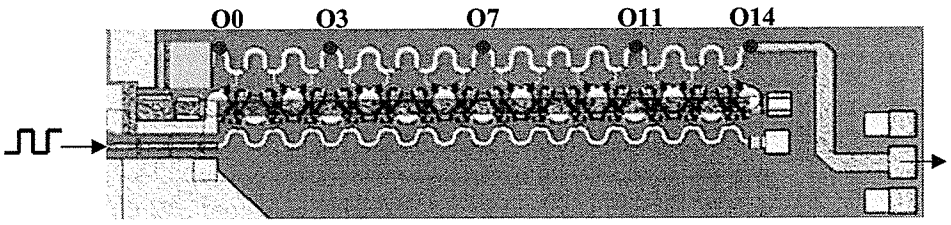

Fig. 3.39. A microphotograph of the wideband distributed amplifier with internal probing nodes on the output transmission line.

Internal circuit probing was also performed on the distributed amplifier's on-die output transmission line at the distributed amplifier as shown in Fig. 3.39. The width of the transmission line is approximately 20 μm. The EFM measurements for a 2.5 Gb/s input signal were performed on the test points O0, O3, O7, O11 and O14. The measured wave forms are all smoothed and shown in Fig. 3.40. The five wave forms look similar and are vertically shifted for comparison. The propagation delays between these measurements are indistinguishable since the delay from O0 to O14 (with a distance less than 2 mm) is less than 20 ps according to the delay measurement in Fig. 3.38.

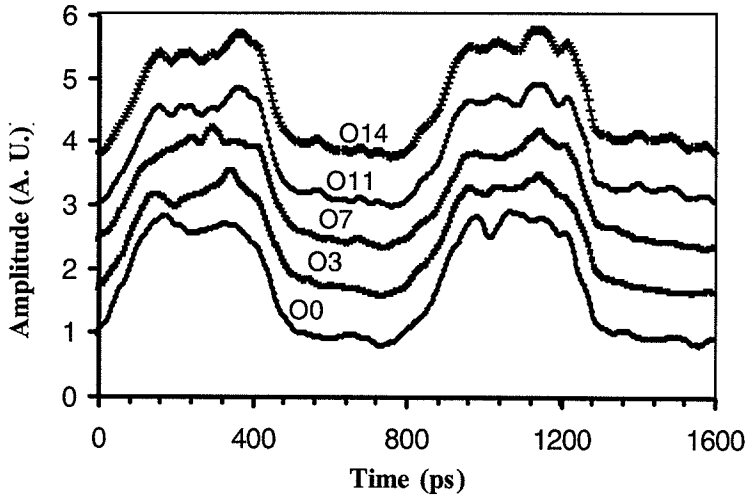

Fig. 3.40. The EFM pulse sampled wave forms on the five test points of the amplifier's output transmission line.

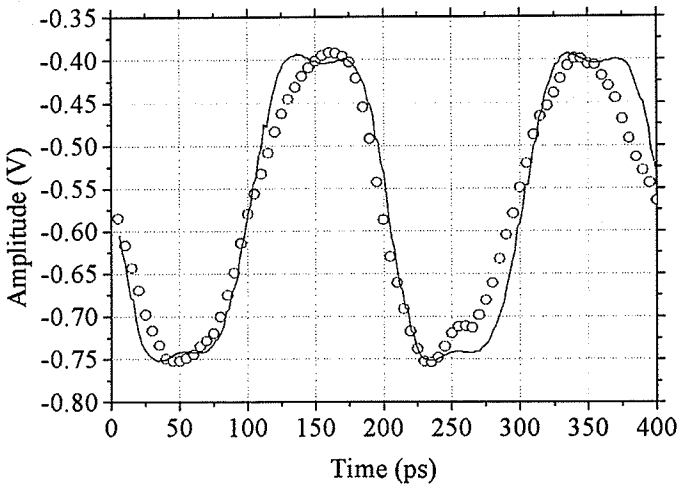

Fig. 3.41. The measured 10 Gbit/s signal (dots) on the amplifier's off-die input transmission line compared with the circuit input data measured by a 20 GHz digital oscilloscope (the solid curve).

10 Gbit/s pulse sampling measurements were performed on the amplifier's off-die input transmission line (test point *A*). The sampling rate was 200 Gsample/s (5 ps sample step). The sampled waveform is compared with the circuit input data measured by a 20 GHz digital sampling oscilloscope as shown in Fig. 3.41. The circuit input signal had about a 30 ps rise/fall time. The rise/fall time for the EFM sampled wave form was much worse due to the bandwidth limit of the sampling pulse and the circuit under test.

10 Gbit/s EFM pulse sampling measurements were also performed on the test point *B* while the amplifier was operated under saturation. Figure 3.42 shows the measured wave form in comparison with the circuit output measured by the oscilloscope. On-die measurements had similar results as in the Fig. 3.40. These measurements show that the amplifier has a bandwidth much less than expected by the designer (the output signal transient time increased significantly). The circuit has a limited bandwidth and has internal reflection problems as seen from the vector voltage measurements [68].

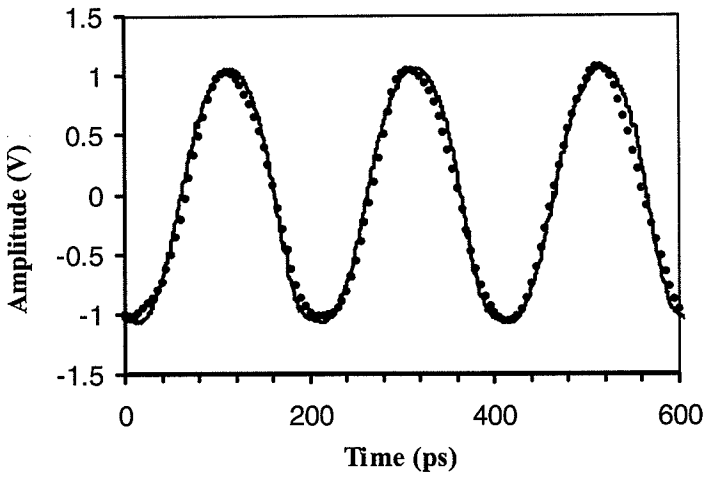

Fig. 3.42. The measured 10 Gbit/s signal (dots) on the amplifier's off-die output transmission line compared with the circuit output data measured by a 20 GHz digital oscilloscope (the solid curve).

Chapter 4

System Characterization and Improvement

Chapter 4 characterizes the DEFM pulse sampling system performance factors, such as the bandwidth, the temporal resolution, the phase shift linearity, the noise, the voltage sensitivity and the probe invasiveness. It describes a noniterative deconvolution process which could be used to compensate for the waveform distortion caused by the nonideality of the sampling pulse.

4.1 The system bandwidth and the rise time

The bandwidth of a dynamic measurement system is generally defined as the frequency range within which its frequency response is constant within which the deviations are of less than 3 dB. For an oscilloscope, this range usually starts at dc, and its upper limit is shown as f_{-3dB}, as illustrated in Fig. 4.1. It means that ≥ 30% error in amplitude measurements can be expected for frequencies above this frequency.

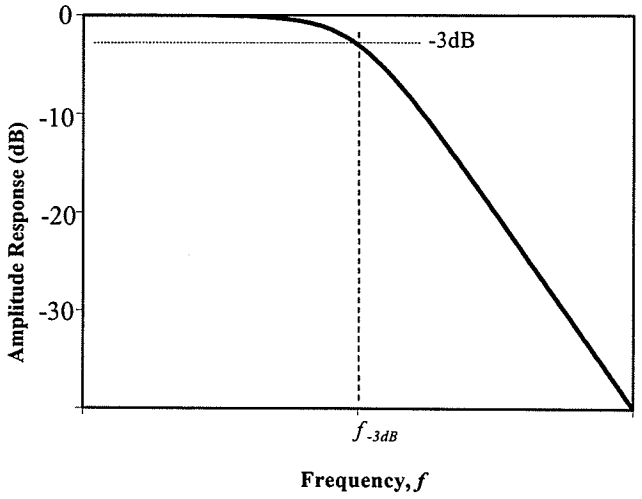

Fig. 4.1. The bandwidth of a dynamic measurement system.

The bandwidth of an EFM pulse sampling system is determined by the sampling pulse generated by the electronics and the circuitry for propagating the pulse to the tip, including the printed circuit board, the cantilever supporting chip and the cantilever, as discussed in the previous chapter. If $V_c(\omega)$, $H_p(\omega)$ and $H_g(\omega)$ are the Fourier transforms of the circuit signal $v_c(t)$, the pulse $v_p(t)$ and the pulse propagation circuitry frequency response respectively, then the measured deflection signal is

$$A_z(\tau) \propto F^{-1}[H_p(\omega)H_g(\omega)V_c(\omega)], \tag{4.1}$$

where F^{-1} is an inverse Fourier transform operation. The frequency response of the pulse propagation circuitry, $H_g(\omega)$, has been discussed in the previous chapter. In this section, we are going to discuss the sampling pulse frequency response, $H_p(\omega)$.

It is interesting to see how close the EFM sampling pulse frequency response is to a Gaussian frequency response. The 37 ps sampling pulse (dots) used in the EFM system (in the Fig. 3.11) is similar to a Gaussian pulse (solid curve) as seen in Fig. 4.2 except for some small ringing (< 5% energy of the pulse).

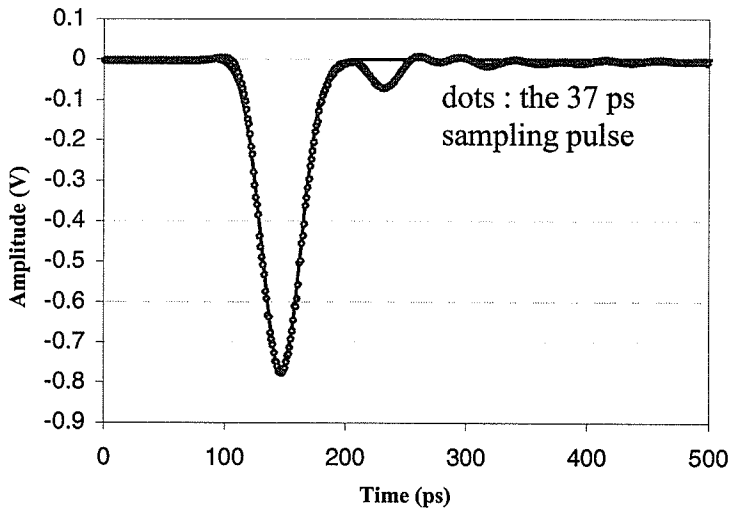

Fig. 4.2. The 37 ps sampling pulse (dots) used in the EFM system is similar to a Gaussian pulse (the solid curve) except for some small ringing.

The calculated frequency response $H_p(\omega)$ for the 37 ps sampling pulse in comparison with the Gaussian pulse frequency response is shown in Fig. 4.3. The sampling pulse behaves as a low pass filter with a 3dB bandwidth of > 8 GHz. However, due to the non-ideal Gaussian pulse shape, some high frequencies can pass through as shown in Fig. 4.3.

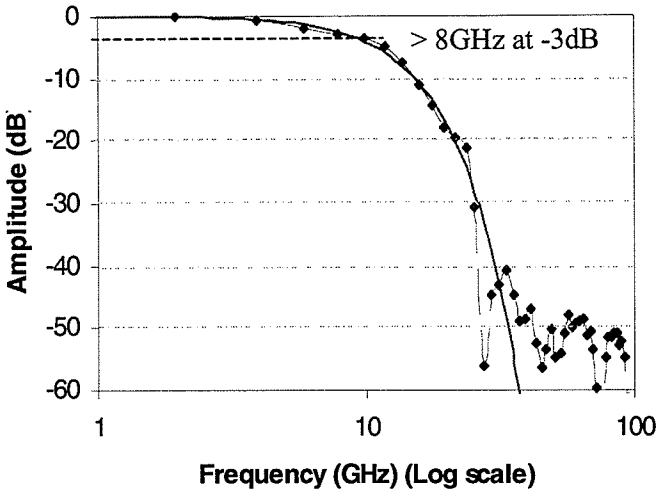

Fig. 4.3. The frequency response of the 37 ps sampling pulse (dots) in comparison with the Gaussian frequency response (the solid curve). The pulse has a 3 dB bandwidth of > 8 GHz.

A Gaussian impulse response and its frequency response are

$$g(t) = \frac{1}{\sqrt{\pi}\tau} e^{-t^2/\tau^2} \text{ and } G(\omega) = e^{-\omega^2\tau^2/4}, \tag{4.1}$$

respectively. τ is the standard deviation.

From the Gaussian impulse response, the full-width half-maximum (FWHM), δ, is

$$\delta = 2(\ln 2)^{1/2}\tau. \tag{4.2}$$

The bandwidth of a Gaussian frequency response can be calculated to be

$$f_{-3\text{dB}} = \frac{\sqrt{2\ln 2}}{2\pi\tau} \approx \frac{0.31}{\delta}. \tag{4.3}$$

From this equation, theoretically, the bandwidth of the 37 ps sampling pulse $f_{-3\text{dB,pulse}}$ = 8.4 GHz. This is consistent with the Fig. 4.3.

The term rise time t_r usually denotes the time a step response needs to rise from 10% to 90% of its final value. The rise time of a step response with a Gaussian frequency response (4.1) can be computed to be,

$$t_{r,10-90} \approx 1.81 \cdot \tau. \tag{4.4}$$

It is related to the 3dB bandwidth by the formula,

$$t_{r,10-90} (\approx 1.09\delta) \approx 0.34 / f_{-3\text{dB}}. \tag{4.5}$$

For high speed oscilloscopes, the 20% to 80% rise time definition is also commonly used since the 20% to 80% rise time is less affected by aberrations than the 10% to 90% rise time and, consequently, often exhibits a smaller standard deviation. The theoretical relationship between the 20% to 80% rise time and the 3dB bandwidth for a Gaussian system is,

$$t_{r,20-80}(\approx 0.72\delta) \approx 0.22/f_{-3\text{dB}}. \tag{4.6}$$

In this thesis, the 10% to 90% rise time definition is used by default.

The frequently used transformation formula for the bandwidth and the rise time of a measurement system is actually derived from a low-pass of the first-order (the transfer function is $1/[1+j\omega/\omega_{3\text{dB}}]$) [69]:

$$t_{r,10-90} \approx 0.35/f_{-3\text{dB}}. \tag{4.7}$$

For an actual measurement system, assuming a low-pass of second-order (the transfer function is $1/[1+2D(j\omega/\omega_{3\text{dB}})+(j\omega/\omega_{3\text{dB}})^2]$ with the damping factor D of 0.7) leads to a considerably better approximation [69]. For the second-order frequency response, the bandwidth-rise time product is near 0.34 (which is close to that of a Gaussian frequency response, (4.5)) [69]. A comparison of the frequency response for the three systems with the same 3 dB bandwidth is shown in Fig. 4.4.

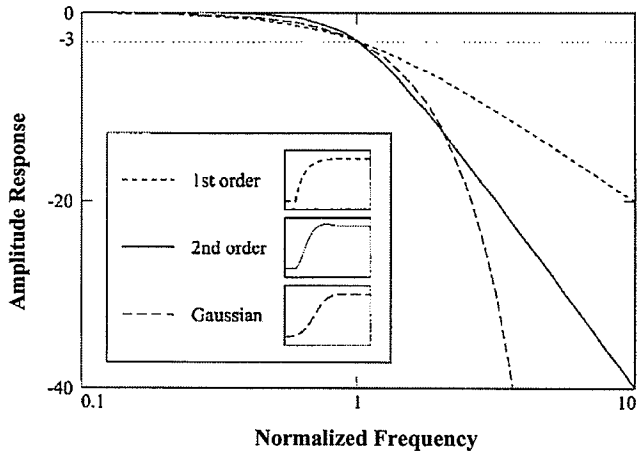

Fig. 4.4. The frequency responses for the first-order low-pass, the second-order low-pass and the Gaussian system (step responses are shown in the legend) [69].

The bandwidth of a system limits the rise time measurement accuracy of transient signals. Assuming the measurement system has a Gaussian frequency response and the signal has a Gaussian signal shape, the measured signal rise time can be estimated by

$$t_{r,meas} = \sqrt{t_{r,signal}^2 + t_{r,sys}^2}, \tag{4.8}$$

where $t_{r,signal}$ ant $t_{r,sys}$ are the rise time of the signal and the measurement system, respectively [69]. This equation is also valid for the 20% to 80% rise time definition. From this equation, the rise time measurement error due to the system bandwidth limit is small for $t_{r,signal} > 3\ t_{r,sys}$ (< 6%) and large for $t_{r,signal} \sim t_{r,sys}$ (> 40%).

Figure 4.5 shows a sample of the rising edge measurements of the HP80000 data generator output signal. The curve (1) was measured directly by a 20 GHz sampling oscilloscope. The curve (2) was measured on a 50 Ω transmission line by the EFM probe (with the sampling pulse shown in the Fig. 3.10).

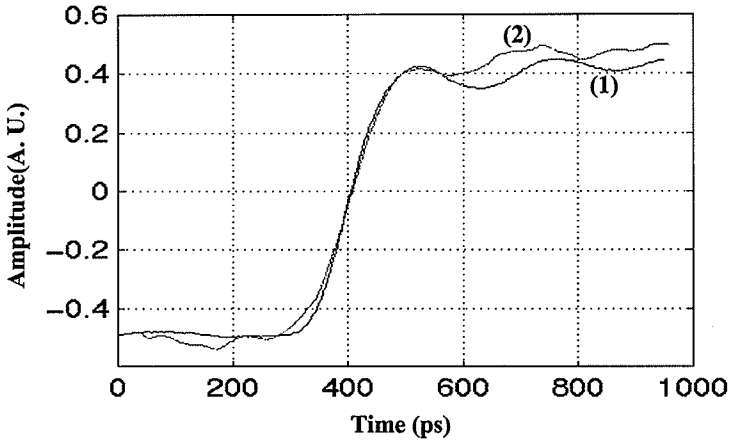

Fig. 4.5. The measured results for the rising edge of the HP data generator's output signal. The curve (1) was measured directly by a 20 GHz sampling oscilloscope. The curve (2) was measured on a 50 Ω transmission line by the EFM probe.

In the first measurement, $t_{r,signal}$ (~120 ps) > 3 $t_{r,sys}$ ($t_{r,sys}$ ~17 ps). The measured signal rise time is $t_{r,meas} \sim t_{r,signal}$. The second measurement shows the rise time measurement error (~17%) due to the EFM sampling system bandwidth limit.

It seems to be promising to use equation (4.8) as a correction term to recalculate the $t_{r,signal}$ from the $t_{r,meas}$,

$$t_{r,signal} = \sqrt{t_{r,meas}^2 - t_{r,sys}^2}. \tag{4.9}$$

However, the idea is not very practicable due to the noise, the reading error, the non-ideal Gaussian shape signal and the EFM sampling system frequency response. The accuracy of the estimation given by (4.9) is limited even for a measurement system with much higher time resolution.

As previously discussed, the EFM sampling system frequency response is determined by the frequency response of the sampling pulse and the pulse propagation circuitry. It can be calibrated by using a higher standard internal probing system. Unfortunately, it is not currently available.

The rise time measurement of a known signal could be used for a rough estimation of the EFM sampling system's bandwidth, i.e. the overall bandwidth of the EFM sampling system is

$$f_{-3dB,EFM} \approx 0.34/\sqrt{t_{r,meas}^2 - t_{r,signal}^2}. \tag{4.10}$$

The overall bandwidth of the EFM sampling system could be used for further estimation of the bandwidth of the pulse propagation circuitry, $f_{-3dB,prop}$, using

$$f_{-3dB,prop} = 1/\sqrt{1/f_{-3dB,pulse}^2 - 1/f_{-3dB,EFM}^2}, \tag{4.11}$$

assuming the frequency response of the pulse propagation circuitry, $H_g(\omega)$, is also a Gaussian response.

In this example (Fig. 4.5), assuming the circuit signal rise time is the same as that of the data generator's output signal, the roughly estimated overall bandwidth of the EFM sampling system with the sampling pulse as shown in Fig. 3.18 is 5 GHz. Theoretically, the bandwidth of the 57 ps sampling pulse $f_{-3dB,pulse}$ is about 5.4 GHz. From the equation (4.11), the bandwidth of the pulse propagation circuitry is > 10 GHz. This means that the sampling pulse response is dominant for the EFM sampling measurement up to 10 Gbit/s.

Theoretically, the bandwidth of the 37 ps sampling pulse is > 8 GHz. The rise time of the pulse signal is $t_{r,pulse} \leq t_{r,signal}$ for measuring a 10 Gb/s circuit signal, which means the estimation of the rise time of the circuit signal from the measured rise time is difficult. A faster sampling pulse generator has to be used. For measurements above 10 Gbit/s, the pulse propagation circuitry's frequency response has to be improved as well.

4.2 Deconvolution

The EFM measurement temporal resolution is determined by the system bandwidth. If the system bandwidth is not sufficiently high, the measurements are distorted and some high frequency components in the original signal are lost. When the system response is known, a deconvolution process [70] can be used via inverse filtering to remove the effects of the system response and improve the temporal resolution of the measurement to obtain a more accurate estimation of the original signal.

However, with inverse filtering, noise in a measurement system becomes highly amplified at frequencies where the system response approaches zero. Various filtering approaches are used to control this behaviour, but any filter necessarily removes some components of the original signal that is being recovered. The main concern of a deconvolution process is in finding an optimum filter. In this work, the analytic high order Gaussian filters and an optimization criteria based on the least mean-squared error [70] were used for the initial simulation of the deconvolution process.

4.2.1 Background

For a sampling system with an impulse response $p(n)$, the output resulting from an input $v(n)$ is given by

$$m(n) = v(n) \otimes p(n) + n(n), \tag{4.12}$$

where $\otimes$ designates a convolution operator, $n(n)$ represents the added measurement noise and n is the discrete time. In the frequency domain, this is represented as

$$M(k) = V(k)P(k) + N(k), \tag{4.13}$$

where $M(k)$, $V(k)$, $P(k)$, and $N(k)$ are the discrete Fourier transforms of $m(n)$, $v(n)$, $p(n)$, and $n(n)$ respectively and k is the discrete frequency indices..

Although the noise $N(k)$ is unknown, an estimate of the original signal can be obtained from a direct deconvolution in the frequency domain as

$$V_e(k) = \frac{M(k)}{P(k) + N_e(k)}, \tag{4.14}$$

where $N_e(k)$ includes any errors in the estimate of $V(k)$ and is usually small.

A frequency-domain low-pass filtering team $F(k)$ has to be used for the raw deconvolution result given by Eq. (4.14) to avoid the amplification of noise. Thus Eq. (4.14) becomes

$$V_e(k) \approx \frac{M(k)}{P_e(k)} F(k), \tag{4.15}$$

where $P_e(k)=P(k)+N_e(k)$. The commonly used frequency-domain filtering team, $F(k)$, is

$$F(k) \approx \frac{|P_e(k)|^2}{|P_e(k)|^2 + \gamma |L(k)|^2}, \tag{4.16}$$

where $|L(k)|^2$ is the squared magnitude of the discrete Fourier transform of the second difference operator and may be written as

$$|L(k)|^2 = 6 - 8\cos(\frac{2\pi k}{N_0}) + 2\cos(\frac{4\pi k}{N_0}), \qquad k = 0,1,2,...,N_0 - 1, \tag{4.17}$$

where N_0 is the number of sampled points [71]. Variable γ is called the regularization parameter.

In this work, filters of the form given in Eq. (4.16) were used

$$F(k) = e^{-|k/\sigma|^m/2}, \tag{4.18}$$

where $m>1$ (m=2 for the Gaussian form), and σ is a variable related to the cutoff frequency of the filter. Two example filters are shown in Fig. 4.6. All these filters have a linear phase response. It has been found that these filters performed best for $m \geq 2$ (the time-domain ringing produced by the higher order filters dies out quickly [70]). The sharp cutoff afforded by larger values of m allowed the cutoff to be extended to higher frequencies, resulting in less attenuation of the signal components [70].

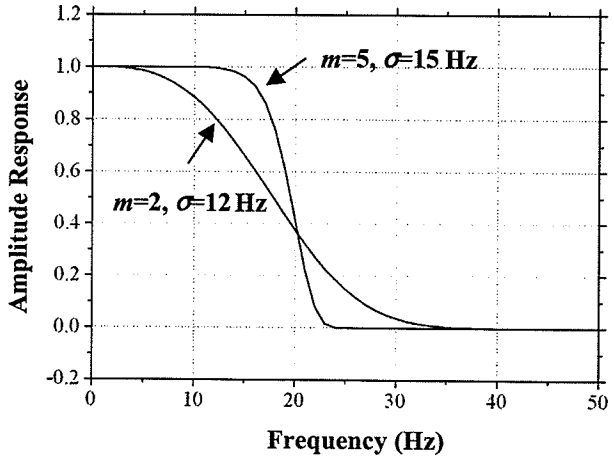

Fig. 4.6. The two filters of the form given in Eq. (4.16).

In this work, the criteria used for the optimization of the parameter m and σ is based on the least mean-squared error between the estimate and the true signal, yielding a cost function [70]

$$Cost = \frac{1}{N_0} \sum_{n=0}^{N_0-1} [v_e(n) - v(n)]^2, \tag{4.19}$$

to be minimized, where $v_e(n)$ represents the deconvolution output and N_0 is the total sampled number. Since the real signal $v(n)$ is unknown, an indicated error $i(n)$ given by Eq. (4.20) could be used to substitute $v_e(n)$-$v(n)$ [70]

$$i(n) = F^{-1}[\frac{M(k)}{\hat{P}(k)} F(k)(1-F(k))]. \tag{4.20}$$

4.2.2 Simulations

Simulations were run using a Gaussian pulse with a small ringing (imitating the 37 ps sampling pulse as shown in Fig. 4.2) as $p(n)$ and using a Gaussian-shape square wave as $v(n)$. The "measured" signal, $m(n)$, was analytically generated from the convolution of $p(n)$ and $v(n)$ plus some uniform white noise $n(n)$.

Figure 4.7 shows an example of an input signal $v(n)$ (dots) with a rise time t_v = 20 ps and a period of 1 ns, the "measured" signal $m(n)$ (the curve a) with a prearranged SNR of 60 and the deconvolution output $v_e(n)$ (the curve b), which was optimized at m = 3 and σ =14 GHz (the cutoff frequency is about 13 GHz) for the minimum cost.

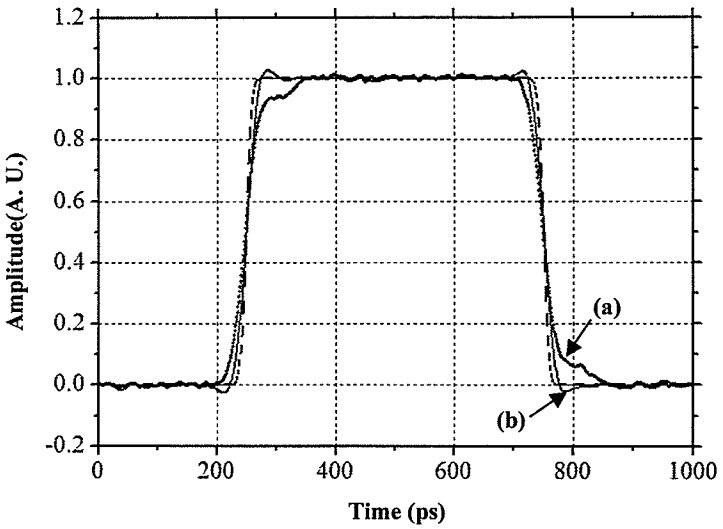

Fig. 4.7. Curve (a) is the "measured" signal $m(n)$, which is a convolution of the Gaussian pulse $p(n)$ and $v(n)$ (the input signal with a 20 ps transition time (dots)). White noise was added with a prearranged SNR of 60. Curve (b) is the deconvolution result.

Figure 4.8 is a zoom-in picture of the rising edges of the three curves shown in Fig. 4.7. The rise time of the deconvolution output (the curve b), t_{ve}, is approximately 35 ps. It is much smaller than the rise time of the "measured" signal (the curve a), t_m ($= \sqrt{t_p^2 + t_v^2}$ = 45 ps, where t_p = 1.09*37 ~ 40 ps). The measurement bandwidth is equivalently improved from 8.4 GHz (for the 37 ps sampling pulse) to approximately 12 GHz. Figure 4.7 also shows that the shoulders on the measured signal curve (a) (due to the contribution of the small pulse ringing) have disappeared. This means that the deconvolution process helps removing the effects of the pulse response from the "measured" signal.

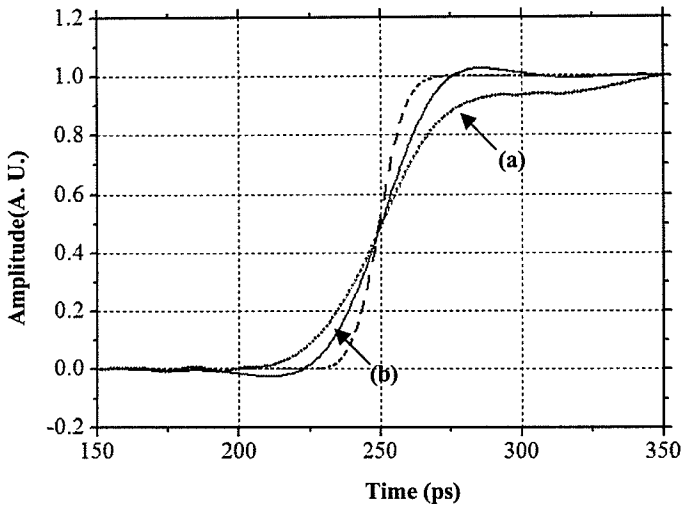

Fig.4.8. Curve (a) is the rising edge of the "measured" signal. Curve (b) is the deconvolution result with a rise time of approximately 34 ps. The input signal is shown as dots.

Unfortunately the actual pulse at the EFM probe tip is unknown since the frequency response of the pulse output circuitry, $H_g(\omega)$, is difficult to be determined as discussed in Chapter 3. Ideally, if the input signal $v(t)$ is known, the deconvolution process could also be used to find out the pulse $p(t)$ from the measured signal $m(t)$, similar to the Eq. (4.15):

$$P_e(k) \approx \frac{M(k)}{V(k)+N_e(k)} F(k). \tag{4.21}$$

To simulate the process, the previous 20 ps rise time input signal $v(n)$ and the "measured" signal $m(n)$ (= $v(n) \otimes p(n)+n(n)$, assuming a SNR of 600) were used. The optimized deconvolution output, $p_e(n)$, is shown in Fig. 4.9 (solid curve) in comparison with the sampling pulse, $p(n)$, which is a Gaussian pulse with a small peak. It is found that it is much more difficult to do the pulse deconvolution than the input signal deconvolution. A larger SNR is usually required for the pulse deconvolution and the deconvolution output is usually associated with the more ringing as shown in Fig. 4.9 since the higher order Gaussian filters have to be used to minimize the cost.

These simulations suggest that the deconvolution process could be used for improving the EFM measurement bandwidth and for finding the pulse at the probe tip. However, the deconvolution process creates errors as discussed in reference [70].

So far, the deconvolution process has shown very little improvement in terms of the measurement bandwidth for the actual EFM pulse sampled results. Further investigation is needed.

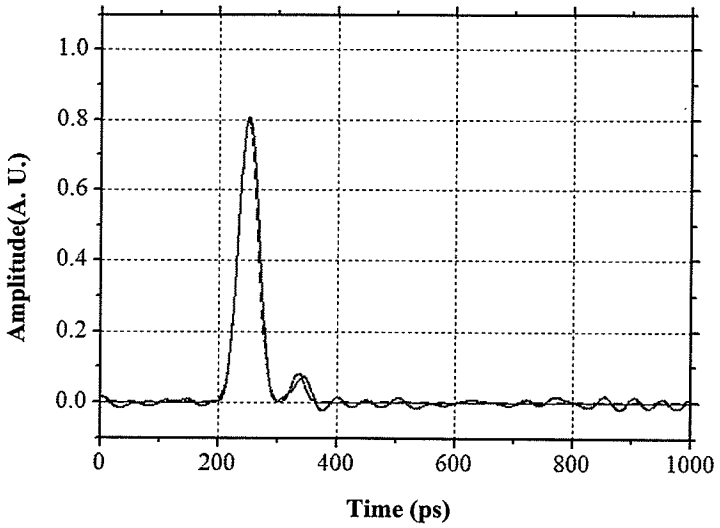

Fig. 4.9. The solid curve is the deconvolution output of the pulse. The dashed curve is a Gaussian pulse with a small peak imitating the 37 ps sampling pulse.

4.3 The delay resolution

The delay resolution is the minimum propagation delay detectable by the instrument. The delay resolution of the EFM instruments used in this work is mainly determined by the HP80000 data generator's internal delay line. The delay line has a minimum delay increment of 2 ps, however its phase jitter is typically 10 ps [56].

The amplitude noise of the sampling pulse is another source of the measurement temporal error. The effect of the amplitude noise on the temporal error is illustrated in Fig. 4.10. The voltage sensitivity and the slope of the signal transition time can be utilized to formulate the relationship between the amplitude noise and the timing ambiguity [72] as

$$\Delta T = \varepsilon \cdot \frac{dt}{dv}, \tag{4.18}$$

where ΔT is the resulting phase jitter from the amplitude error ε, and dt/dv is the inverse slope of the signal transition.

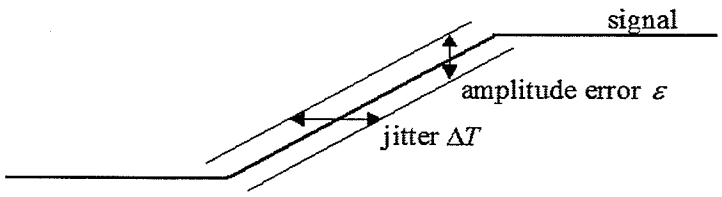

Fig. 4.10. The effect of the amplitude noise on the temporal inaccuracy.

The actual delay resolution of the EFM measurements should be an accumulation of the jitter ΔT caused from the amplitude noise and the delay line's phase jitter. The phase lock loop used in Fig. 3.34 for synchronizing the pulse signal with the circuit stimulus might add in extra phase jitter to the measurements.

The delays from the HP80000 data generator's delay line were carefully measured by using a 20 GHz digital sampling oscilloscope. The measured delay versus the expected delay is shown in Fig. 4.11. It shows that the delay line has a fairly poor linearity and that errors of as much as 50 ps exist for some points.

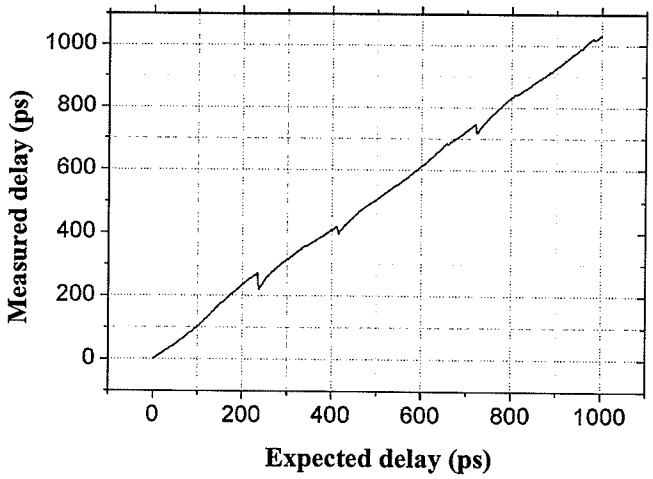

Fig. 4.11. The measured delay versus the expected delay from the HP80000 data generator's delay line.

To solve the problem, software was used to calibrate the delay by using a lookup table based on the measurement. The measured delay versus the expected delay after the software calibration was measured again by the scope as seen in Fig. 4.12. It is found that the improved delay accuracy is better than 10 ps. The software was used for all other measurements in this work.

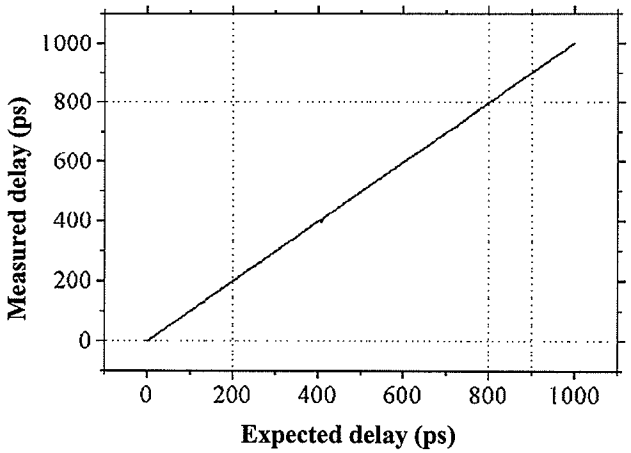

Fig. 4.12. The measured delay verses the expected delay after the software calibration.

Using the software, the EFM delay measurements were performed on a microstrip transmission line. The circuit input signal was a 2.5 Gbit/s square wave. The sampling pulse was the EFM 37 ps pulse. Five rising edge measurements were shown in Fig. 4.13 to correspond with the circuit signal's initial delay of 0, 10, 20, 30 and 50 ps.

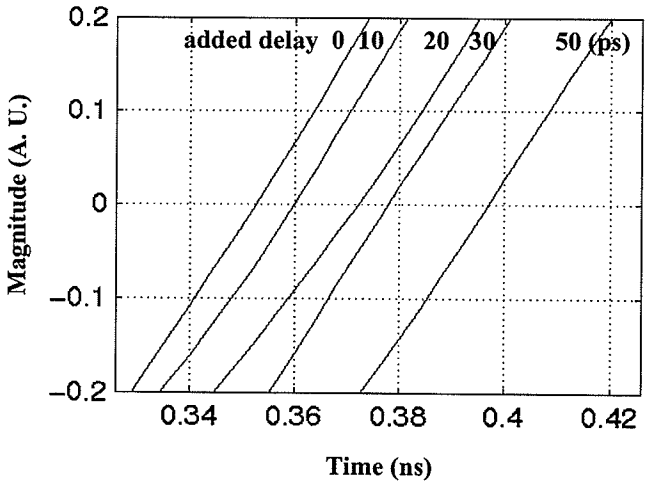

Fig. 4.13. The EFM delay measurements of 2.5 Gbit/s signal with the added initial delays of 0, 10, 20, 30, and 50 ps on a microstrip transmission line.

The same delay measurements were also performed on the Nortel distributed amplifier as shown in Fig. 3.35. Five initial delays of 0, 10, 20, 40 and 90 ps were added to the circuit input signal. The measured rising edges on the test point *B* corresponding to these delays are given in Fig. 4.14. Fig. 4.15 shows five measurements for the circuit under the same conditions to demonstrate the measurement repeatability.

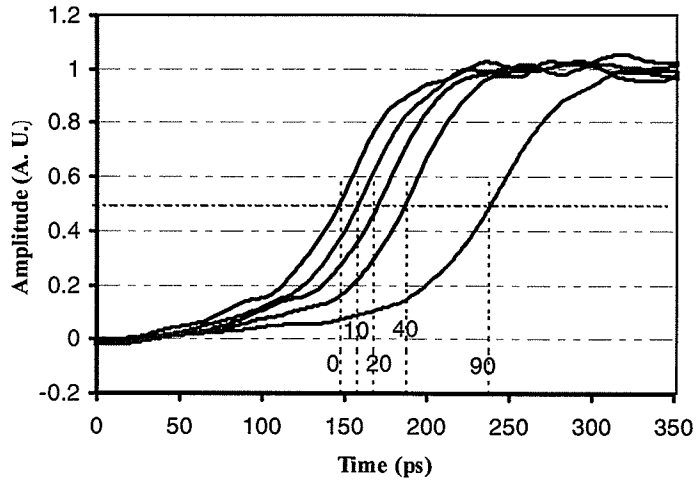

Fig. 4.14. The EFM delay measurements of 2.5 Gbit/s signal with the added initial delays of 0, 10, 20, 40, and 90 ps on the Nortel distributed amplifier.

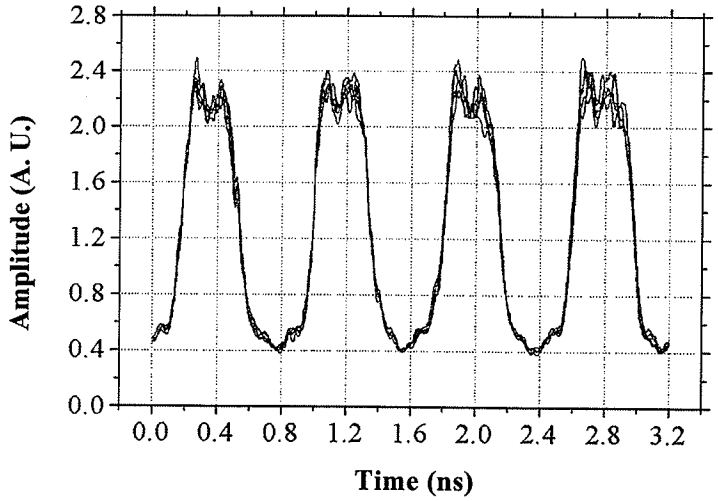

Fig. 4.15. Five measurements to demonstrate the measurement repeatability.

These measurements show that EFM has the capability to measure internal circuit signal delays with a delay resolution of less than 10 ps.

4.4 Noise and voltage sensitivity

4.4.1 Noise and voltage sensitivity

There exists mechanical noise from the probing structure, thermal noise, laser source noise, electrical noise coupled to the detection electronics, and $1/f$ noise [73]. With good design, mechanical noise, laser and electrical noise can be made negligibly small. The fundamental noise source for the SPM based techniques is the mechanical noise vibration

of the cantilever due to its equilibrium with the finite temperature environment. The cantilever vibrates with a total thermal energy of $k_BT/2$, where T is the temperature of the thermal environment (in Kelvins) and k_B is the Boltzmann's constant [73].

Figure 4.16 is an example of the measured noise power spectra by using a spectrum analyser. The noise floor was measured without input signals. The electrical noise was measured when the laser was turned off. When the probe was touching the sample surface, the measured spectrum included the overall mechanical noise (the contact probe formed a loop between the probe station and the probe head), electrical noise and laser noise. The thermal-mechanical noise spectrum (the noncontact probe) has the same Lorentzian shape as the cantilever frequency response except for the mechanical noise at low frequencies.

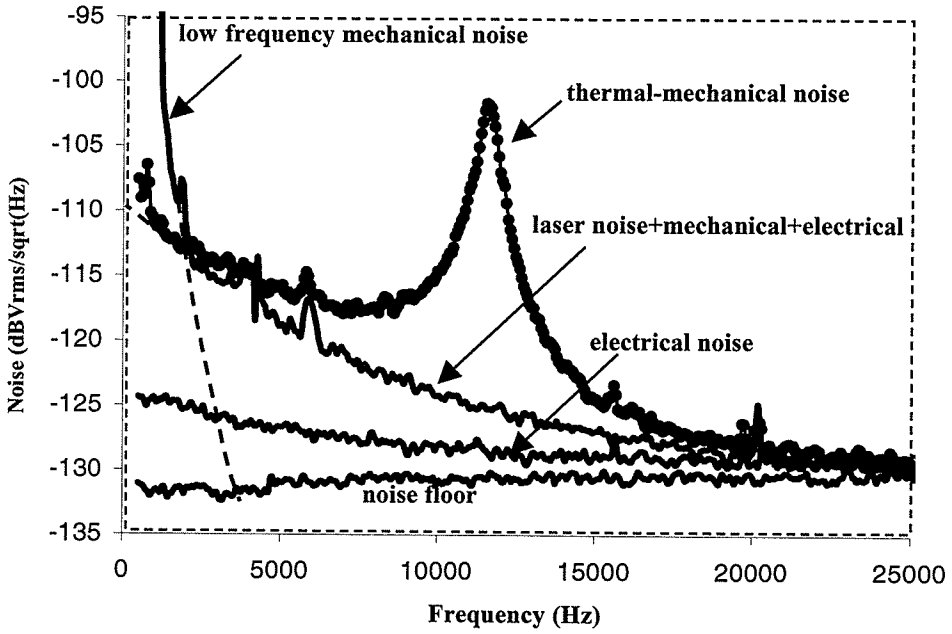

Fig. 4.16. The measured noise power spectra using a spectrum analyzer.

As seen in Fig. 4.16, the total noise for the EFM measurement system (measuring the desired signal at the cantilever resonance frequency) is usually dominated by the thermal-mechanical noise (the thermal peak is 20 dB higher or 10 times larger than the other noise). At the resonance, the thermal noise induced cantilever deflection is also enhanced by the cantilever's mechanical response, the same as the EFM force induced cantilever deflection.

Since the EFM measurements are made at the resonance frequency of the cantilever, the RMS thermal noise is approximately [73]

$$[\Delta z_{th}^2 |_{\omega_r}]^{1/2} = \sqrt{\frac{4k_B TQB}{k\omega_r}}, \tag{4.19}$$

where B is the deflection measurement bandwidth. When the lock-in technique is used for the desired signal, B is determined by the lock-in amplifier's time constant setting.

For the silicon cantilever used in this work, Q~20, k~0.055 N/m, f_r~11.675 kHz, and B ~ 1 Hz. Using Eq. (4.19) and the photo detector's transfer function (~1 V/μm), the calculated thermal noise is approximately 0.01 nm/$\sqrt{\text{Hz}}$. This is in close agreement with the measured noise value (-102 dBV$_{\text{rms}}$/ $\sqrt{\text{Hz}}$ ~ 0.01 nm/ $\sqrt{\text{Hz}}$) at the cantilever's resonance frequency.

The EFM measurement voltage sensitivity is defined as the minimum root-mean-square (RMS) detectable voltage (or amplitude) that is limited by the thermal noise. From the Eq. (4.19) and (2.31), the minimum detectable voltage amplitude (rms) is

$$V_{c,rms}(\tau)\Big|_{\min} = \frac{\pi}{2} \frac{1}{\frac{\partial C}{\partial z}\frac{\delta}{T} V_p} \sqrt{\frac{4k_B TkB}{Q\omega_r}}. \tag{4.20}$$

This expression indicates that the mechanical properties of the cantilever greatly affect the noise performance of the instrument. For a better noise performance, the cantilever should have a large Q and a small spring constant k. For example, for the case where the tip-to-sample distance ~ 200 nm, $\partial C/\partial z$ ~ 80 pF/m (estimation), and the pulse duty cycle is δ/T = 1/8 and V_p = 1 V, the estimated voltage sensitivity is about 2 mV/$\sqrt{\text{Hz}}$.

Experimentally, the voltage sensitivity was measured using the 1 ns pulse with δ/T = 1/8 and V_p = 2 V. The tip-to-sample spacing was about 800 nm on an unpassivated transmission line. The lock-in amplifier time constant was 30 ms. The RMS noise was measured to be about 11 mV, which was found out by comparing it with the measurement of known voltage. The measured EFM voltage sensitivity was approximately 5 mV/$\sqrt{\text{Hz}}$. It is greater than the estimated value 2 mV/$\sqrt{\text{Hz}}$. The error may come from the estimation of the $\partial C/\partial z$ value and the tip-to-sample distance.

The SNR of the measurement is defined as

$$SNR = \frac{[A_z^2|_{\omega_r}]^{1/2}}{[\Delta z_n^2|_{\omega_r}]^{1/2}} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N} A_{zi}^2|_{\omega_r}}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} \Delta z_{ni}^2|_{\omega_r}}} \tag{4.21}$$

where $[A_z^2]^{1/2}$ is the RMS signal (or signal amplitude) induced by the electrostatic force; $[\Delta z_n^2]^{1/2}$ is the RMS noise mainly determined by the thermal noise and N is the total number of sampled points. The minimal detectable signal is limited by the noise at SNR=1.

4.4.2 The noise filtering

The high-frequency noise in the EFM sampled signal can be filtered out by taking advantage of an oversampling method. The Nyquist criterion [74] requires the sampling frequency to be at least twice the maximum frequency of the measured signal to avoid aliasing. In EFM, the maximum sampling rate could be 500 Gsample/s (using 2 ps delay increment) which is much higher than the maximum frequency of the desired signals. This oversampling method could be taken advantage of for filtering out unwanted high frequency noise and to improve the measurement signal-to-noise ratio.

Usually, Butterworth filters were used for the noise filtering with a transfer function as

$$H(f) = \frac{1}{1+0.414(f/f_c)^{2n}}, \tag{4.22}$$

where f_c is the cutoff frequency. Equation (4.22) is known as n^{th} order Butterworth low pass filter response. Butterworth filters are causal and have linear phase response. The Butterworth filters of 2nd, 3rd, and 5th order with a cutoff frequency of 10 GHz are shown in Fig. 4.17. Usually, the cutoff frequency is selected to be two or three times larger than the system bandwidth. For the 37 ps pulse sampling measurement, the filter's cutoff frequency was selected to be 20 GHz.

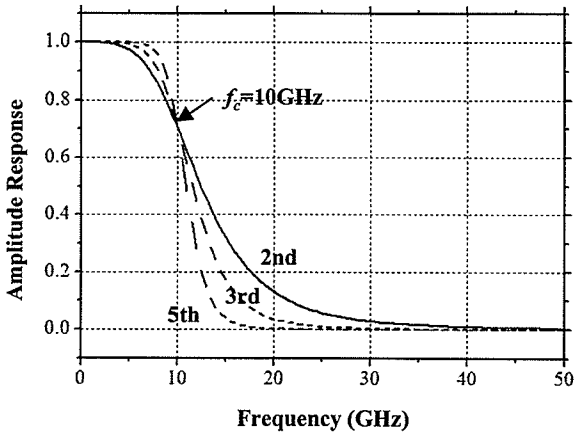

Fig. 4.17. The Butterworth filters of the 2nd, 3rd, and 5th order with a cutoff frequency of 10 GHz.

One example of the EFM sampled original and filtered wave forms is shown in Fig. 4.18. The sampling pulse was the SDL amplifier output 57 ps pulse with a sampling rate at 250 Gsample/s. The filter used was a 2nd order Butterworth filter with a cutoff frequency of 20 GHz. The *SNR* calculated from Eq. (4.21) using the sampled wave form (the left) and the noise (measured when the circuit signal was off) was approximately 22. The *SNR* for the filtered wave form (the right) was approximately 57, which is about a three times increasing in the *SNR* for this case. The measured wave forms shown in this thesis have all been noise filtered.

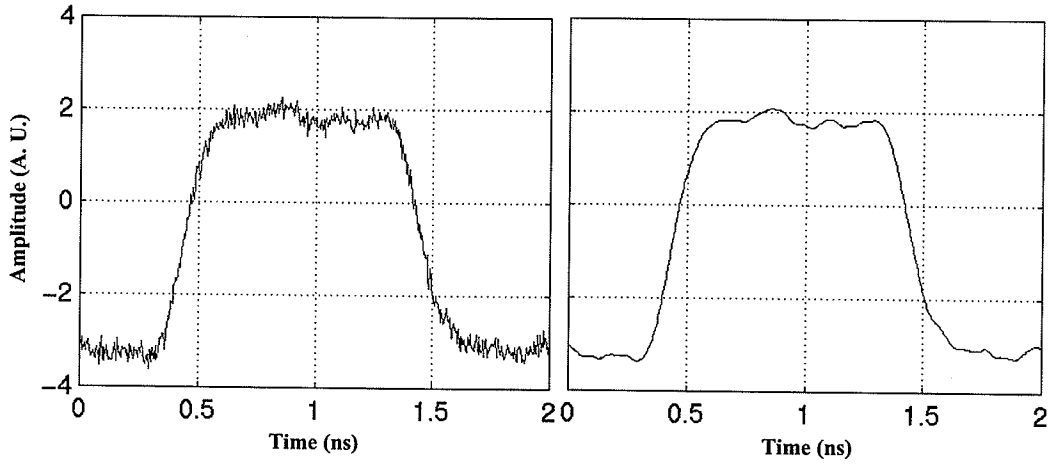

Fig. 4.18. EFM sampled wave form (the left) and its filtered result (the right).

4.5 The invasiveness

The EFM probe has two types of the invasiveness, the electrical loading due to the impedance disturbance and the pulse coupling due to the signal injection.

Due to the non-contact nature of the EFM probe, the parasitic loading of the EFM probe on the circuit is dominantly capacitive, determined by the coupling capacitance between the probe and the circuit under test (C_{pc}) as seen in the probe's equivalent circuit (Fig. 3.20). The probe lumped elements C_p, R_p and L_p can be neglected.

The coupling capacitance C_{pc}, which is dependent on the probe/circuit geometry, is simulated and discussed in Chapter 6. As a sample, when the silicon cantilever is located above three 2 μm wide interconnects with 1 μm separation, the total coupling capacitance is simulated to be < 0.2 *f*F at the tip-to-sample distance of 100 nm as shown in Fig. 6.18. This is small compared to the capacitive loading of the commercial contact probes, which are typically 0.01-0.1 pF [4].

The typical probe pulse coupling was measured by placing the probe at a typical measuring distance above a transmission line as shown in Fig. 4.19. In the measurement, the probe was positioned approximately 0.5 μm above the transmission line and the probe signal was the HP80000 1 ns pulse (as shown in Fig. 3.4). The circuit input signal was a 250 Mbit/s square-wave signal with an amplitude of 100 mV_{pp}. The disturbed circuit output signal after being amplified with a gain of 24 dB, $v_o(t)$, was measured by a 20 GHz digital sampling oscilloscope as shown in Fig. 4.20. Two small peaks in the $v_o(t)$ were the pulse coupling from the probe.

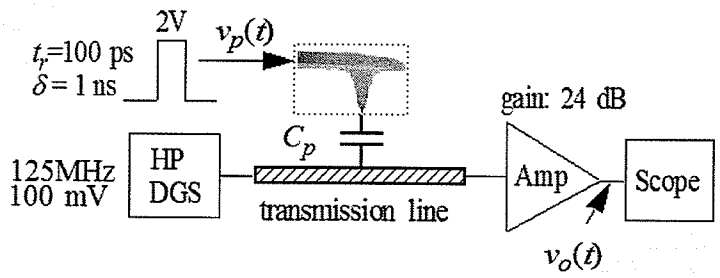

Fig. 4.19. The EFM probe pulse coupling measurement setup.

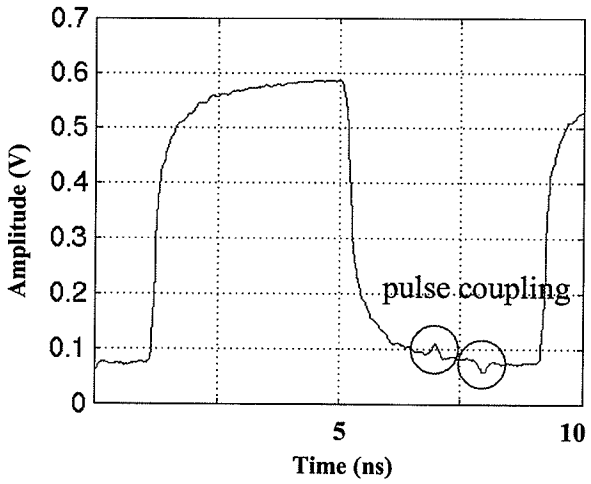

Fig. 4.20. The disturbed circuit output wave form (the two small peaks came from the probe pulse coupling).

The coupling capacitance, C_{pc}, can also be estimated from this measurement. A simple RC circuit model as seen in Fig. 4.21 was used to estimate the C_{pc}. The resistance R represents the two parallel 50Ωs (=25Ω). The capacitive impedance is $Z_c = 1/j\omega C_{pc}$. The probe pulse coupling is $v_{pc}(t)=[R/(R+Z_c)]v_p(t) = [j\omega RC_{pc}/(1+j\omega RC_{pc})]v_p(t) \approx (\omega RC_{pc})v_p(t)$ since the coupling capacitance C_{pc} is usually very small.

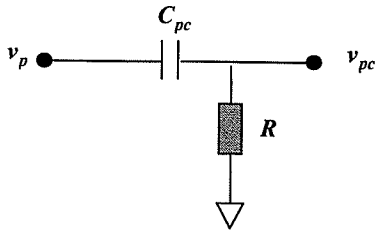

Fig. 4.21. The RC model for the probe pulse coupling measurement.

In this measurement, the pulse amplitude $V_p \sim 2$ V, the small peak $V_{pc} \sim 2$ mV (considered in the amplifier's gain factor) and $\omega \sim 2\pi/t_r$, where t_r is the probe pulse's transition time (~100 ps). The coupling capacitance C_{pc} is estimated to be ~ 1.5 fF. It is larger than the simulated 0.2 fF since the width of the transmission line was much larger than the 2 μm used for the simulation. Larger contributions came from the coupling capacitance between the cantilever beam and the transmission line. At 100 GHz, the parasitic shunt impedance $1/2\pi fC_{pc}$ is about 10 kΩ. This means that the capacitive coupling from the EFM probe is practically negligible when measuring 50 Ω systems.

4.6 Summary

The EFM measurement temporal resolution is limited by the system bandwidth which is mainly determined by the sampling pulse (the pulse generation system and the pulse propagation circuitry). The error due to the pulse nonidealities could be partially compensated by using a deconvolution process which requires the knowledge of the pulse response and the measurement. To improve the system bandwidth (10 GHz or above), a picosecond pulse generation system (such as an edge compressor) should be used and the pulse generation system has to locate as close as possible to the cantilever to avoid cable loss effects on the high frequency components in the pulse. A CPW structure cantilever will be used in future.

The voltage sensitivity of an EFM instrument is mainly determined by the thermal noise. To improve the measurement sensitivity, the probe should be placed as close as possible to the circuit (to increase $\partial C/\partial z$); a moderate vacuum system could be used to increase the cantilever vibration quality factor Q; and a multiple sampling pulse approach could be used to effectively increase the pulse duty time δ/T, which is very useful for long pattern signal measurement [75].

To reduce the probe capacitance loading (and also improve the spatial resolution), a longer tip with smaller tip radium should be used. Newly designed cantilevers with such a tip are commercially available [6]. Such a tip could be easily broken and a tradeoff between the loading and the voltage sensitivity exists.

The DEFM method assumes that the probe operates in the electrostatic regime (probe dimensions << wavelength). This is valid at mm-wave frequencies [76].

Chapter 5

Quantitative Voltage Measurements

As shown in the last two chapters, EFM pulse sampling technique provides only qualitative measurements, yielding the signal's relative wave shape and information such as rise and fall times, ringing, and timing. Up to now, quantitative voltage magnitude of high-speed arbitrary signals is still difficult to measure at internal nodes of an integrated circuit by using non-contact probing techniques. However, the internal signal magnitude information is very useful in trouble-shooting and important for many applications such as testing mixed signal high-speed ICs.

It is usually very complicated for the EFM probe to calibrate an accurate voltage scale since the force signal scale would alter and change according to any small changes in the tip-to-sample spacing, tip geometry and surface environment, e.g. cleanliness, etc.

In this chapter, an efficient null-force technique is employed in EFM to measure up to 250 Mbit/s signal voltages with the advantage of being free of those complicated calibration procedures. Measurement results using this technique are linear in a voltage range (-2V~ +2.5V). Measurements on a Nortel wideband distributed amplifier using the technique show a significant agreement to the results measured by a digital sampling oscilloscope and a vector network analyzer.

5.1 The null-force technique

As seen in Eq. (2.1), the electrostatic force depends on the ill-defined parameter *dC/dz*. Obtaining quantitative voltage information by detecting the force requires complex calibration or precise and bothersome probe positioning since *dC/dz* is determined by the physical geometry and the relative position of the cantilever and the circuit. A null-force technique was used to perform the quantitative voltage measurements. The nulling approach avoids the need for complicated calibration or for the knowledge of the ill-conditioned parameters.

The null-force technique can be realized by driving the probe with a modulated sampling pulse as

$$v_p(t) = (A + K\cos(\omega_r t))v_s(t), \tag{5.1}$$

where A has to be an adjustable variable, K could be an arbitrary non-zero constant. From the force equation Eq. (2.1), the force component at the cantilever's resonance frequency is

$$F_z(t)\big|_{\omega_r} = \frac{\partial C}{\partial z} K[A\langle v_s(t), v_s(t)\rangle - \langle v_s(t), v_c(x,y,t)\rangle + \Delta\Phi(x,y)\langle v_s(t)\rangle]\cos(\omega_r t), \tag{5.2}$$

and the force induced cantilever deflection is

$$\Delta z(t)\big|_{\omega_r} = \frac{\partial C}{\partial z}\frac{Q}{k} K[A\langle v_s(t), v_s(t)\rangle - \langle v_s(t), v_c(x,y,t)\rangle + \Delta\Phi(x,y)\langle v_s(t)\rangle]\sin(\omega_r t), \tag{5.3}$$

where the <> was defined in Eq. (2.28) and (2.29).

In Eq. (5.3), if the sampling signal is an ideal pulse, $v_s(t) = G_\delta(t-\tau)$, and assuming its bandwidth is much larger than the test point signal bandwidth ($\delta << t_r, t_f$) with $\delta/T << 1$, this probe deflection monitored by a lock-in amplifier can be approximated as

$$A_z(\tau)\big|_{\omega_r} = \frac{\partial C}{\partial z}\frac{Q}{k} K\frac{\delta}{T}[A - v_c(x,y,t=\tau) + \Delta\Phi(x,y)]. \tag{5.4}$$

The probe deflection will be nulled when $A = v_c(x,y,t=\tau) - \Delta\Phi(x,y)$. Here the dc offset, which is usually very small, will still be present in the measured signal voltage when the probe deflection is nulled. To avoid this, an ac-coupled sampling pulse,

$$v_s(t) = G_\delta(t-\tau) - \langle G_\delta(t)\rangle, \tag{5.5}$$

could be used, such that the probe deflection will be (neglected the $(\delta/T)^2$ term)

$$A_z(\tau)\big|_{\omega_r} = \frac{\partial C}{\partial z}\frac{Q}{k} K\frac{\delta}{T}[A - v_c(x,y,t=\tau) + \langle v_c(x,y,t)\rangle]. \tag{5.6}$$

The probe deflection will be nulled when $A = v_c(x,y,t=\tau) - \langle v_c(x,y,t)\rangle$ where $\langle v_c(x,y,t)\rangle$ is the dc component of the test point signal. Only the ac component of the test point signal is measured. This arrangement is preferred since the unknown (and slowly time varying) dc error $\Delta\Phi$ will not be presented in the measured signal voltage (since $\langle v_s(t)\rangle = 0$). The ac-coupled pulse arrangement is also useful for testing high-speed circuits which can be dc sensitive.

Experimentally, the sampling pulse Eq. (5.1) could be implemented by using a high-bandwidth switch and two adjustable-gain amplifiers with the amplitude V_{s1} and V_{s2} as seen in Fig. 5.1. If the switch is toggled at a frequency of $\omega_r=2\pi f_r$, with a 50% duty cycle, the resulting signal applied to the probe, as seen in Fig. 5.2, is (Fourier series)

$$v_p(t) = (A + K \sum_{n=1,odd}^{\infty} \frac{1}{n} \cos(n\omega_r t)) v_s(t), \tag{5.7}$$

where $A=(V_{s1} + V_{s2})/2$ and $K=2(V_{s1} - V_{s2})/\pi$, ignored the high order ($O(\omega_r^2)$) harmonics. This gives the equation (5.1).

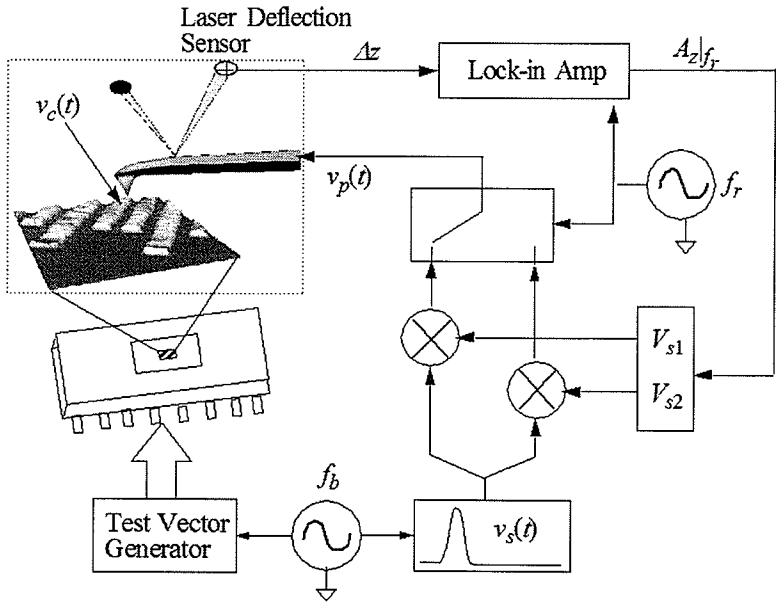

Fig. 5.1. Block diagram of the noncontact scanning probe microscopy based probe which employs a null-force technique to perform quantitative voltage measurements of the high-frequency internal IC nodes.

The sampling pulses used in this work were the output pulses of the HP80000 data generator which has a total of four output channels. The two pulses were synchronized and shifted in phase as seen in Fig. 5.2 to a wideband switch, which applied a square wave modulation at the cantilever resonance frequency, f_r, for the probe signal. Two identical wideband dc blockers could be used to block the pulse dc component for the ac-coupled measurements. The circuit under test was externally driven by the HP80000 data generator with a test pattern $v_c(t)$. $v_c(t)$ was configured to be an N-bit length vector with a period $T_o=N/f_b$, where f_b was the system clock rate (highest clock rate is 1 GHz). The

entire waveform $v_c(x,y,t)$ can be captured by shifting the pulse delay τ which is the same as that of the sampling process.

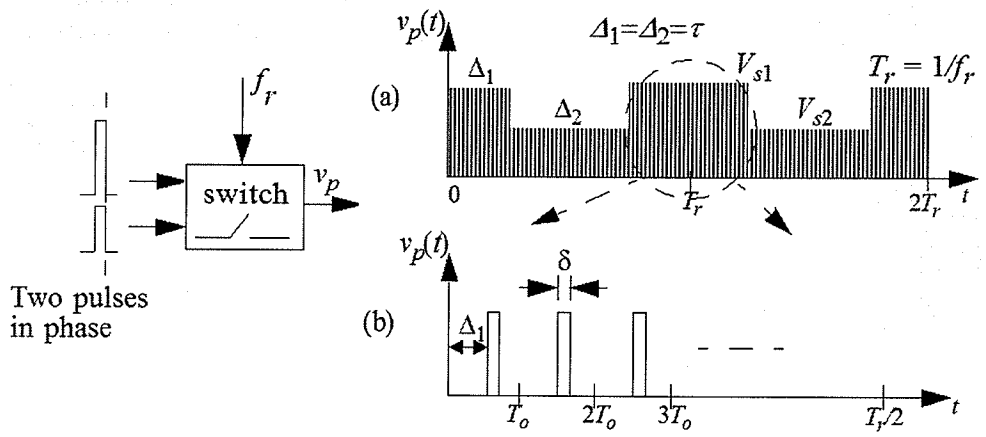

Fig. 5.2. The pulse amplitude modulation for the null-force technique.

Similar to the sampling process, the force nulling process is also highly automated by using the LabVIEW programs through GPIB cables. The nulling VI programs use the deflection signal, A_z (the output signal of the lock-in amplifier), as an error signal to automatically adjust the pulse parameter A (V_{s1}, V_{s2}) as shown in Eq. (5.7) until the probe deflection signal is zeroed (or within a pre-selected tolerance). The feedback control system is demonstrated in Fig. 5.3. Block D represents all the stages that process the pulse parameter to the deflection signal, which includes the probe, the deflection detector and the lock-in amplifier etc. Block C represents the stages that process the error signal (E) and drive the probe.

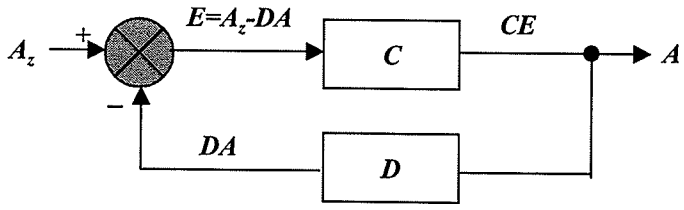

Fig. 5.3. The feedback control system for the automated nulling process.

The cantilever deflection, A_z, as a function of the applied pulse parameter A can be measured experimentally. Figure 5.4 shows a sample of the measurement. The curve demonstrates the linearity of the system and its slope depends on the phase setting and

the sensitivity setting on the lock-in amplifier, etc. The gain of the feedback control loop is dependent on the slope.

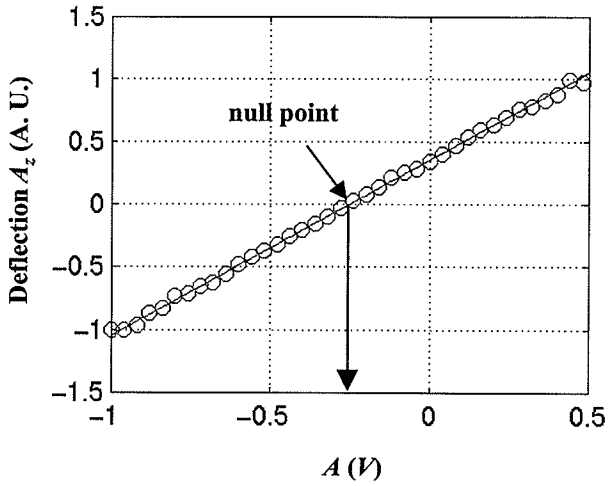

Fig. 5.4. The measured deflection signal, A_z, as a function of the applied pulse parameter A.

The delay for each iteration of the control system must to be larger than the time constant of the lock-in amplifier. This is to ensure that the lock-in amplifier has reached a steady state for the next error computation. The number of iterations for the nulling process can be adjusted by the program and these processes usually take less than five minutes.

With the described null-force technique, accurate signal voltages can be extracted without knowledge of dC/dz, Q, or k, thus eliminating the need for complex calibration and accurate probe positioning above the test point and since the approach is independent of C, as discussed in following chapter. As a result, measurement of passivated circuits can be performed. This eliminates the need for milling access holes to the desired test points of the circuit, which can be costly, time consuming and possibly destructive.

5.2 The null-force measurements

One of the challenges in this work is being able to perform quantitative voltage measurements in conjunction with very high bandwidth measurements (using a sampling pulse with a small δ). Previous results were only for 1 Mb/s [77] and 2 Mb/s [78] data rates. In this work, HP data generator's 1 ns sampling pulses are used to measure circuits operating up to 250 Mb/s [79], which is closer to the performance of current IC

technology. Two different nulling approaches have been used here: dc-coupled nulling (using a dc-coupled pulse) and ac-coupled nulling (using an ac-coupled pulse).

5.2.1 The dc-coupled nulling measurements

From the Eq. (5.4), using a dc-coupled pulse, the measured circuit voltage $v_c(x,y,t=\tau) = A+\Delta\Phi(x,y)$ and $\Delta\Phi$ could be measured by grounding the circuit.

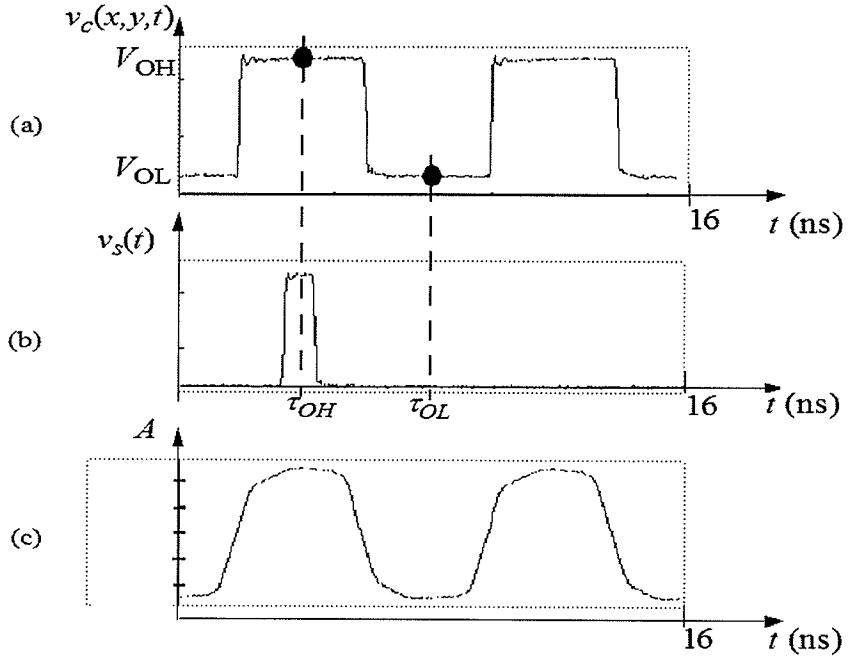

Fig. 5.5. (a) An example of a 250 Mbit/s data pattern to be measured, (b) a typical 1 ns wide sampling pulse $v_s(t)$ applied to the probe, and (c) the resulting wave form measured.

Experimentally, a 250 Mb/s square-wave data pattern as shown in Fig. 5.5(a), was applied to the matched 50 Ω microstrip transmission line. The voltage high, V_{OH}, and voltage low, V_{OL}, levels of the data signal were varied in order to mimic different technologies and logic families. Fig. 5.5(b) is a typical 1 ns wide sampling pulse, which is generated by the HP80000 data generator, the same as discussed in Chapter 3. Fig. 5.5(c) shows a sample of the corresponding wave form measured by the sampling pulse technique. The wave form may include some small dc offset $\Delta\Phi$. The dc-coupled nulling measurements were performed on the microstrip line as shown in Fig. 4.3 with the cantilever fully overhanging the line to avoid coupling from the ground plane.

Figure 5.6 shows the measured nulling value, A, as a function of the applied circuit signal voltage V_c. The points represent experimental data and the curve represents an

ideal mapping. The results show the linear relationship between the nulled values, A and the applied circuit signal voltages, V_c, along with some dc offset which is less than 50 mV ($\Delta\Phi$>-50 mV) as seen in the Fig. 5.6. The RMS value of the experimental points is less than 15 mV. The voltage accuracy is better than 30 mV in the dc measurement. The dynamic dc measurement range is –1.5 V to 1.5 V.

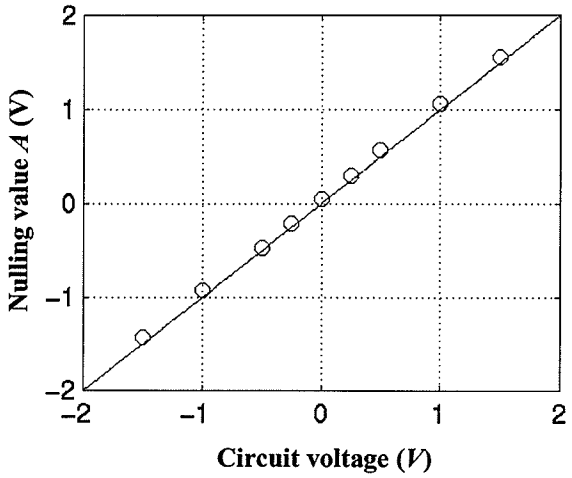

Fig. 5.6. The measured nulling value, A, as a function of the applied circuit signal voltage V_c.

Nulling measurements have been performed on the microstrip line with the probe located at different tip-to-sample distances of, 100 nm, 350 nm, 1 μm, and 2 μm. As shown in Fig. 5.7, the cantilever deflection was nulled at the same nulling value within the measurement error. This demonstrates that the probe is capable of direct measurement on a circuit without the passivation removed.

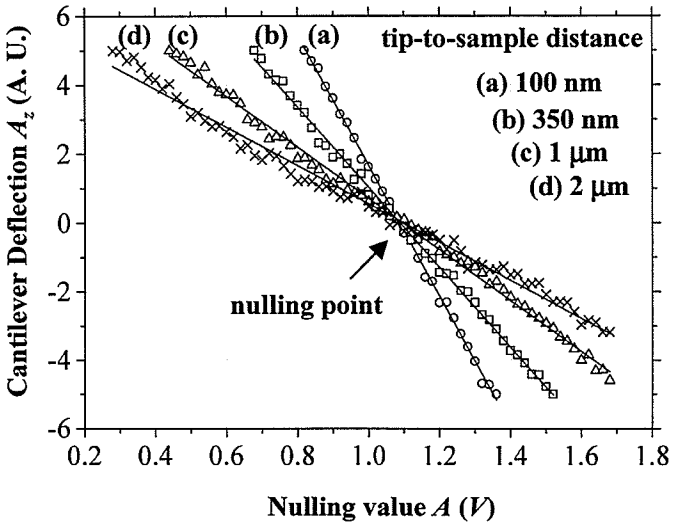

Fig. 5.7. The measured cantilever deflection, A_z, as a function of the applied pulse parameter A at different tip-to-circuit distances, i.e. 100 nm, 350 nm, 1 μm, and 2 μm.

5.2.2 ac-coupled nulling measurements

From the Eq. (5.6), using an ac-coupled pulse, the probe deflection is nulled at $A = v_c(x,y,t=\tau)-\langle v_c(x,y,t)\rangle$. Experimentally, ac-coupled nulling measurements were applied for the same 250 Mb/s square-wave circuit signal as shown in Fig. 5.8(a). Figure 5.8(b) is the ac-coupled sampling pulse. Figure 5.8(c) is the measured waveforms, which is also ac-coupled without dc error $\Delta\Phi$.

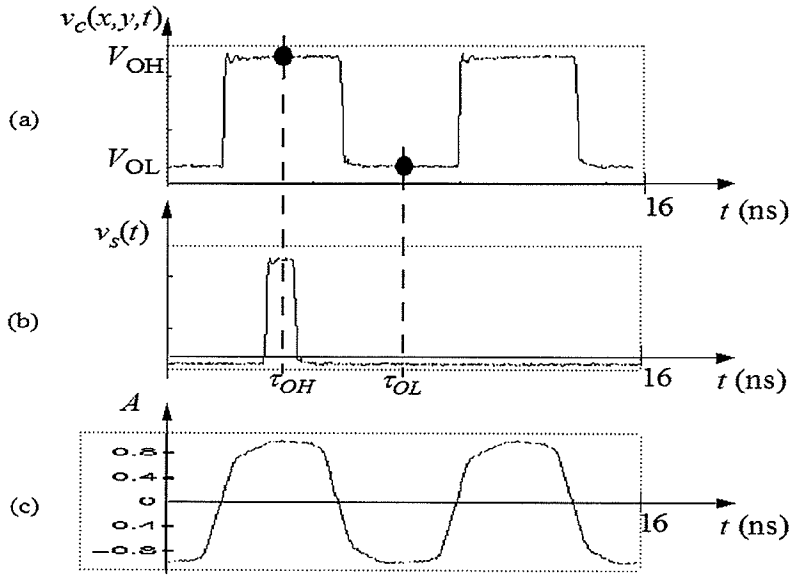

Fig. 5.8 Similar waveforms as in Fig. 5.5 except that here the pulse (b) is ac-coupled and the measured waveform (c) has no dc component.

The measured voltage swing, (A_H-A_L), as a function of the applied voltage swing, (V_{OH}-V_{OL}), is shown in Fig. 5.9. $A_H = V_{OH} - V_{offset}$; $A_L = V_{OL} - V_{offset}$ and $V_{offset} = \langle v_c(x,y,t) \rangle$. These results are obtained by performing the nulling procedure with the sampling pulse placed first at the middle of a logic-high state and then at a logic-low state.

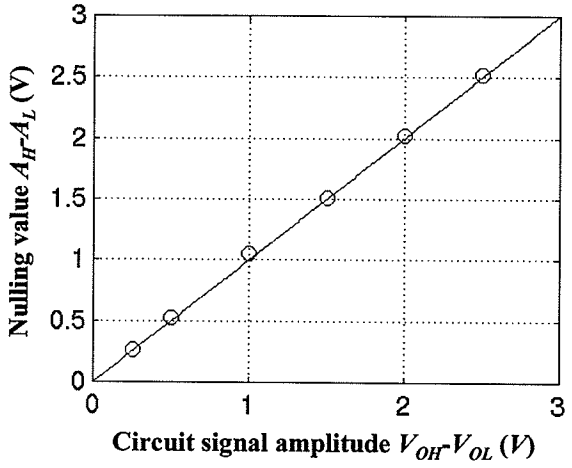

Fig. 5.9. Measured voltage swing, (A_H-A_L), as a function of the applied circuit voltage (V_{OH}-V_{OL}), obtained by performing the nulling procedure with the sampling pulse positioned at the middle of a logic-high state and then a logic-low state.

Also the ac voltage measurement accuracy is better than 30 mV and the dynamic range of the ac measurements is 2.5 V. The voltage sensitivity of the probe, measured on

the 50 Ω transmission line with a tip-to-circuit distance of ~ 1 μm was approximately 5 mV/√Hz.

To demonstrate the capabilities of the probe, a GaAs wideband amplifier IC used in an OC-192 fiber communications system was measured. The IC, mounted in its fixture, is shown in Fig. 5.10. The amplifier has a gain of 23 dB for up to 5 GHz. The input/output match is quite good (S_{11}, S_{22} < -10 dB) at this frequency range.

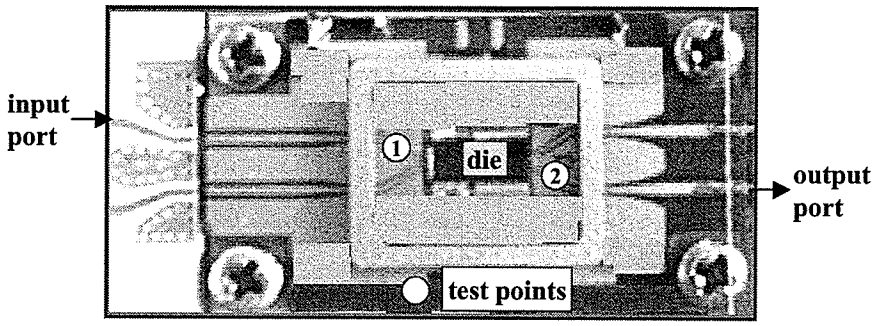

Fig. 5.10. A fixtured GaAs wideband distributed amplifier IC indicating two internal test points, located just off the die, on the circuit's input and output transmission lines.

A 250 Mb/s clock signal, with V_{OL}=-50 mV and V_{OH}=50 mV, was applied to the input of the amplifier. The signal on the input transmission line, just off the die, as shown in Fig. 5.10 (test point 1), was measured with the probe. The measured signal has an amplitude of 105 mV. The waveform shown in Fig. 5.11 is compared with the measurement made at the input connector of the fixture using a 20 GHz sampling oscilloscope. The probe measured voltage signal on the output transmission line, as shown in Fig. 5.10 (test point 2), has an amplitude of 1.6V. The resulting waveform is shown in Fig. 5.12 and compared with the measurement made at the output connector of the fixture using the 20 GHz sampling oscilloscope. These results of the probe are in good agreement except for the risetime, which is bandwidth-limited due to the 1 ns sampling pulse width.

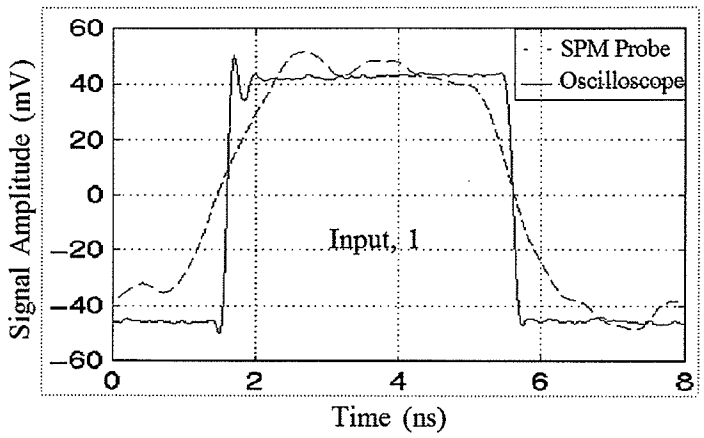

Fig. 5.11. The measured wave form at test point 1 in Fig. 5.10 compared with the measurement made at the input connector of the fixture using a 20 GHz sampling oscilloscope.

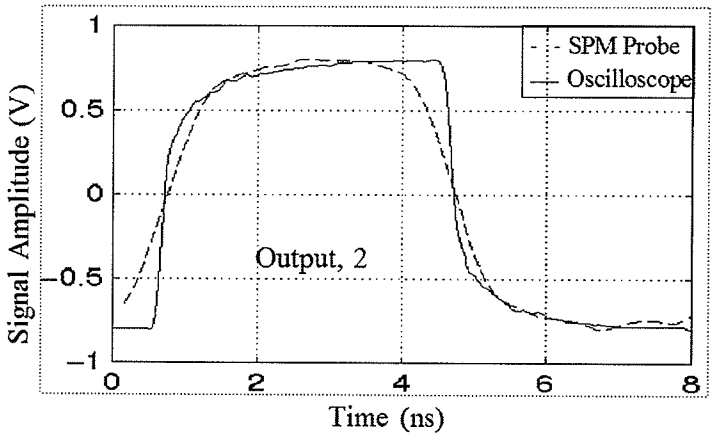

Fig. 5.12. The measured wave form at test point 2 in Fig. 5.10 compared with the measurement made at the output connector of the fixture using a 20 GHz sampling oscilloscope.

An optical photograph of the amplifier IC is shown in Fig. 5.13 and consists of two cascaded distributed amplifier stages. Internal measurements were performed at the input and output transmission lines (approximately 10 μm and 30 μm wide, respectively) of the two amplifier stages. Waveforms obtained at the I/O ports of each stage are shown in Fig. 5.14. From these results, the gain (signal reflections are not considered) of each stage was measured to be 13 dB. The total gain comes out to be 26 dB, which compares well with the 24 dB obtained using a 20 GHz sampling oscilloscope and the 23 dB gain using a network analyzer. The amplifier IC was passivated for all of the measurements with the cantilever tip typically located 0.3 mm above the circuit test points.

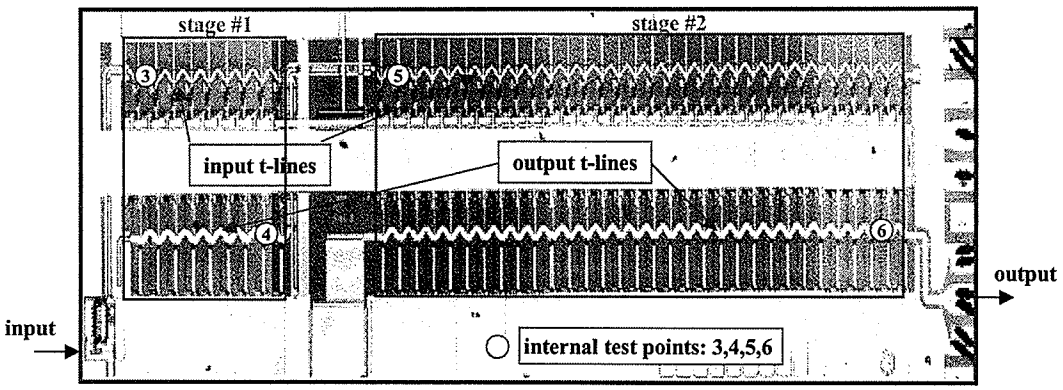

Fig. 5.13. A microphotograph of the two-stage distributed amplifier IC die in Fig. 5.10, indicating internal test points at locations 3, 4, 5, and 6 of the input and output transmission lines of each stage.

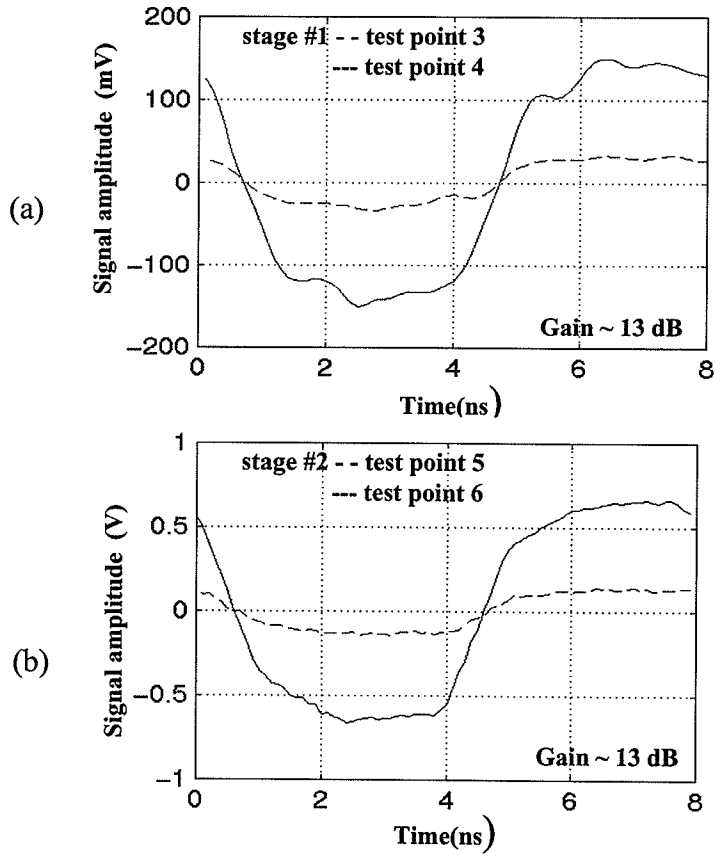

Fig. 5.14. A comparison of (a) the measured wave forms at the stage 1 input and output transmission line (test points 3 and 4) and (b) the measured wave forms at the second stage input/output (test points 5 and 6).

5.2.3 The error due to the pulse nonidealities

An ideal sampling pulse should have a rectangle shape with a small duration δ. The nonideal sampling pulse shape will effect the <> terms calculated in the Eq. (5.3). However, since $v_s(t)$ can easily be measured externally, the terms $\langle v_s(t), v_s(t)\rangle$ and $\langle v_s(t)\rangle$ can be predetermined independently of the signals being measured on the test circuit. For reasonably flat digital signals, the inner product of $v_c(t)$ and $v_s(t)$, can be evaluated as $\langle v_c(t), v_s(t)\rangle \sim \langle v_s(t)\rangle v_c(t=\tau)$. Thus, when the parameter A is adjusted to null the probe deflection, the unknown signal voltage levels can be determined as [77]

$$v_c(x,y,t=\tau) = \frac{\langle G_\delta(t-\tau), G_\delta(t-\tau)\rangle}{\langle G_\delta(t-\tau)\rangle} A + \Delta\Phi, \tag{5.9}$$

for $v_s(t)=G_\delta(t-\tau)$ and

$$v_c(x,y,t=\tau) - \langle v_c(x,y,t)\rangle = \frac{\langle G_\delta(t-\tau), G_\delta(t-\tau)\rangle}{\langle G_\delta(t-\tau)\rangle} A, \tag{5.10}$$

for $v_s(t)=G_\delta(t-\tau)-\langle G_\delta(t-\tau)\rangle$.

From the 1 ns sampling pulse in Fig. 3.4, $\langle G_\delta(t-\tau), G_\delta(t-\tau)\rangle = 0.0682$, and $\langle G_\delta(t-\tau)\rangle = 0.0910$. The scale factor is about 1.33. This scale factor can also be calculated from the results measured on the microstrip transmission line, assuming the test point voltage was the same as the input signal voltage.

The measurement results presented in previous sections are all corrected by this scale factor.

5.3 Pulse duration modulation approach

The modulated sampling pulse in the Eq. (5.1) for the null-force technique could also be implemented by using the pulse duration modulation approach (instead of the pulse amplitude modulation approach as used in the previous sections).

The pulse duration modulation approach is illustrated in Fig. 5.15 in comparison with the pulse amplitude modulation approach. Two sampling pulses (one is the exact opposite of the other) are switched on and off with the duty cycle D and 1-D, respectively, at a rate

of ω_r. The duty cycle was set to D = 0.5 (for a square wave modulation) in the pulse amplitude modulation approach, which is discussed in this chapter.

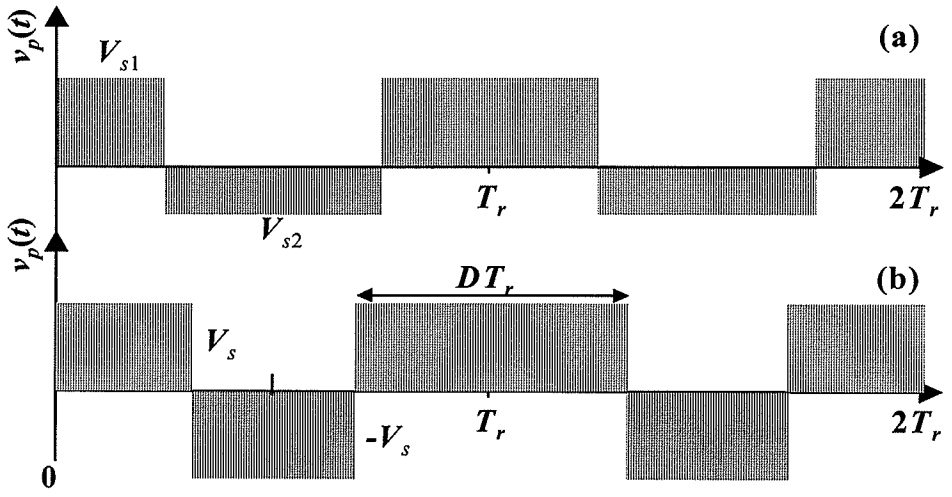

Fig. 5.15. The pulse (a) amplitude modulation and (b) duration modulation, where *D* is the duty cycle parameter.

Ignoring the high order $O(\omega_r^2)$ harmonic teams in its Fourier series, the modulated pulse $v_p(t)$ can be written as

$$v_p(t) = (A + K\cos(\omega_r t))v_s(t), \tag{5.8}$$

where $A=(2D-1)V_s$ and $K=(4/\pi)V_s\sin(D\pi)$. The force nulling could be approached by adjusting the parameter D (0→1) (while keeping the ω_r signal in constant phase) to null the probe deflection at $A=v_c(x,y,t=\tau)-\langle v_c(x,y,t)\rangle$ for an ac-coupled sampling pulse $v_s(t)=G_\delta(t-\tau)-\langle G_\delta(t-\tau)\rangle$.

The implementation of the pulse duration modulation could be easier since it only requires pulse inverting to have two exact opposite sampling pulses. There is no need to regulate (and invert) the pulse amplitudes as is required in the pulse amplitude modulation. The disadvantage for the pulse duration modulation is that the K value changes with the parameter D ($K=(4/\pi)V_s\sin(D\pi)$). K is only half of its maximum at D=1/6 or 5/6 and will decrease significantly as D getting near to 0 and 1. This means that the voltage measurement dynamic range will relate to the measurement sensitivity. Since $K \propto \sin(D\pi)$, this also means that the gain in the feedback control loop (Fig. 5.9) has to change sign while D crosses the point of 0.5.

5.4 Summary

In this chapter, an efficient null-force technique is employed in EFM to measure up to 250 Mbit/s signal voltages with the advantage of being free of those complicated calibration procedures. Two nulling approaches have been introduced and discussed. The performance of the nulling technique is demonstrated by the measurements on a Nortel wideband distributed amplifier.

Chapter 6

Electrostatic Force Microscopy Spatial Resolution

In the previous chapters, dynamic electrostatic force microscopy is used for high-speed voltage signal measurements. It is a non-contact probing technique which utilizes the nonlinear Coulomb force interaction between the probe tip and the circuit test point to extract localized voltages. When a standard SPM probe is used and is in close proximity (< 100 nm) to the circuit the EFM should have a theoretical spatial resolution in the order of the tip-to-sample distance [36,80]. This is often not achievable in the practical probing of an operating IC however, due to the large tip-to-sample distance required when the probing is operated on a probe station or when a circuit passivation layer is present. This results in strong coupling of interfering signals to the probe tip sidewall and to the cantilever [81,82,83]. Theoretical and experimental studies [84,85] show that this can lead to inaccurate measurements and a dramatically reduced spatial resolution.

In this chapter, the spatial resolution of a DEFM instrument is studied.

6.1 Spatial resolution problems

6.1.1 Coupling capacitances

The spatial resolution of an EFM instrument generally refers to the locality of the voltage measurements. Since current EFMs use a direct force detection approach, the operative parameter for assessing the spatial resolution is the coupling capacitance gradient as given by the force equation (ignoring the dc offset $\Delta\Phi$ for conciseness)

$$F_z(z,t) = \frac{1}{2}\frac{\partial}{\partial z}C(x,y,z)[v_p(t) - v_c(x,y,t)]^2. \qquad (6.1)$$

Note that the capacitance in this equation is from all sources on the circuit surface to the entire cantilever. At a large tip-to-sample distance (z > the radius of the tip curvature), significant coupling to the probe can result from points other than from the test point under the tip apex [81-85]. The resulting spatial resolution depends, not only on the cantilever geometry, but also on the circuit geometry. This chapter investigates the spatial resolution problems in internal circuit signal probing.

In internal circuit signal probing, the probe spatial resolution problem can lead to incorrect absolute voltage and waveform measurements due to the interference from sources other than the test point. As an example, internal probing of three parallel conducting traces with a standard SPM probe is investigated as shown in Fig. 6.1. If the three conducting lines are energized with the potential $v_{ci}(x,y,t)$; i=0,1,2, and the entire cantilever is equipotential, then the resulting total electrostatic force on the cantilever can be expressed as

$$F_z(z,t) = \frac{1}{2}\sum_{i=0}^{2}\frac{\partial C_i(x,y,z)}{\partial z}[v_p(t) - v_{ci}(x,y,t)]^2. \tag{6.2}$$

where $C_i(x,y,z)$; i=0,1,2, are the coupling capacitances between the entire cantilever and each of the three traces, respectively. $F_z = \sum F_{zi}$ and F_{zi} is a complicated average of electrostatic forces induced by all points on the trace i. Assuming the averaged forces are dominantly acting at the end of the cantilever, the resulting cantilever deflection at the cantilever resonance frequency, ω_r, (using sampling pulse $v_p(t)$) can be simply expressed as

$$A_z(z,\tau)\big|_{\omega_r} = \sum_{i=0}^{2} W_i v_{ci}(x,y,t=\tau), \tag{6.3}$$

which can also be written in two parts

$$A_z(z,\tau)\big|_{\omega_r} = W_1 v_{c1}(x,y,t=\tau) + \sum_{i=0,2} W_i v_{ci}(x,y,t=\tau), \tag{6.4}$$

where the first part is the desired signal and the second part represents the interfering signals from the adjacent traces. W_i; i=0,1,2, is a weight factor which is a complex function of the cantilever spring constant, the quality factor, especially the capacitance derivative and the force average effect. According to the Eq. (2.3), the cantilever is much

less susceptible to the forces acting on the points further away from the end of the cantilever.

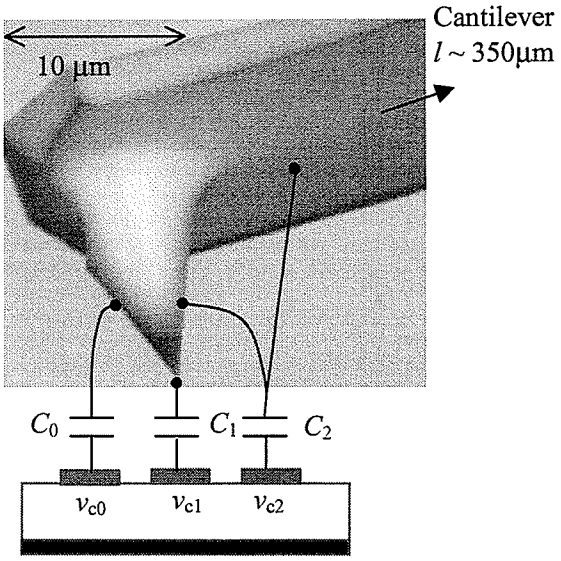

Fig. 6.1. The coupling capacitances between three parallel interconnects and a standard silicon SPM probe.

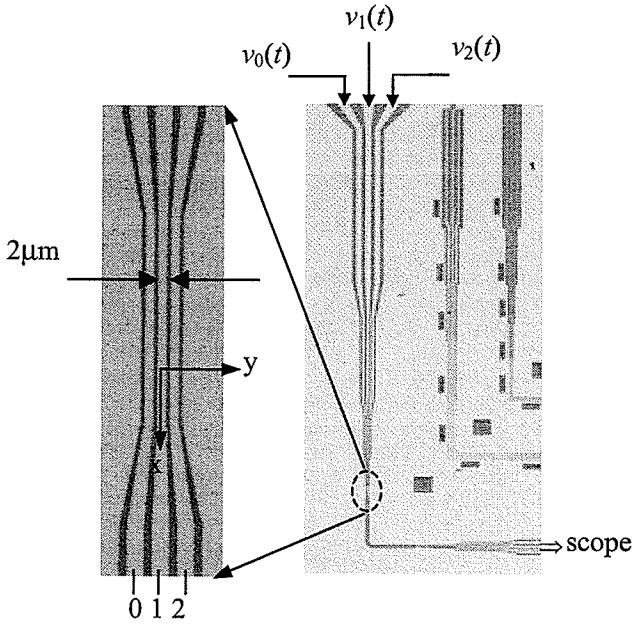

Fig. 6.2. A high-speed test circuit showing five sections of three conducting traces (light in color), which are surrounded by a large ground plane, and the applied signals. The smallest traces (zoom-in picture) are 2 μm wide with a 1 μm separation.

Experimental investigation of the spatial resolution of the force technique was performed using the high-speed test circuit shown in Fig. 6.2, as called ITG [86]. The

circuit consists of five sections of three parallel conducting traces of the same width, w, and a separation $w/2$ with w =2, 4, 8, 16, and 32 μm. The traces are surrounded by a large ground plane. The circuit was carefully mounted and leveled on the probe station (slope < −0.01 in y-axis).

6.1.2 Spatial resolution effects on the absolute voltage measurements

Using the null-force technique discussed in Chapter 5, the measured cantilever deflection can be expressed as

$$A_z(z,\tau)\big|_{\omega_r} = W_1(A - v_{c1}(x,y,t=\tau)) + \sum_{i=0,2} W_i(A - v_{ci}(x,y,t=\tau)). \tag{6.5}$$

The nulling value A is

$$A = \frac{\sum_{i=0}^{2} W_i v_{ci}(x,y,t=\tau)}{\sum_{i=0}^{2} W_i}. \tag{6.6}$$

If $v_{ci}(x,y,\tau)$=0; i=0,2, then the nulling value is

$$A = \frac{W_1 v_{c1}(x,y,t=\tau)}{\sum_{i=0}^{2} W_i}. \tag{6.7}$$

From Eq. (6.7), the measured voltage is smaller then the circuit voltage due to the capacitive coupling between the cantilever and the adjacent traces (W_0 and W_2) even if they are grounded. The error for the voltage measurement could be significant for the case when the measured signal trace is surrounded by a larger ground plane.

Experimental investigation was performed on the trace 1 of the ITG test circuit with traces 0 and 2 grounded (see Fig. 6.3 for trace number). In the measurement, the circuit signal was supplied by the HP80000 data generator. The voltage on the circuit was approximately -0.75 V. The sampling pulse was a 100 ns pulse at a clock rate of 10 MHz. During the measurement, the probe tip was usually positioned at the center of the trace 1 as seen in the Figure 6.3 and the cantilever was extended to the right with a tilted angle 15° to the circuit surface.

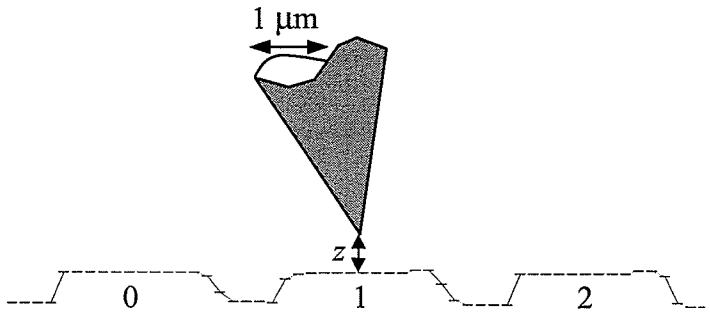

Fig. 6.3. The probe tip located at the center of the trace 1 and the cantilever extended to the right.

The measured nulling values at the different tip-to-sample distances on the trace 1 (with a different width w=2 and 16 μm) are shown in Fig. 6.4 (a) and (b). The results are compared with the nulling values measured on a 300 μm wide microstrip line with the probe cantilever half loaded on the transmission line (Fig. 6.4(c)).

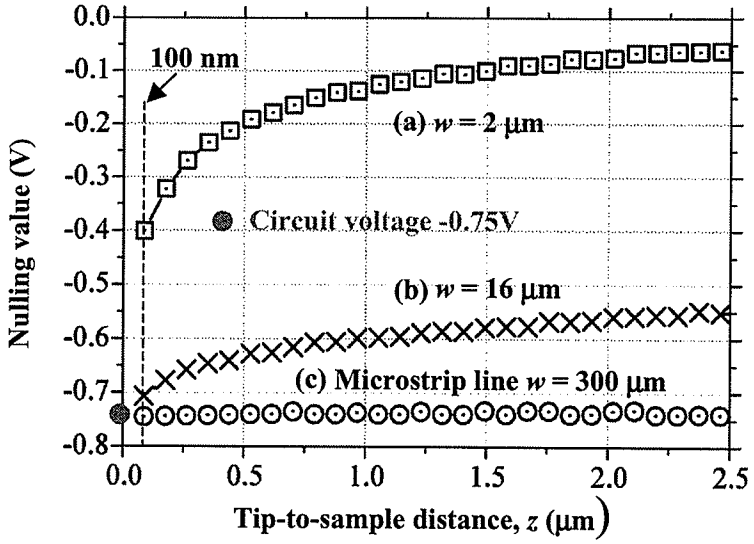

Fig. 6.4. The nulling value in a function of the tip-to-sample distance, z. Curves (a) and (b) are the measurement results on the ITG circuit trace 1 with w=2 and 16 μm respectively. Curve (c) is the measurement results on a microstrip line with a width of 300 μm.

From Fig. 6.4(c), the nulling measurements are accurate (the measurement error ~ 1%) when measuring a 300 μm wide microstrip line, where the probe capacitance coupling to the ground plane is very small. There exist significant errors when measuring the center trace of three parallel interconnects. The measured voltage had about a 5% error for the 16-μm wide trace 1 (Fig. 6.4(b)) and about a 45% error for the 2-μm wide trace 1 (Fig. 6.4(a)) at the probe tip-to-sample distance ~ 100 nm. The measured error increased to

about 20% for the 16-μm wide trace 1 and > 80% for the 2-μm wide trace 1 at the tip-to-sample distance 1 μm.

The relative sizes of the weight factors, which determine the interference level, could be estimated from these measurements.

6.1.3 The spatial resolution effecting on wave form measurements

To demonstrate the interference signals in the wave form measurement, HP 80000 2.5 Mbit/s signals were applied to the circuit as shown in Fig. 6.5(a); $v_1(t)$ and $v_2(t)$ were both square waves with the same amplitude (-0.75V to +0.75V) and a phase shift of 90° between them. $v_0(t) = 0$. A 20 ns sampling pulse with amplitude 2V and $T = 800$ ns was applied to the probe. In Fig. 6.5(b) the probe tip was located 100 nm above trace 1 to sample $v_1(t)$, with $v_2(t)$ acting as an interfering signal. In Fig. 6.5(c) the probe tip was located 100 nm above trace 2 to sample $v_2(t)$, with $v_1(t)$ acting as an interfering signal. These wave forms were sampled by shifting the pulse delay, at 20 ns per step, along the circuit signals.

From Eq. (6.4), the sampled wave form is

$$A_z(z,\tau) = A_{z0}(z,\tau) + A_{z1}(z,\tau) + A_{z2}(z,\tau), \tag{6.8}$$

where $A_{zi}(z,\tau)=W_i v_{ci}(x,y,t=\tau)$, $i=0,1,2$. $A_{z0}(z,\tau)=0$ since the trace 0 was grounded.

In Fig. 6.5(b), the sampled wave form obviously includes the inference signal from the trace 2, i.e. $A_{z1}(z,\tau)+A_{z2}(z,\tau)$. The result indicates an interference level (A_{z2}/A_{z1}) of as much as 30%. In Fig. 6.5(c), the sampled waveform includes the inference signal from the trace 1. The interference level (A_{z1}/A_{z2}) was over 10%.

When the probe tip was located 100 nm above the trace 0, which was grounded, the sampled wave form was similar to the Fig. 6.5(b). In this case, both $v_1(t)$ and $v_2(t)$ acted as interfering signals to the measurement. The ratio of the interference levels for the two adjacent traces (A_{z2}/A_{z1}) was approximately 50%.

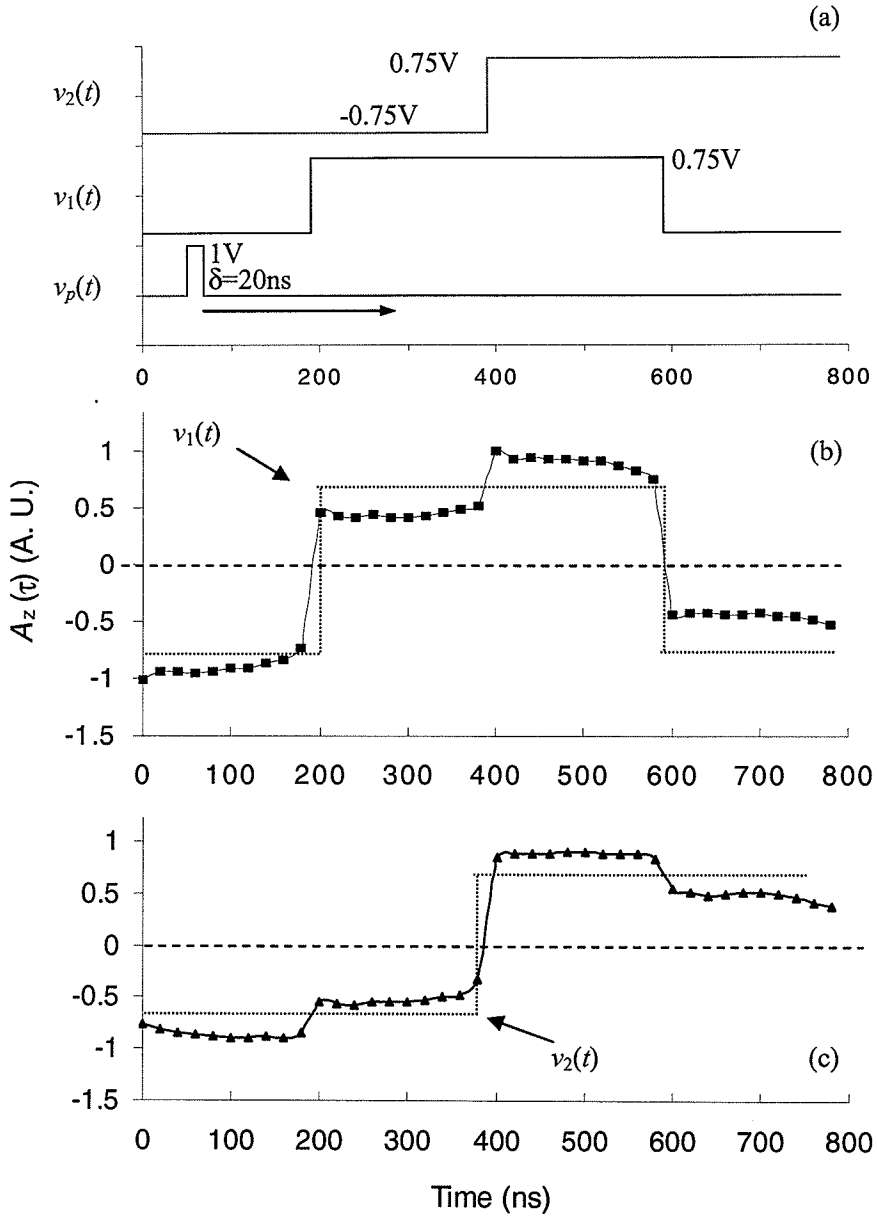

Fig. 6.5. (a) The square-wave 2.5 Mbit/s signals applied to traces 1 and 2, with $v_0(t)$=0. The probe signal is a 20 ns sampling pulse with T=800 ns. (b) The measured deflection $A_z(z,\tau)$ for the probe 100 nm above trace 1. (c) The measured deflection for the probe 100 nm above trace 2.

More measurements were done for the different tip-to-sample distances and the measured signal and the inference signal levels are listed in Table 2 and 3.

Table 2 lists the measured signal and the inference signal levels (normalized to the signal A_1 at 100 nm) when the probe was located above the trace 1 at three different tip-to-sample distances, i.e. 100 nm, 500 nm and 1 μm. Table 3 lists the measured signal and the inference signal levels (normalized to the signal A_0 at 100 nm) when the probe was

located above the trace 0. From these measurements, the interference levels from the adjacent traces increased a lot as the tip-to-sample distance increases from 100 nm to 1000 nm. When there is an adjacent trace, the probe should be positioned as close as possible to the signal trace under test.

Table 2. The measured signal and the inference signal levels (normalized to the signal A_1 at 100 nm) when the probe was located above the trace 1 at three different tip-to-sample distances.

z (nm)	A_0	A_1	A_2
100	0.12	1	0.3
500	<0.1	0.5	0.25
1000	<0.1	0.35	0.21

Table 3. The measured signal and the inference signal levels (normalized to the signal A_0 at 100 nm) when the probe was located above the trace 0 at three different tip-to-sample distances.

Z (nm)	A_0	A_1	A_2
100	1	0.35	0.15
500	0.5	0.25	0.13
1000	0.35	0.21	0.11

6.1.4 1D force scan

To further demonstrate the spatial resolution using direct force measurement, one-dimension (1D) force scans were performed across the three traces.

First, a topographic scan was acquired by the SPM probe. The probe was operated in the contact mode, the same as atomic force microscopy (AFM) [28]. In this work, the deflection signal A_z was used as a feedback control signal to perform "constant-deflection scanning" as seen in Fig. 6.6. The voltages (V_y, V_z), which were applied to the 3D piezoelectric positioner, produced the topographic information (1D) of the circuit surface.

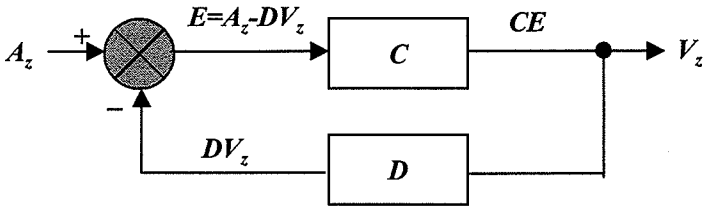

Fig. 6.6. The feedback control system used for a topographic scans.

The piezoelectric positioner used in this work has a movement range of 20 μm (for the 75 V maximum output voltage of the piezo controller) for all three axes. The piezo coordinates were calibrated by a standard Si/SiO_2 grating (TGZ03) from the SPMTIPS [6], which is a 1D arrays of rectangular SiO_2 steps on a Si wafer as seen in Fig. 6.7.

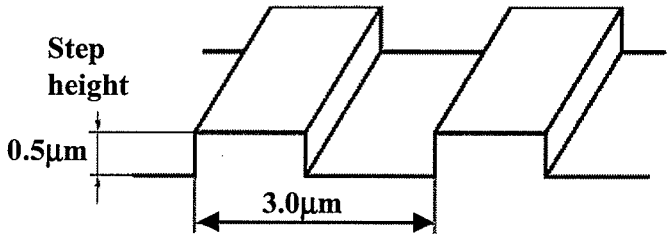

Fig. 6.7. The schematic of a standard Si/SiO_2 grating (TGZ03) with a step height of 0.5 μm ±1% and a pitch size of 3.0 μm.

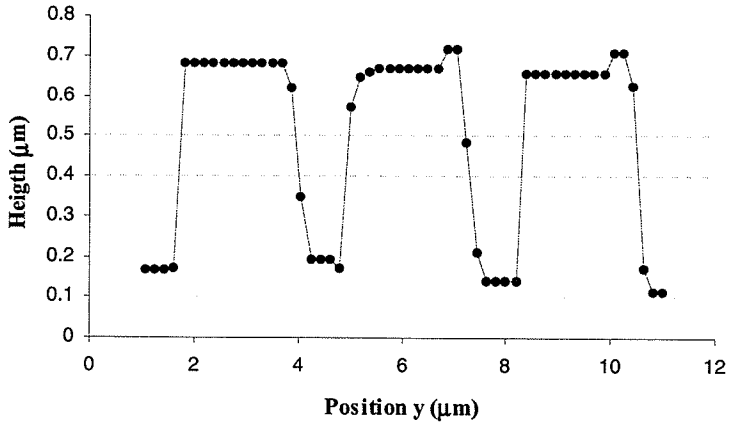

Fig. 6.8. A sample of a 1D topographic scan of the test circuit with the smallest line width of 2 μm and a separation of 1 μm. The thickness of the three traces is about 0.5 μm.

A sample of a 1D topographic scan of the ITG circuit's three traces of 2 μm in width is shown in Fig. 6.8. From this measurement, the thickness of the three metal traces was shown to be about 0.5 μm.

Next, the 1D force scan was performed in the noncontact mode. During the second scan as seen in Fig. 6.9, the probe was lifted and kept at a constant distance z above the interconnect traces (z is denoted as the distance between the tip apex and the top of the traces). Since the test circuit may be tilted at a small angle (could be adjusted to be as small as 0.1%), the piezo voltage V_z had to be corrected by this angle for the constant z. The two scans were both controlled automatically by VI programs.

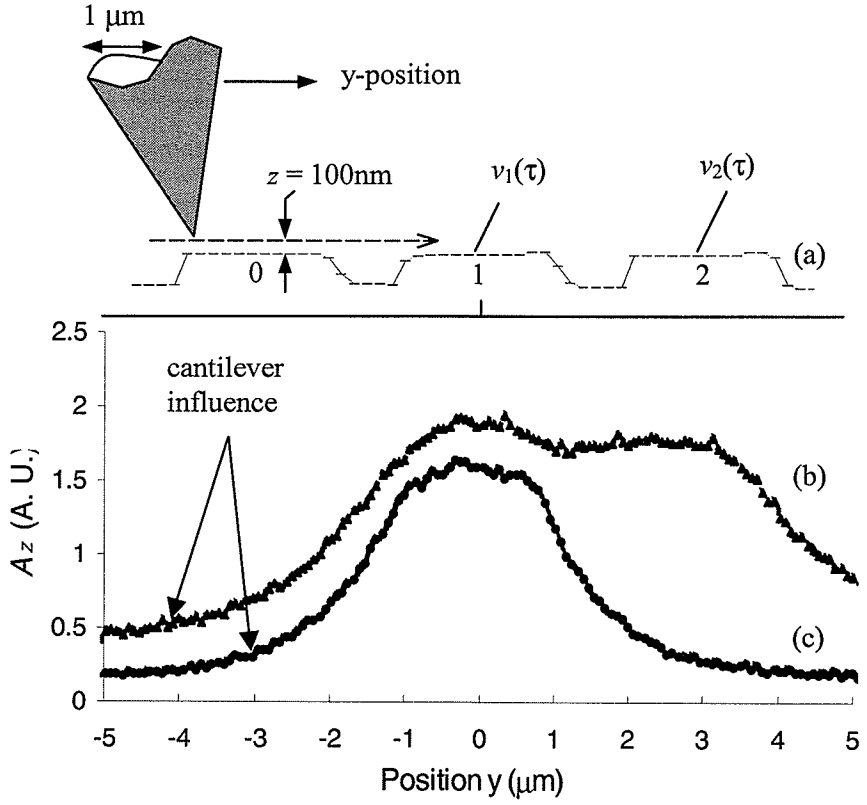

Fig. 6.9. Curve (a) is a 1D topographic scan of the test circuit's three traces with 2 μm width. Curve (b) and (c) show two 1D force scans, A_z, for z = 100 nm. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_0(\tau)$ = 0 and $v_1(\tau) = v_2(\tau)$ = 0.75 V in (b) and $v_0(\tau) = v_2(\tau)$= 0 V, $v_1(\tau)$=0.75 V in (c).

In Fig. 6.9, the curves (b) and (c) are two 1-D force scans, A_z, for z = 100 nm. For curve (b), a 2.5 Mbit/sec square-wave signal was applied to the three traces such that $v_0(\tau)$ = 0, and $v_1(\tau) = v_2(\tau)$ = 0.75 V. For curve (c), $v_0(\tau) = v_2(\tau)$ = 0 V, $v_1(\tau)$ = 0.75 V. Curve (b) shows that a large force was measured even when the tip was over the grounded trace (trace 0). According to Eq. (6.8), this undesirable background signal came from the signal A_{z1} and was mainly the force contribution of the cantilever (so called the cantilever inference). A larger background signal was measured for curve (c) when traces 1 and 2

were both energized. In this case, $A_z=A_{z1}+A_{z2}$, and both A_{z1} and A_{z2} were the interfering signals.

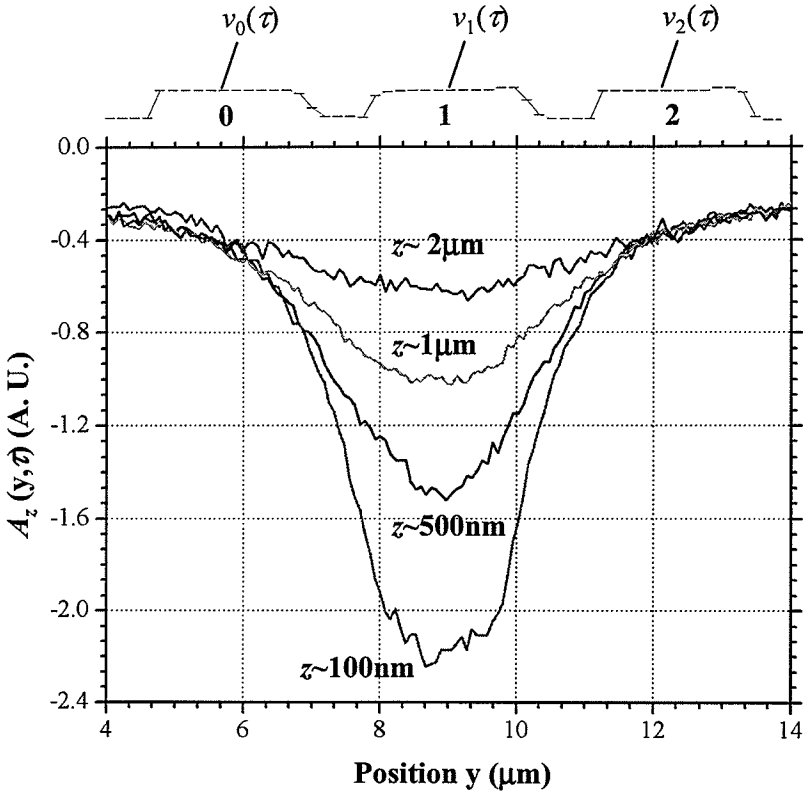

Fig. 6.10. Four 1D force scans at four different tip-to-sample distances, i.e. 100 nm, 500 nm, 1 μm, and 2 μm. The three traces were stimulated with $v_0(\tau) = v_2(\tau) = 0$ and $v_1(\tau) = -0.75$ V.

Figure 6.10 shows 1D force scans for four different tip-to-sample distances at 100 nm, 500 nm, 1 μm, and 2 μm. These curves show that the background signal was almost independent of the probe height. If this background signal was mainly the contribution of the cantilever, it should have been insensitive to the small change in the probe height, due to the much larger distance between the cantilever and the trace. The signal of interest (due to the $v_1(\tau)$) decreased much quicker as the tip-to-sample distance increased.

Figure 6.11 and 6.12 are 1D force scans with signals on trace 1 and 2 with the width of 2 μm at tip-to-sample distance z=100 nm. In Fig. 6.11, the circuit was energized with $v_0(\tau) = 0$ and $v_1(\tau) = v_2(\tau) = 0.75$ V in curve (a) and $v_0(\tau) = 0$ V, $v_1(\tau)$=-0.75 V and $v_2(\tau) =$ 0.75V in curve (b). In Fig. 6.12, the circuit was energized with $v_0(\tau)= 0$ and $v_1(\tau) = 0.75$V, $v_2(\tau) = -0.75$ V in curve (a) and $v_0(\tau) = 0$ V, $v_1(\tau)=v_2(\tau)= -0.75$ V in curve (b).

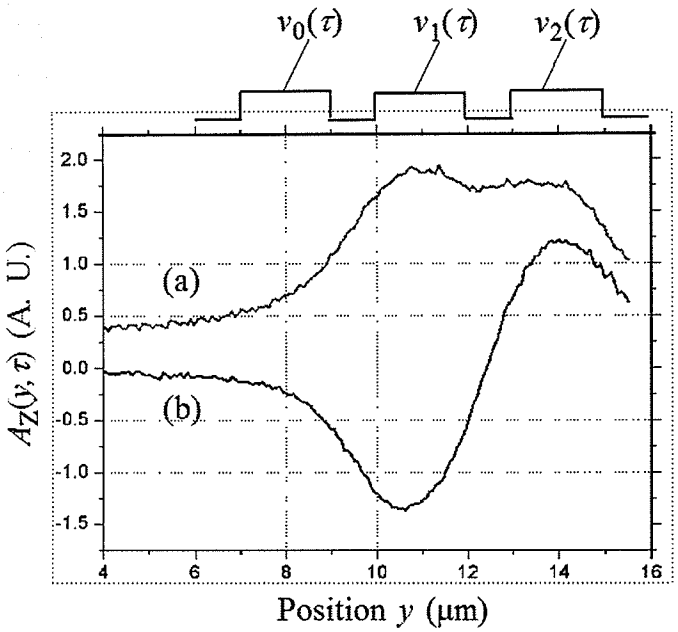

Fig. 6.11. The 1D force scan of the test circuit's three traces with a 2 μm width for z=100 nm. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_0(\tau)$=0 and $v_1(\tau) = v_2(\tau)$ =0.75 V in (a) and $v_0(\tau)$ = 0V, $v_1(\tau)$=-0.75 V and $v_2(\tau)$=0.75 V in (b).

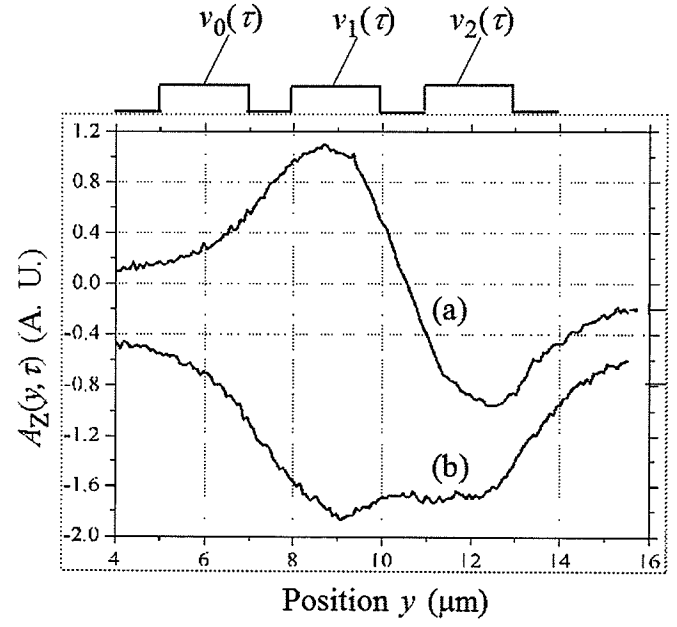

Fig. 6.12. The 1D force scan of the test circuit's three traces with a 2 μm width for z=100 nm. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_0(\tau)$=0, $v_1(\tau)$ = 0.75 V and $v_2(\tau)$ =-0.75 V in (a) and $v_0(\tau)$ = 0 V and $v_1(\tau)=v_2(\tau)$=-0.75 V in (b).

Figure 6.13 is a 1D force scan on trace 1 and 2 with the width of 4 μm and a separation of 2 μm at a tip-to-sample distance of 100 nm. The circuit was energised with $v_0(\tau)$=0 and $v_1(\tau)$=$v_2(\tau)$=-0.75 V.

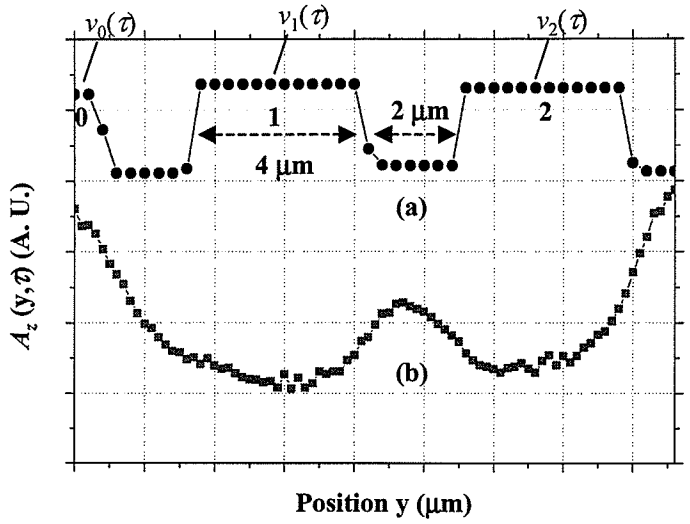

Fig. 6.13. The curve (a) is the topographic scan of the test circuit's three traces with a 4 μm width. The curve (b) is the 1D force scan for z = 100 nm. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_0(\tau)$ = 0, $v_1(\tau)$ = $v_2(\tau)$= -0.75V.

A 1D force scan was also performed on the two edges of the 32 μm-wide trace 1 (0.75 V) with the traces 0 and 2 grounded. Figure 6.14 shows the measured results on the falling edge and Fig. 6.15 on the rising edge at two different tip-to-sample spacings. The cantilever extended to the right of the trace for each measurement (as shown in Fig. 6.14 and 6.15, but not scaled). In Fig. 6.14, the cantilever deflection signals drop dramatically (due to the tip contribution) when the probe tip was started moving away from the falling edge but drop slowly (mainly came from the tip sidewall coupling) to zero when the probe tip was far away. However, in Fig. 6.15, non-zero small deflection signals were detected even when the probe tip was already far away from the rising edge. These non-zero small signals were obviously the contribution of the coupling capacitance between the cantilever and the signal trace since the cantilever (with a length of 350 μm) was hanging over the signal trace. At y=0, the cantilever-to-trace distance can be estimated by the tip height (10 μm) and the cantilever tilted angle 15° to be ~ 14 μm. The cantilever influence is insensitive to the tip-to-sample distance as seen in these results because the cantilever-to-sample distance is > 10 times the tip-to-sample distance in these cases.

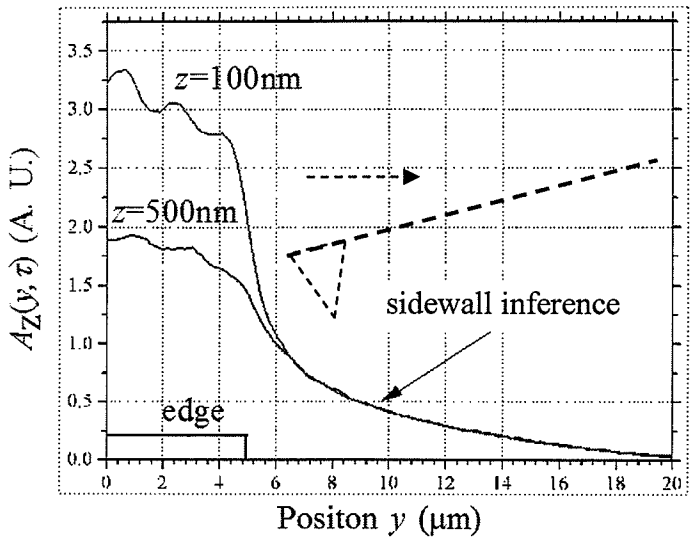

Fig. 6.14. The 1D force scan on the falling edge of the trace 1 for z=100 and 500 nm. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_1(\tau)$= 0.75 V, and $v_0(\tau) = v_2(\tau)$=0 V.

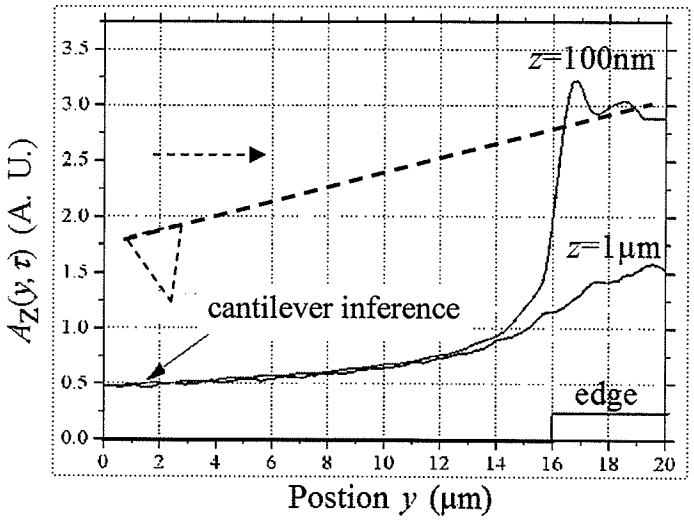

Fig. 6.15. The 1D force scan on the rising edge of trace 1 for z=100 nm and 1 μm. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_1(\tau)$= 0.75 V, and $v_0(\tau) = v_2(\tau)$=0 V.

6.1.5 Summary

From these measurements, the spatial resolution of the EFM probe is limited due to the contribution of the coupling capacitance between the tip sidewall and especially the cantilever and the circuit traces. The simple point load model can not be used for solving the probe spatial resolution problem. More complicated models have been used for the tip and the cantilever [81-83]. However, the probe spatial resolution could be roughly estimated from these measurements. Commonly, the spatial resolution is defined as the

ability of resolving two adjacent signals (with a -3 dB separation). As seen in Fig. 6.11, 6.12 and 6.13, the probe with a tip-to-sample distance of 100 nm was unable to resolve two adjacent signal traces with a 1 μm separation, but was able to resolve the two adjacent signal traces with a 2 μm separation. This means that the probe spatial resolution is less than 2 μm but larger than 1 μm in this case, according to this criteria.

The resolution can also be estimated as the length of the transition from 25 to 75% of the maximum force [87]. According to Fig. 6.14 and 6.15, the resolution is also larger than 1 and less than 2.

If assuming $G_{tip}(x,y,z)$ and $G_{cantilever}(x,y,z)$ are the spatial responses of the tip and the cantilever, respectively, then the total probe spatial response $G(x,y,z)$ is

$$G(x,y,z) = G_{tip}(x,y,z) + G_{cantilever}(x,y,z). \tag{6.10}$$

Employing the method in [80] used for static surface potentials, the total force on the cantilever is approximately

$$A_z(x,y,z,\tau) = \int_{-\infty}^{+\infty} G(x,y-y',z)v_c(x,y',t=\tau)dy', \tag{6.9}$$

where $v_c(x, y, t=\tau)$ is the circuit potential distribution at $t = \tau$. The mathematical inner product represents a weighted sum of all of force contributions. To further understand the effect of distributed forces on the probe, a coupled electrostatic-mechanical analysis was performed for the EFM probes [88].

6.2 Simulation

6.2.1 Theory

As previously discussed, the spatial resolution of the electrostatic force measurement depends on the probe-circuit geometry and the interaction between the probe and the circuit, which is, in fact, a complicated average of all the forces on the probe induced by each point of the circuit interconnects. To analyse the details of the spatial resolution, one would use a full three-dimensional numerical solution of Laplace's equation to calculate the probe-circuit interaction.

A more practical solution is to estimate the coupling capacitance between the probe and the circuit which could be calculated from the Poisson's equation by using a basic mathematical technique for reducing functional equations to matrix equations, i.e. the

Method of Moments [89]. The method will be briefly described and has been used for the calculation of the coupling capacitances between the probe and the three parallel metal traces. The simplified model of the physical system is shown in Fig. 6.16.

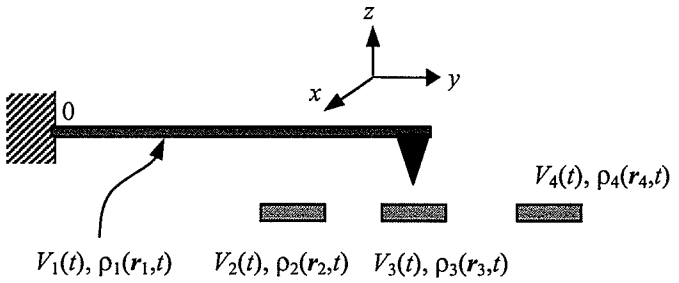

Fig. 6.16. The simplified model of the physical system.

Considered in a region of constant permittivity ε and constant volume charge density ρ, the electrostatic potential should satisfy the Poisson's equation

$$\nabla^2\phi(x,y,z) = -\frac{\rho}{\varepsilon}. \tag{6.11}$$

The potential at the point (x,y,z) is

$$\phi(x,y,z) = \iiint_V \frac{\rho(x',y',z')}{4\pi\varepsilon\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}}dx'dy'dz'. \tag{6.12}$$

For all perfect electric conductors (with boundary condition $\phi=V$) without dielectrics (the free space, $\varepsilon=\varepsilon_0$), the problem can be numerically solved by discretizing the surfaces of the conductors into sub-areas (patches) and approximating the charge on each sub-area as an impulse point charge concentrated at the centre of each sub-area (finite elements). The Eq. (6.12) becomes

$$V_m = \sum_{n=1}^{N} \frac{\alpha_n}{4\pi\varepsilon\sqrt{(x_m-x_n)^2+(y_m-y_n)^2+(z_m-z_n)^2}}A_n + \alpha_m\sqrt{A_m}, m = 1,...,N, \tag{6.13}$$

where α_n is the charge density on the patch n with the area A_n and N is the total number of the patches (i.e. the conductor number one has patches from number 1 to N_1, the conductor number two has patches from number N_1+1 to N_2, ..., and the conductor I has patches from number N_{I-1}+1 to N).

Equation (6.13) an be expressed in a matrix form as $[V_m]=[l_{mn}][\alpha_n]$. The charge density can be solved from this matrix equation, i.e. $[\alpha_n]=[l_{mn}]^{-1}[V_m]$. The corresponding capacitance for each conductor is

$$C_{jk} = \frac{Q_j}{V_k} = \frac{1}{V_k}\sum_{j}^{N_j}\alpha_j A_j = \sum_{j}^{N_j}\sum_{n}^{N} l_{nj}^{-1} A_j. \tag{6.14}$$

where C_{jk} is the coupling capacitance between the conductors j and k, and Q_j is the total charges on conductor j.

6.2.2 Simulation results

To compare the theory with the experiment, the method was used to simulate the EFM probe and the smallest three traces of the ITG circuit. The accuracy of the method depends on the matrix, $[l_{mn}]$, and how accurate it can represent the charge distribution. The exact geometry of the probe has to be carefully taken into consideration for discretization.

In this work, the geometry of the EFM probe was obtained from its data sheet [41] and from the scanning electron microscopy images as shown in Fig. 2.2. The tip was modelled as a cone with a hemisphere apex as shown in Fig. 6.17. The three conductive traces and the cantilever beam were modelled as zero height conductors. The cantilever support block was modelled with a size of 1 μm × 1 μm. The circuit ground plane was located at 1 μm below the three traces with a substrate of free space between them. The length of the traces was much longer than the width of the cantilever.

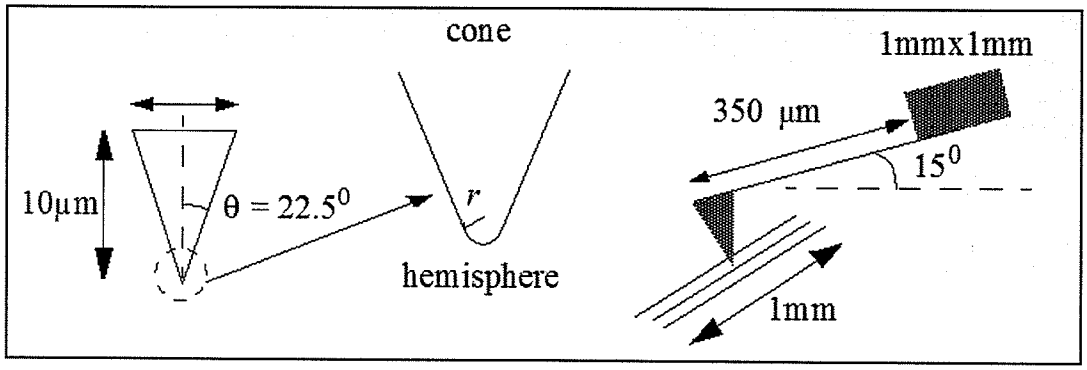

Fig. 6.17. The silicon tip, the cantilever and the cantilever supporting block geometry model.

For stability, the discretization avoids long and narrow patches and the distance between each two adjacent patches is $\sim\sqrt{A}$, where A is the patch area. Two adjacent

patches should have approximately the same area while the total area of all the patches should be equal to the total physical area. In the simulation, 2971 patches in total were generated for the probe and the three traces. The geometric model with these patches as seen in Fig. 6.18 can be viewed through a program called GeoView [88] for verification.

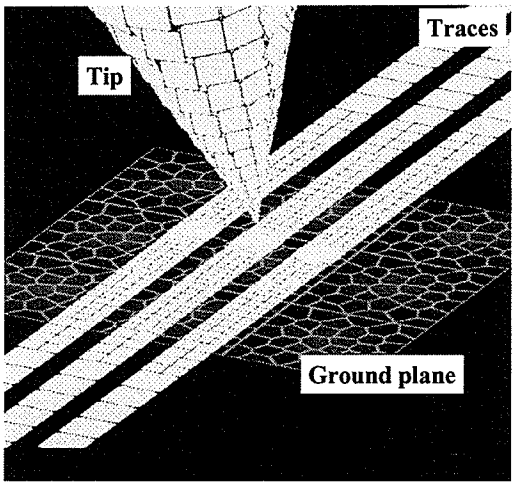

Fig. 6.18. The patches for the tip and the circuit traces with and a ground plane below. The trace sizes are 2 μm with a 1 μm separation, i.e. 2:1:2:1:2 [88].

The coupling capacitances between the probe and the trace 1 and all three traces were calculated as seen in Fig. 6.19 and 6.20, respectively, for the probe tip position, y, scanned from –3 μm to +3 μm (y=0, the centre of the trace 1) and tip-to-sample distance, z. z changes from 100 nm to 2 μm, which is the typical range for the EFM operation. From these capacitance profiles, the capacitance derivatives (the 2D vertical force mapping) could be calculated. The capacitance derivative calculated from Fig. 6.19 is shown in Fig. 6.21, and correspondingly the force calculated from Fig. 6.20 is shown in Fig. 6.22.

The calculated capacitance derivative results (Fig. 6.21 and 6.22) show a better probe spatial resolution in comparison with the experimental results (Fig. 6.9). Using the same definition as discussed in the previous section, the simulated probe spatial resolution for the tip-to-sample distance of 100 nm is about 1.5 μm.

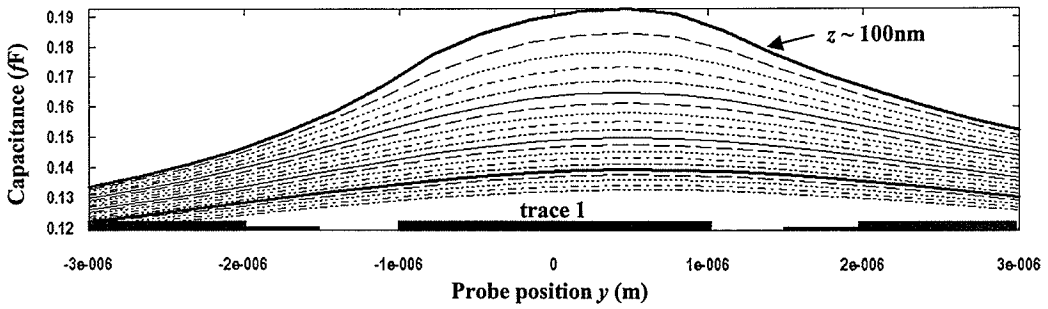

Fig. 6.19. The simulated coupling capacitance between the probe and the trace 1. The tip-to-circuit distance changes from 100 nm to 2 μm (100 nm for each increment).

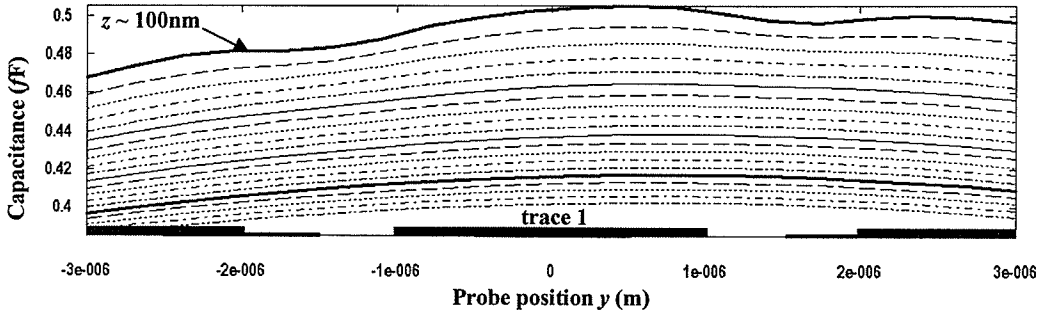

Fig. 6.20. The simulated coupling capacitance between the probe and all three traces. The tip-to-circuit distance changes from 100 nm to 2 μm (100 nm for each increment).

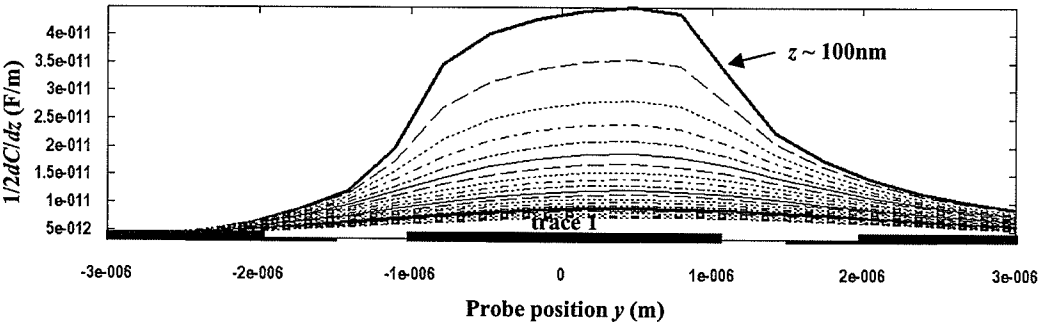

Fig. 6.21. The capacitance derivative calculated from Fig. 6.19. The tip-to-circuit distance changes from 100 nm to 2 μm (100 nm for each increment).

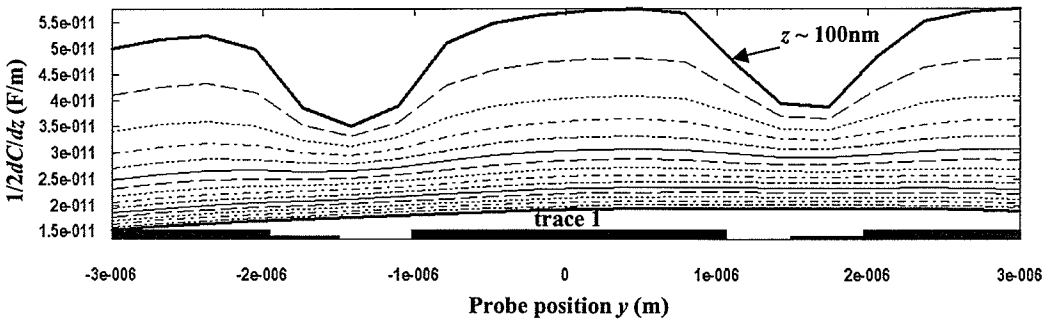

Fig. 6.22. The capacitance derivative calculated from Fig. 6.20. The tip-to-circuit distance changes from 100 nm to 2 μm (100 nm for each increment).

The normalized theoretical capacitance derivative, $\frac{1}{2}dC_i/dz$, between the probe, which is positioned above the center of the trace 1, and the trace i=0,1,2, is given in Fig. 6.23 as a function of tip-to-sample distance. The results are compared with the measured electrostatic force induced cantilever deflection A_{zi}, due to the same signal $v_i(t)$, which was applied individually to the three traces.

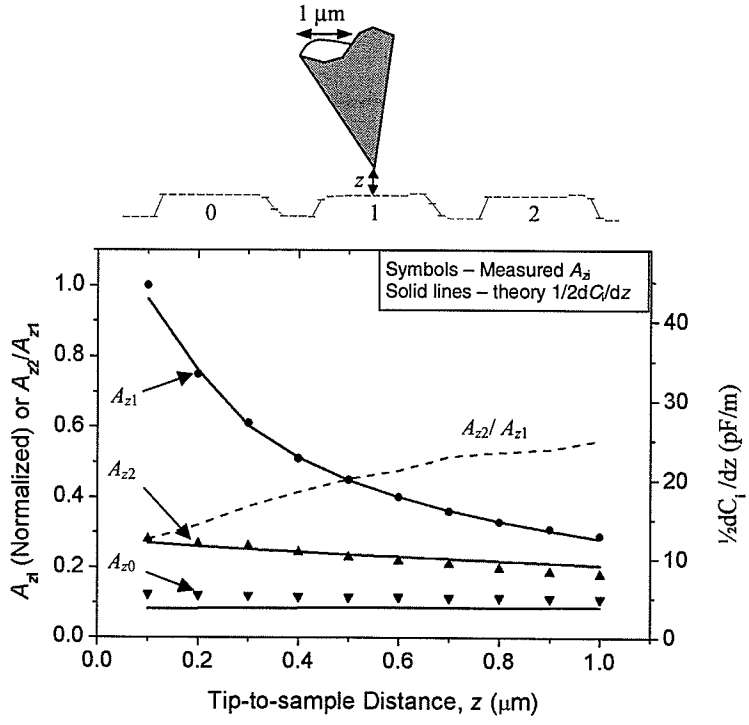

Fig. 6.23. The measured cantilever deflection A_{zi} [due to the same signal $v_i(t)$, which was individually applied to the three traces] as a function of the probe-to-sample distance z with the probe above the trace 1. The normalized deflection A_{zi} is compared with the simulated $1/2dC_i/dz$ values. The dashed line indicates the measured interference level, A_{z2}/A_{z1}, due to the trace 2.

Figure 6.23 also shows the interference level, A_{z2}/A_{z1}, due to the trace 2, which varies from 30% at z=100 nm to > 50% at z = 1 μm. This result is in agreement with Table 2. The figure shows that the deflection caused by the signals on the adjacent traces is far less sensitive to the tip-to-sample distance and their influence would be reduced using a force-gradient method as discussed in the next chapter.

Figure 6.24 also shows the capacitance second order derivative $½d^2C/dz^2$ calculated from Fig. 6.22 for comparison and to show that the higher order capacitance derivative (force-gradient) has better spatial resolution. This is the focus of the next chapter.

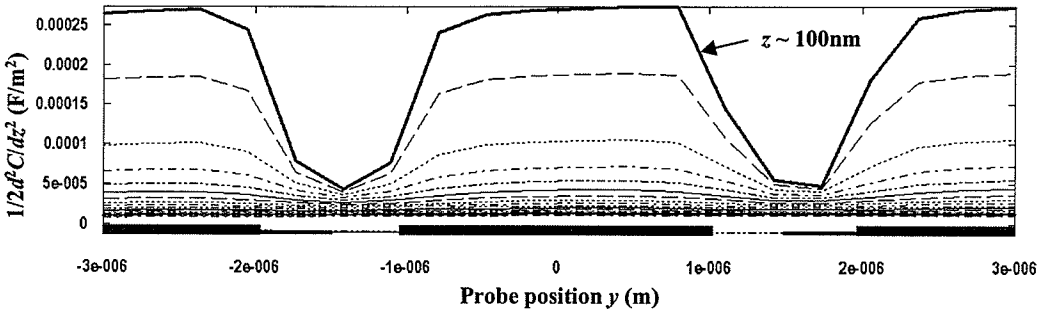

Fig. 6.24. The capacitance second order derivative calculated from Fig. 6.21. The tip-to-circuit distance changes from 100 nm to 2 μm (100 nm for each increment).

Chapter 7

Electrostatic Force-Gradient Microscopy

The results of Chapter 6 implies that detection schemes based on higher order capacitance deriatives can yield a better spatial resolution. Chapter 7 proposes two dynamic force-gradient microscopy approaches to improve the probe spatial resolution. The first one is a numerical force-gradient detection approach, in which the force gradient is extracted numerically from directly sensing the force induced deflections for two tip-to-sample distances. The second approach enables direct detection of high-frequency force-gradients by employing a high-frequency implementation of a mechanoelectric cross-modulation scheme. It involves using an external vibration oscillator to mechanically modulate the tip-to-sample distance. Measurements on three parallel interconnects of an IC operating up to 500 Mbit/s demonstrate a submicron spatial resolution and a significant reduction of interference by using the force-gradient approaches.

7.1 Force-gradient method in DEFM

As discussed in the previous chapter, the electrostatic force on the cantilever is not only a function of time, but is also a function of the tip-to-sample distance, z, as expressed in Eq. (6.1). The force gradient can be derived by differentiating the Eq. (6.1) with respect to the tip-to-sample distance, z, as

$$\frac{\partial}{\partial z}F_z(z,t)=\frac{1}{2}\frac{\partial^2}{\partial z^2}C(x,y,z)[v_p(t)-v_c(x,y,t)]^2. \tag{7.1}$$

The force gradient depends on the second derivative of the probe-circuit coupling capacitance and the square of the potential difference. From this equation, the circuit voltage $v_c(x,y,t)$ can be extracted by detecting the force gradient instead of the force as is discussed in previous chapters. Direct force-gradient detection methods will be discussed in the next section.

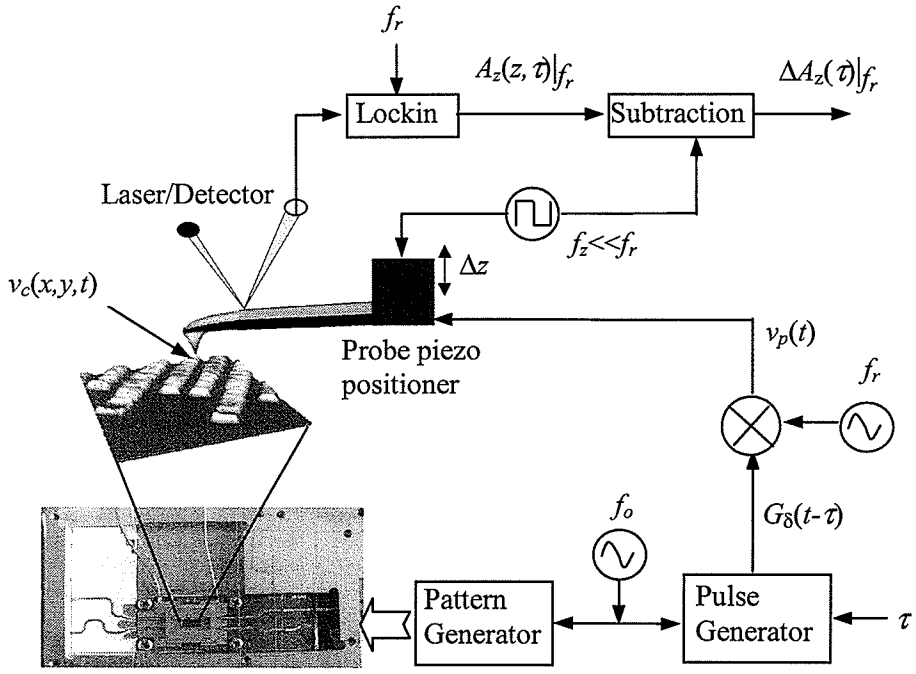

Fig. 7.1. Block diagram of the dynamic electrostatic force probing instrument which employs a force-gradient method.

In this section, a numerical force-gradient method is proposed. The force-gradient method is based on sensing the force as the tip-sample distance is modulated [90]. A schematic diagram of the instrument is shown in Fig. 7.1. The tip-to-sample distance is step modulated at a slow rate, $f_z<<f_r$, by means of a piezo as shown in Fig. 7.1. The change in probe deflection amplitude for the change of the tip-to-sample distances, $\Delta z=(z_1-z_2)>0$, is implemented numerically and can be described as [90]

$$\left.\frac{\Delta A_z(z,\tau)}{\Delta z}\right|_{f_r,f_z} = \left.\frac{[A_z(z_1,\tau)-A_z(z_2,\tau)]}{\Delta z}\right|_{f_r} \approx \frac{\partial^2 C(x,y,z)}{\partial z^2}\frac{Q}{k}\frac{\delta}{T}v_c(x,y,t=\tau)\sin(2\pi f_r t). \quad (7.2)$$

Equation (7.2) is an approximation of the force gradient at a tip-to-sample distance $z = (z_1+z_2)/2$. For a small Δz, the dominant error for the equation (7.2) can be estimated from the Taylor series expansion of $\Delta A_z(z,\tau)$,

$$A_z(z_1,\tau)-A_z(z_2,\tau) \approx \frac{\partial A_z(z,\tau)}{\partial z}\Delta z + \frac{1}{3}\frac{\partial^3 A_z(z,\tau)}{\partial z^3}(\frac{\Delta z}{2})^3 + O((\Delta z)^5), \quad (7.3)$$

to be proportional to $(\Delta z^2/24)(\partial^4 C/\partial z^4)$, notes that $A_z \propto \partial C/\partial z$. The accuracy of the equation (7.2) will be further discussed later in this section.

For probing three parallel conducting traces as discussed in the previous chapter, the detected force gradient can also be expressed in three parts

$$\frac{\Delta A_z(z,\tau)}{\Delta z} = \frac{\Delta A_{z0}(z,\tau)}{\Delta z} + \frac{\Delta A_{z1}(z,\tau)}{\Delta z} + \frac{\Delta A_{z2}(z,\tau)}{\Delta z}, \tag{7.4}$$

where one of the three parts is the desired signal and the others are the interfering signals from the adjacent traces. $\Delta A_{zi}(z,\tau)$; i=0,1,2, is proportional to $\partial^2 C_i(x,y,z)/\partial z^2$, respectively. Since the coupling capacitance has an inverse relationship with respect to the tip-to-sample distance, the force-gradient method using the second-order derivative of the capacitance yields better spatial resolution. This means that the interfering signals will be much smaller than the desired signal in the force-gradient measurement due to the relatively larger distances of the adjacent traces to the probe tip.

Figure 7.2 shows the force-gradient, ΔA_{zi}, calculated numerically using (7.2) and the measured A_{zi} values from Fig. 6.23. The ΔA_{zi} result is compared with the theoretical $\frac{1}{2}\partial C_I^2/\partial z^2$ values determined by simulation and indicates the validity of the numerical approach. The interference level, $\Delta A_{z2}/\Delta A_{z1}$, is <10% when the tip-to-sample distance z is <300 nm (for comparison, recall that A_{z2}/A_{z1}= 30% at z=100 nm in Fig. 6.23). The force-gradient method should achieve a significant reduction in cross talk.

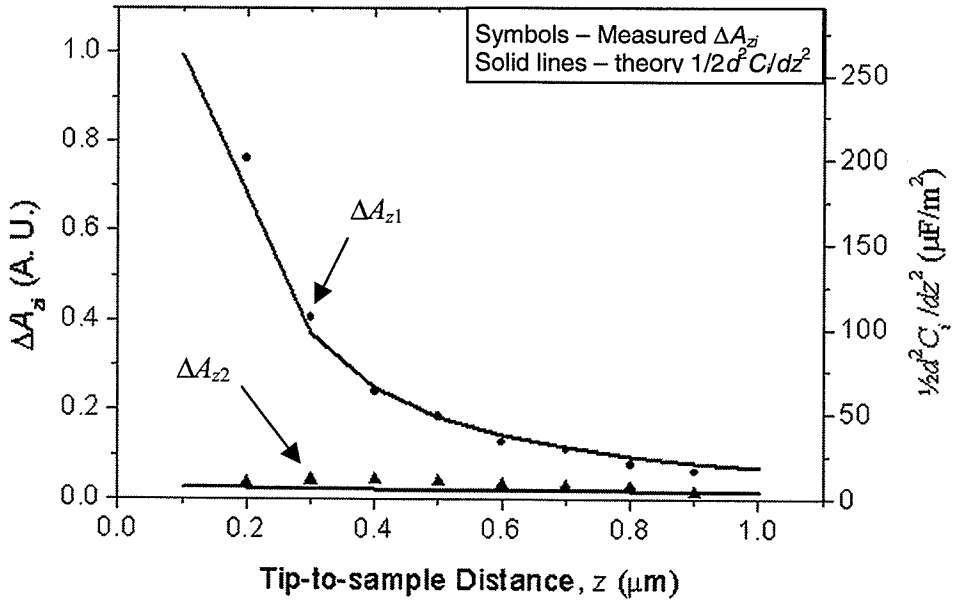

Fig. 7.2. The numerically calculated force-gradient, ΔA_{zi}, using the equation (7.2) and the measured A_{zi} values from Fig. 6.23. The ΔA_{zi} result is compared with the simulated $\frac{1}{2}\partial^2 C_i/\partial z^2$ values. The interference level, $\Delta A_{z2}/\Delta A_{z1}$, is less than 10% when z < 300 nm.

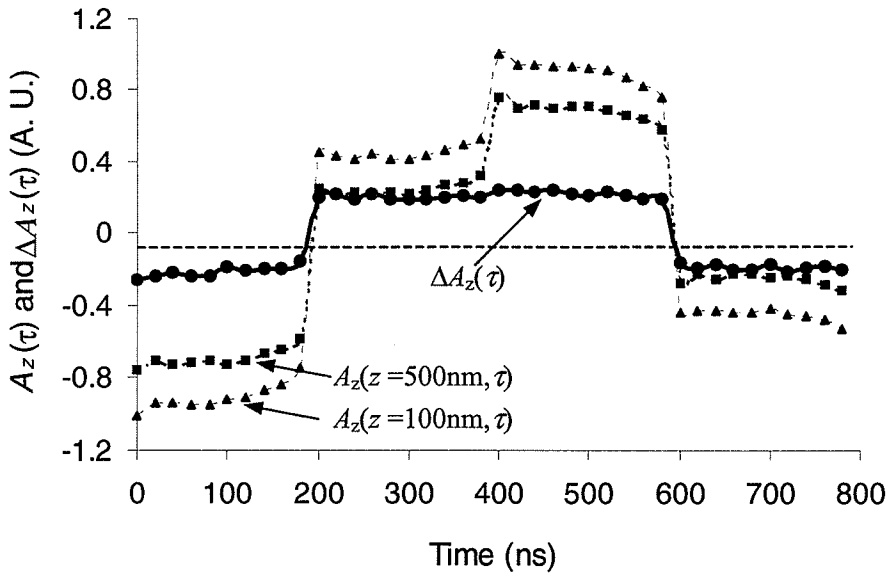

Fig. 7.3. The force-gradient, $\Delta A_z(\tau)$, calculated from the measured deflection signals, $A_z(z,\tau)$, when the probe tip-to-sample distance is modulated from z_1=100 nm to z_2= 500 nm. The interference signal was significantly reduced.

The force-gradient method was applied to the high speed waveform measurements. The measurement setup is similar to that used in Fig. 6.5 in Chapter 6. The probe tip is located over the trace 1 to sample $v_1(t)$ and an interference signal $v_2(t)$ is applied to the trace 2. Fig. 7.3 shows the force-gradient, $\Delta A_z(\tau)$, calculated numerically from the deflection signals, $A_z(z,\tau)$, when the probe tip-to-sample distance is modulated from z_1=100 nm to z_2=500 nm (Δz=400 nm). $\Delta A_z(\tau)$ is clearly a better representation of the signal $v_1(t)$ than the force induced deflection signals, $\Delta A_z(z,\tau)$. Similarly, in Fig. 7.4, $v_1(t)$ is an 8-bit [01011100], 100 ns/bit, repeating pattern; $v_2(t)$ is an 8 bit [00001111] pattern and $v_0(t)$=0. The force-gradient signal in Fig. 7.4, $\Delta A_z(\tau)$, is a better representation of the signal of interest, $v_1(t)$. The interference level presented in the force-gradient measurements is significantly reduced in comparison to the reduction of the interference levels presented in the force measurements.

To improve the measurement bandwidth, a 1 ns sampling pulse was used for higher bandwidth force-gradient measurements.

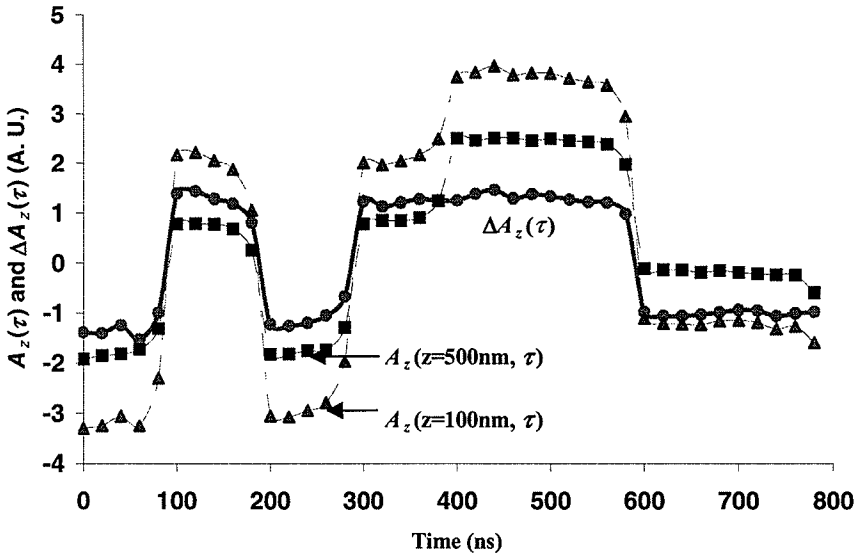

Fig. 7.4. The force-gradient, $\Delta A_z(\tau)$, calculated from the measured deflection signals, $A_z(z, \tau)$, when the probe tip-to-sample distance is modulated from z_1=100 nm to z_2=500 nm. The circuit signal $v_1(t)$ is an 8-bit [01011100], 100 ns/bit, pattern, $v_2(t)$ is an 8-bit [00001111] pattern and $v_0(t)$=0.

As an example, force and force-gradient DEFM measurement of 500 Mbit/sec signals applied to the test circuit of the three traces is shown in Fig. 7.5. Here $v_1(t)$ is an 8-bit [01010011], 2 ns/bit, repeating pattern; $v_2(t)$ is an 8-bit [01101010] pattern and $v_0(t)$=0. The probe was located above the trace 1 and a δ=1 ns sampling pulse was used. The waveform measured by direct force probing, $A_z(z=100\text{nm}, \tau)$, is given by the curve (b) and shows obvious interference from $v_2(t)$. The force-gradient signal $\Delta A_z(\tau)$, as the probe tip-to-sample distance was modulated from 100 nm to 500 nm, is given by the curve (c). The actual circuit signal $v_1(t)$ was measured by a 3 GHz bandwidth active contact probe (Picoprobe Model 34A [4]) and is given by the curve (a) in Fig. 7.5 after performing a convolution with the 1 ns probe sampling pulse. The force-gradient signal matches well with the active probe result and even identifies distortion of the signal due to the inter-trace coupling at high frequencies.

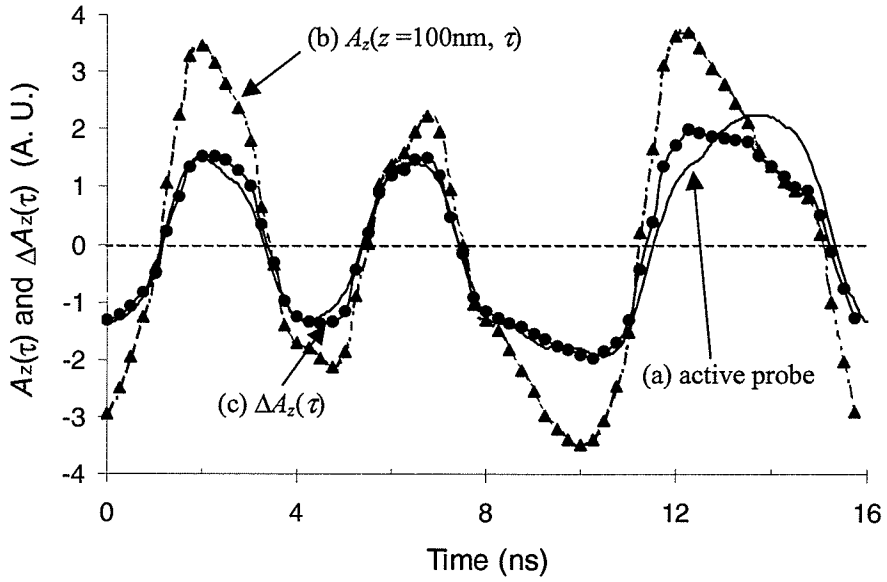

Fig. 7.5. Measurement of 500 Mbit/sec signals applied to the test circuit of the three traces. The signal $v_1(t)$ is an 8-bit [01010011], 2 ns/bit, pattern, $v_2(t)$ is an 8-bit [01101010] pattern and $v_0(t)$=0. Curve (a) shows $v_1(t)$ as measured with a 3GHz active probe and after performing a convolution with the 1 ns probe sampling pulse. Curve (b) shows the waveform $A_z(z$=100 nm, $\tau)$ measured by direct force probing. Curve (c) shows the force-gradient $\Delta A_z(\tau)$, when the probe tip-to-sample distance was modulated from 100 nm to 500 nm.

Figure 7.6 shows the force-gradient, ΔA_z, as the probe was scanned across the three traces of the test circuit with the tip-to-sample distance modulated from 100 nm to 500 nm. The experiment was similar to Fig. 6.9(c) (the cantilever extended to the right), only the trace 1 was energized, $v_1(\tau)$= 0.75 V. The result shows that the background signal due to coupling to the cantilever is suppressed and that a submicron spatial resolution is achieved. When performing scanning measurements, error in the gradient may be significant due to variations in the probe height from the piezo non-linearity. Usually point-by-point time-domain waveform measurement is implemented in IC functional testing and then only topographic scanning is needed to position the probe.

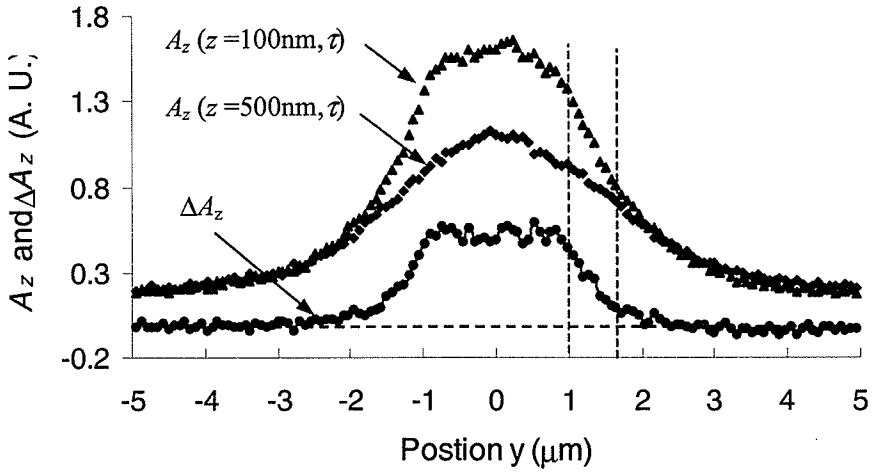

Fig. 7.6. The force-gradient, ΔA_z, signal as the probe is scanned across the three traces of the test circuit. Similar to Fig. 6.9(c), only the trace 1 was energized, $v_1(\tau)$= 0.75 V.

The coupling to the cantilever shown in the force measurements of the rising edge of the 32 μm-wide trace 1 as shown in Fig. 6.14 (the non-zero deflection background signals exist even when the tip is far away from the signal trace) is clearly suppressed in the force-gradient signal, ΔA_z, as shown in Fig. 7.7. Similarly, Fig. 7.8 shows the force-gradient, ΔA_z, as the probe was scanned across the falling edge of the 32 μm wide trace 1. The force-gradient signals drop quickly to zero as the tip was scanned about 2 μm away from the edges.

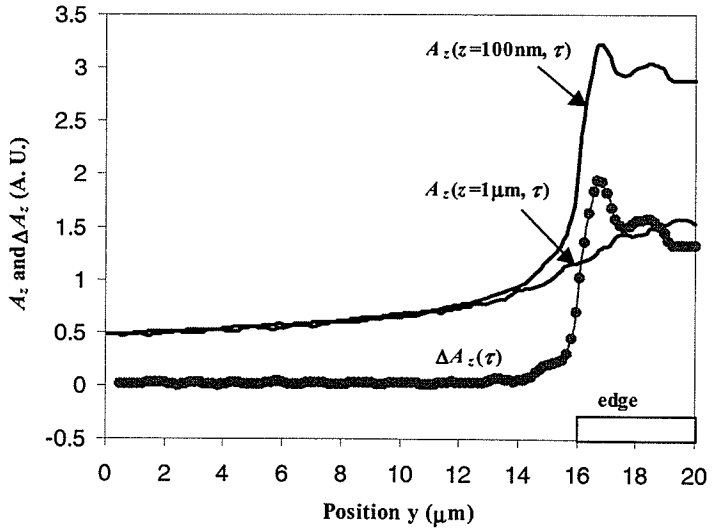

Fig. 7.7. The force-gradient, ΔA_z, signal on the rising edge of the 32 μm-wide trace 1. The circuit was stimulated with 2.5 Mbit/s square-wave signals with $v_1(\tau)$ = 0.75 V, and $v_0(\tau) = v_2(\tau)$= 0 V.

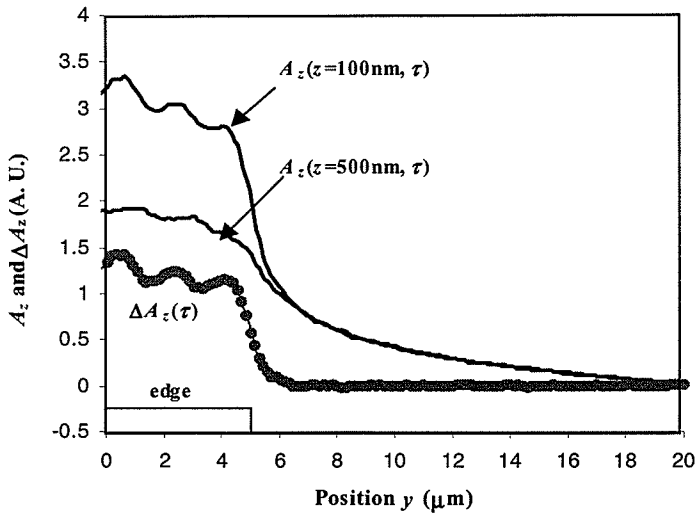

Fig. 7.8. The force-gradient, ΔA_z, signal on the falling edge of the 32 μm-wide trace 1. It was stimulated with 2.5 Mbit/s square-wave signals with $v_1(\tau) = 0.75$ V, and $v_0(\tau) = v_2(\tau) = 0$ V.

The dominant error for the force-gradient calculation using Eq. (7.2) is proportional $(\Delta z^2/24)(\partial^4 C/\partial z^4)$ (for small Δz) and can be estimated by numerically processing the $\frac{1}{2}\partial^2 C/\partial z^2$ data in Fig. 6.22 to be <1% for z = 300 nm and Δz=50 nm. Error in the force gradient could be significant when using Eq. (7.2) for small z (<100 nm) or large Δz (in relation to z). In this specific case, the force-gradient method is used successfully for the reduction of interference from signals on the parallel interconnects and a better spatial resolution is achieved as demonstrated by these measurements.

The cost of employing the force-gradient approach is a decrease in the signal-to-noise ratio of the measurement. Assuming the noise in the system is dominated by the thermal noise of the cantilever, the signal-to-noise ratio (SNR) for direct force measurements is derived in [39]. For a small Δz, the SNR for the force-gradient method employed is scaled by a factor of $\Delta A_z/A_z$, using the approximation $\partial A_z/\partial z \sim \Delta A_z/\Delta z$, to be [90]

$$SNR = \Delta z \frac{\partial^2 C_i}{\partial z^2} \frac{\delta}{T} \Delta v_{crms} \sqrt{\frac{\omega_r Q}{4kk_B TB}}, \tag{7.5}$$

where B is the bandwidth of the system, T is the absolute temperature, and k_B is Boltzmann's constant. In this measurement, B is 1 Hz, and T is the room temperature. Using the mechanical specifications of the cantilever, capacitance derivative data from Fig. 7.2, and for Δz=400nm and δ/T =1/40 as shown in Fig. 7.6, the minimum detectable

signal (at *SNR*=1) is about 14 mV/$\sqrt{\text{Hz}}$. For a small Δz, the SNR of the force-gradient could be too small to be measured by this method. In the following section, direct force-gradient detection methods for the improvement of the spatial resolution will be discussed.

7.2 Direct force-gradient detection in DEFM

The direct force-gradient method has been used for electrical contrast observations and dc voltage measurements in Kelvin probe force microscopy [91,92,93]. Up to now, the force gradient has not been proposed for high frequency voltage measurement. The aims of this section are to discuss the direct force gradient detection methods and how to implement the force gradient measurement in the DEFM for high frequency voltage measurement. The force gradient measurements demonstrate an improved spatial resolution for the high frequency voltage measurements.

7.2.1 Cantilever frequency response to a force gradient

Since the electrostatic force is not a uniform function of z, for a small variation δz to the tip-to-sample distance z_o, the electrostatic force can be expressed in a Taylor series expansion form as

$$F_z(z,t) = F_z(z_o,t) + \delta z F_z'(z_o,t) + O(\delta z^2), \tag{7.6}$$

where the prime in $F_z'(z_o,t)$ represents $\partial/\partial z$ and $O(\delta z^2)$ represents the higher orders of δz which we will ignore.

For a small variation, the static component of the force gradient is usually only considered [91]. The equation of motion Eq. (2.8) is this case becomes

$$m_T \frac{d^2 z}{dt^2} + \gamma \frac{dz}{dt} + kz = W_T + F_z(z_o,t) + F_z'(z_o)z. \tag{7.7}$$

The Fourier transformation is

$$-m_T \omega^2 Z(\omega) + j\omega\gamma Z(\omega) + kZ(\omega) = F_z(z_o,\omega) + F_z'(z_o)Z(\omega). \tag{7.8}$$

We will use the notation: $F_z(\omega)$ for $F_z(z_o,\omega)$, and F_z' for $F_z'(z_o)$.

The cantilever frequency response $|H(\omega)|$ can then be expressed as

$$|H(\omega)| = \frac{Q/k}{\sqrt{(1-\frac{F_z'}{k}) + Q^2(1-\frac{F_z'}{k}-\frac{\omega^2}{\omega_r^2})^2}}. \tag{7.9}$$

Comparing Eq. (7.7) with Eq. (2.15), the effective spring constant of the cantilever in the presence of the force gradient may be given by $k' = k - F_z'$. From the calculated force gradient for a probe located above the trace 1 as shown in Fig. 7.2 in the previous section, $F_z'/k \sim 0.002$ for $z \sim 300$ nm. Figure 7.9 shows an example of the comparison of Q=16 and F_z'/k=0.05. The maximum response of the cantilever occurs at a frequency ω_r' as

$$\omega_r' = \omega_r\sqrt{1-\frac{F_z'}{k}}. \tag{7.10}$$

For a small F_z'/k, the resonance frequency shift is approximately $\Delta\omega \approx -\frac{1}{2}(F_z'/k)\,\omega_r$. This is about 12.5 Hz for $F_z'/k \sim 0.002$ and 312.5 Hz for $F_z'/k \sim 0.05$ (assuming ω_r = 12.5 kHz).

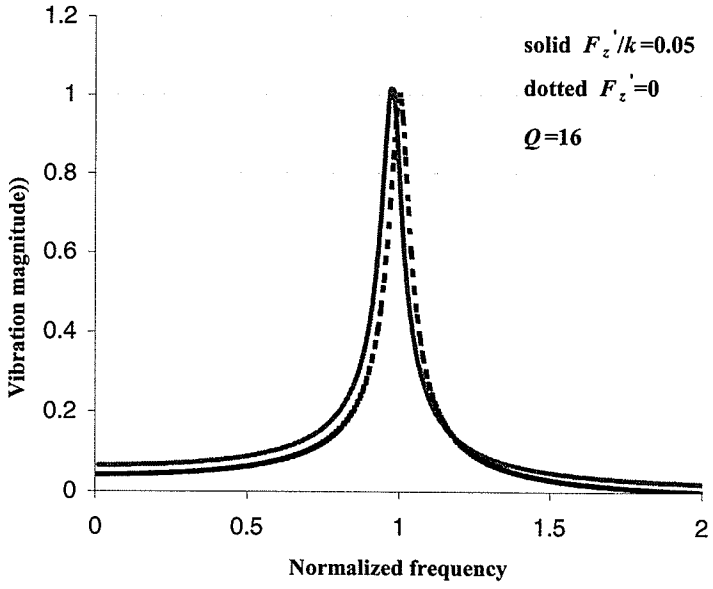

Fig. 7.9. The magnitude of the cantilever mechanical response in the presence of a force gradient (the solid curve), and without a force gradient (the dotted curve). This example calculation is for Q = 16 and F_z'/k = 0.05.

The relative phase of the vibration is

$$\theta(\omega) = -\tan^{-1}\frac{\omega/\omega_r}{Q(1-\frac{F_z'}{k}-\frac{\omega^2}{\omega_r^2})}. \tag{7.11}$$

Figure 7.10 shows an example comparison of the vibration phase with a force gradient and without a force gradient for Q=16 and F_z'/k =0.05. For a small F_z'/k, the maximum variation of the phase is approximately $\Delta\theta \approx -\frac{QF_z'}{k}$. This is about 1.8° for $F_z'/k \sim 0.002$ and 45.8° for $F_z'/k \sim 0.05$ (assuming that Q=16).

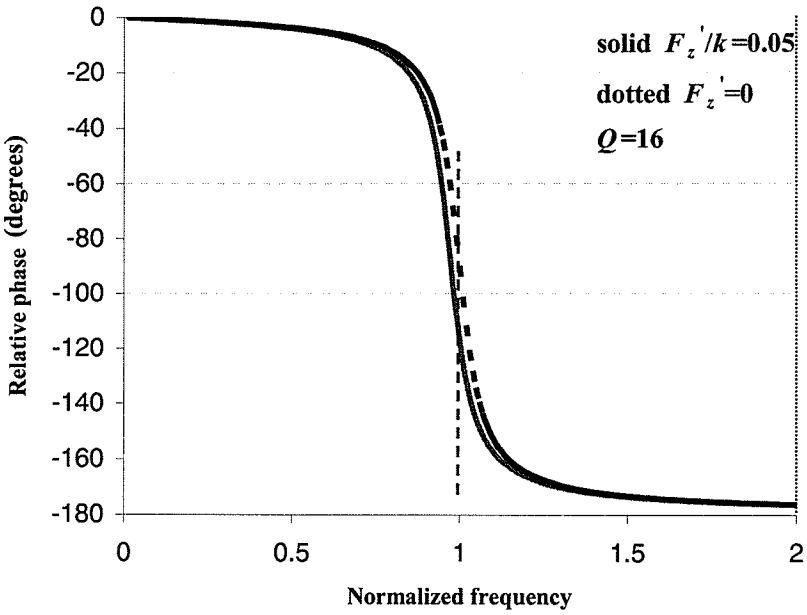

Fig. 7.10. The relative phase of the cantilever mechanical response in the presence of a force gradient (the solid curve), and without a force gradient (the dotted curve). This example calculation is for Q =16 and F_z'/k = 0.05.

7.2.2 Direct force gradient detection

In the most commonly used direct force gradient detection schemes, the cantilever is driven at a constant frequency near its fundamental resonance and the force gradients are detected as variations in the amplitude (amplitude slope detection) [94], the phase (phase detection) [95] and the frequency (frequency modulation detection) [96].

Table 4 lists the detailed information about these three detection methods (assuming a large Q and a small and time-invariant force-gradient, i.e. $F_z'/k \ll 1$). For the amplitude slope detection method, the maximum slope of the cantilever amplitude response occurs at $\omega \approx \omega_r(1 \pm 1/\sqrt{8}Q)$ ($\partial^2|H(\omega)|/\partial\omega^2=0$) and at this point the vibration amplitude change due to the force-gradient F_z' is given by $\Delta A_z \approx -\frac{2QF_z'}{3\sqrt{3}k}A_z$ [94]. Similarly, for the phase detection, the maximum phase variation occurs at $\omega \approx \omega_r(1-1/8Q^2) \approx \omega_r$ ($\partial^2\theta(\omega)/\partial\omega^2=0$) [95]. At this frequency, the phase variation due to the force-gradient F_z' is given by $\Delta\theta \approx -\frac{QF_z'}{k}$, which can be derived from (7.11) using the approximation $\partial\theta(\omega)/\partial\omega \approx 2Q/\omega_r$ and $\Delta\omega_r \approx -\frac{F_z'}{2k}\omega_r$. The minimum detectable force gradients in the table were determined by the thermal noise [94,97].

Table 4 Detailed information about the three force gradient detection methods.

Detection method	Detection variation equation	Maximum variation	Minimum detectable F_z'
FM detection [96]	$\omega_r' = \omega_r\sqrt{1-\frac{F_z'}{k}}$	$\Delta\omega_r \approx -\frac{F_z'}{2k}\omega_r$	$\sqrt{\frac{4kk_BTB}{Q\omega_r A_z^2}}$
Amplitude slope detection [94]	$A_z = \frac{QF_z}{k}\left[\sqrt{(1-\frac{F_z'}{k})-Q^2(1-\frac{F_z'}{k}-\frac{\omega}{\omega_r})^2}\right]^{-1}$	$\Delta A_z \approx -\frac{2QF_z'}{3\sqrt{3}k}A_z$	$\sqrt{\frac{27kk_BTB}{Q\omega_r A_z^2}}$
Phase detection [95]	$\theta = -\tan^{-1}\left[\frac{\omega}{Q\omega_r}\left[1-\frac{F_z'}{k}-\frac{\omega^2}{\omega_r^2}\right]^{-1}\right]$	$\Delta\theta \approx -\frac{QF_z'}{k}$	$\sqrt{\frac{4kk_BTB}{Q\omega_r A_z^2}}$

To improve the force gradient measurement sensitivity, experiments are usually performed in a vacuum. Under vacuum, the force gradient measurement is generally based on frequency shift detection [98]. At about 10^{-3} torr, the cantilever vibration Q factor may increase by at least two or three orders of magnitude [98]. The sensitivity of the force gradient measurement will improve $\sqrt{Q}$ times.

In the following section, the phase detection is studied for EFM voltage measurement.

7.2.2.1 Phase detection

Usually, the force gradient detection methods employ an external vibration oscillator to modulate the probe-to-sample distance and activate the force gradients. In this work, the desired electrostatic force induces a variation of the tip-to-sample distance δz, i.e. the vibration amplitude of the cantilever, also activates a force gradient above the sample. This force gradient depends on the circuit voltage and the second capacitance derivative as given by Eq. (7.1).

In this work, the phase detection method was used for force gradient detection. The phase variation could be measured by the same lock-in amplifier used in the EFM for the cantilever amplitude detection. The phase variation measurement setup is the same as the amplitude measurement setup. In this work, phase variation measurements were performed on the ITG circuit with the three parallel traces (Fig. 6.2).

From the force simulation (Fig. 6.23), the force induced cantilever vibration amplitude (assuming there is a 1 V voltage difference) can be estimated as a function of tip-to-sample distance, as shown in Fig. 7.11. It shows that the probe vibration amplitude is usually much smaller than the probe height (the tip-to-sample distance).

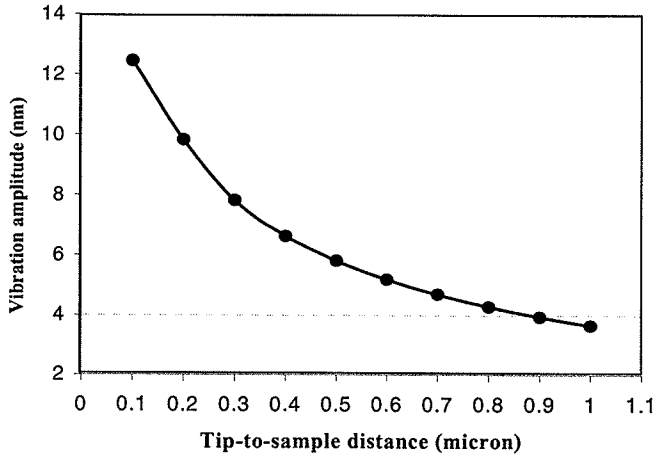

Fig. 7.11. The vibration amplitude due to the calculated force in Fig. 6.22 for the case Q = 16 and k = 0.055 N/m.

The phase variation can be estimated from the force gradient simulation, Fig. 7.2, and is shown in Fig. 7.12 (dotted curve). The phase variation measurement was performed on the ITG circuit using the same probe setup as previously shown in Fig. 6.23. The phase variation, measured at the cantilever resonance frequency, is also shown in Fig. 7.12

(solid curve) for comparison. The results of Fig. 7.12 show that the phase variation induced by the force-gradient is usually quite small, but increases quickly as the probe moves closer to the circuit.

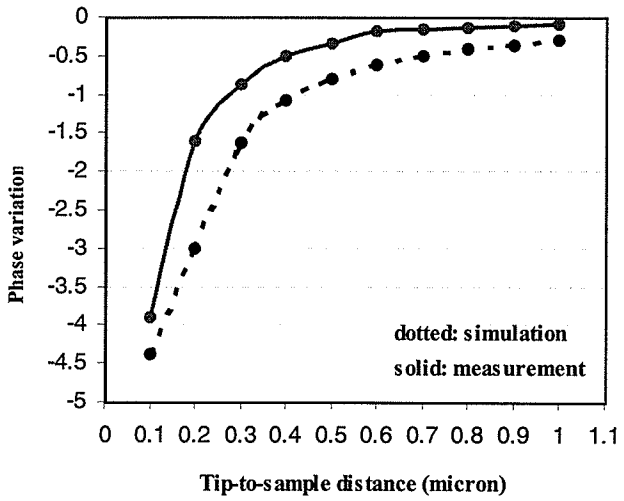

Fig. 7.12. The phase variation due to the calculated force gradient in Fig. 7.2 for Q = 16 and k = 0.055 N/m (dotted curve) and the measured phase variation (solid curve).

One dimensional scans were performed of the phase variation measurements across the three traces of the ITG circuit with the traces 1 and 2 energized and the trace 0 grounded. An example of the measurement result is shown in Fig. 7.13 for a tip-to-sample distance of 100 nm, and is compared to the measured vibration amplitude signal. The result demonstrates a much better spatial resolution (< 1 μm) for the force-gradient measurement when compared with the force measurement. The large slope over the trace 1 and the trace 2, as shown in the phase variation measurement, is mainly due to probe height variations. This may come from piezo nonlinearity and drifting.

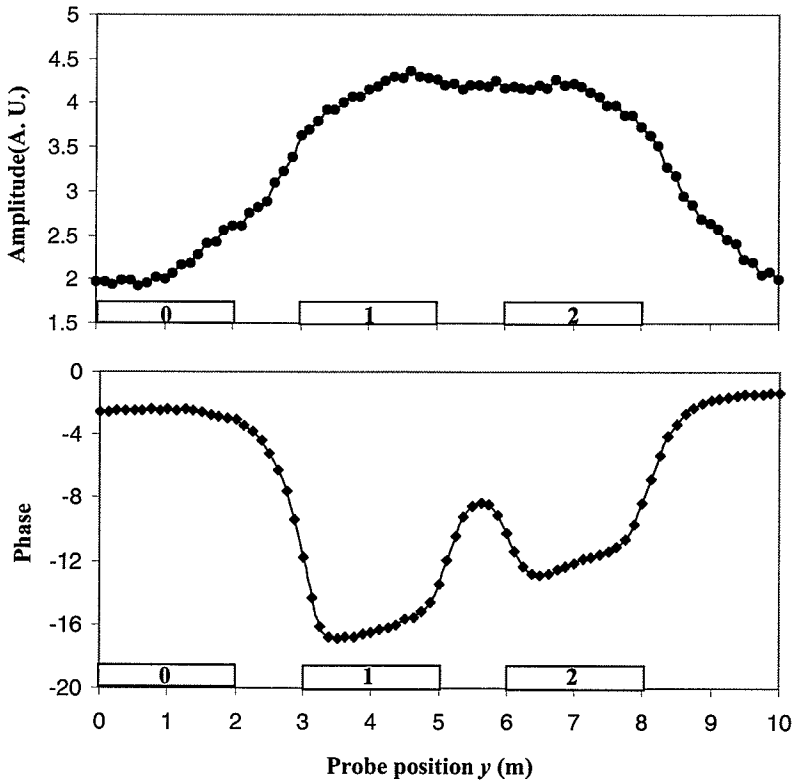

Fig. 7.13. The force (A_z) and the force gradient (the phase variation) signals 1D scan across the three traces of Fig. 6.2. Traces 1 and 2 were energized with a 4 V difference to the probe. The trace 0 was grounded.

The phase detection method avoids the use of external vibration oscillators. The SNR of the force gradient detection depends on the ratio of F_z'/k and the SNR of the force detection,

$$SNR_{phase} = \frac{F_z'}{k} SNR_{amplitude}. \tag{7.12}$$

For the EFM probe over the ITG circuit (the smallest trace 1), $F_z'/k \sim 0.002$ at $z \sim 300$ nm and $F_z'/k \sim 0.006$ at $z \sim 100$ nm. The SNR of the phase measurement method is much smaller than the SNR of the amplitude measurement method. A smaller tip-to-sample distance usually means a much better SNR. However, the smallest tip-to-sample distance is limited by mechanical noise (as discussed in chapter 4) and the probe vibration amplitude which increases as the probe moves to the surface (this was estimated in Fig. 7.11 for this specific case). In this work, the tip-to-sample distance for the optimal observation of the force-gradient is between 100 – 400 nm.

7.2.2.2 DEFM force gradient measurement

The SNR of the force gradient detection method can be improved by increasing the variation of the tip-to-sample distance. An external vibration oscillator is usually used to force the cantilever (or the circuit-under-test) to vibrate at a desired modulation depth, which is usually denoted as d_m. In comparison to the vibration phase detection method, the SNR of the force-gradient measurement will be enhanced by d_m/A_z times where A_z is the force induced vibration amplitude. For the numerical force gradient implementation discussed in the previous section (7.1), the modulation depth d_m was chosen to be as large as 200 nm (Δz = 400 nm), which is many times more than the small force induced vibration amplitude.

In this work, a mechanoelectric cross-modulation scheme [99] is applied for direct force gradient measurement in DEFM. By using this scheme, the force gradient can be detected by monitoring one of the cantilever deflection frequency components. The schematic diagram of the setup for the EFM force-gradient measurement is shown in Fig. 7.14.

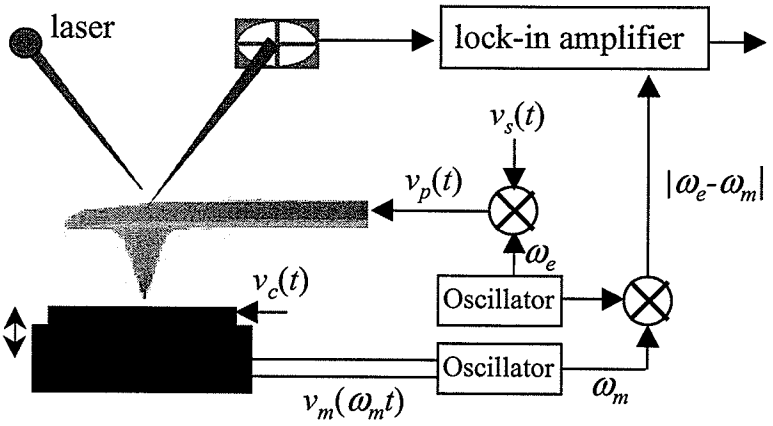

Fig. 7.14. Schematic diagram of the mechanoelectric cross modulation scheme for electrostatic force gradient detection. To make the figure concise, the pulse equivalent time sampling circuit is simplified.

A piezo-electric actuator [100] is used as an external vibration oscillator to vibrate the circuit at a frequency, ω_m, which is usually set to be well above the cantilever resonant frequency in order to avoid spurious oscillations of the cantilever. The probe-to-sample distance is modulated as $z(t) = z_0 + d_m\cos(\omega_m t)$. The probe electric signal is modulated at a frequency ω_e ($\sim\omega_m$). From Eq. (2.28), the electrostatic force component at ω_e is

$$F_z(z(t)) = -\frac{2}{\pi}\frac{\partial C(x,y,z(t))}{\partial z}\langle v_s(t), v_c(x,y,t)\rangle \cos(\omega_e t). \tag{7.13}$$

For a small d_m, (7.13) becomes

$$F_z(z(t)) \approx -\frac{2}{\pi}[\frac{\partial C(z_0)}{\partial z} + \frac{\partial^2 C(z_0)}{\partial z^2}\frac{d_m}{2}(\cos((\omega_e + \omega_m)t) + \cos((\omega_e - \omega_m)t))]\langle v_s(t), v_c(x,y,t)\rangle. \tag{7.14}$$

The desired force-gradient signal can be detected by monitoring the force induced cantilever deflection at the difference frequency $|\omega_e - \omega_m|$, where $|\omega_e - \omega_m| \leq \omega_r$, and is sensed by the cantilever. The signal is proportional to $\frac{\partial^2 C(z_0)}{\partial z^2}$, d_m and $V_c(x,y,t)$.

The essential point of this technique is the decoupling of the spurious and the target oscillation signals in the frequency domain by means of a heterodyne method. The measured signal is proportional to $F_z' d_m/2$, the force gradient. The ratio of the SNR of the force gradient signal to the SNR of the force signal is approximately $\frac{1}{2}F_z' d_m/kA_z = \frac{1}{2}(F_z'/k)(d_m/A_z)$. The force-gradient SNR using this method will increase $\frac{1}{2}d_m/A_z$ times in comparison to the DEFM phase measurement SNR. However, the force-gradient measurement error is proportional to $(d_m)^2$.

Measurements based on this method were performed on the ITG circuit [86]. The circuit is vibrated at 27 kHz with a modulation depth ~ 30 nm. The probe signal is modulated at a frequency (ω_e) so that the cantilever vibration component at the frequency difference $|\omega_e - \omega_m|$ is selected to be near the cantilever resonance frequency and can be measured by a lockin amplifier.

The test circuit contains four traces (widths of 20 μm for traces 0 and 32 μm for traces 1, 2, 3; the spacing between these traces is 16 μm) as shown in Fig. 7.15. The cantilever is extended over the traces while the tip is located at the middle of the trace 0 as seen in Fig. 7.15 (b). To maximize the signal-to-noise-ratio, the probe signal pattern is chosen the same as the circuit signal pattern (both are 4 Mbit/s square waves) with V_p=2V.

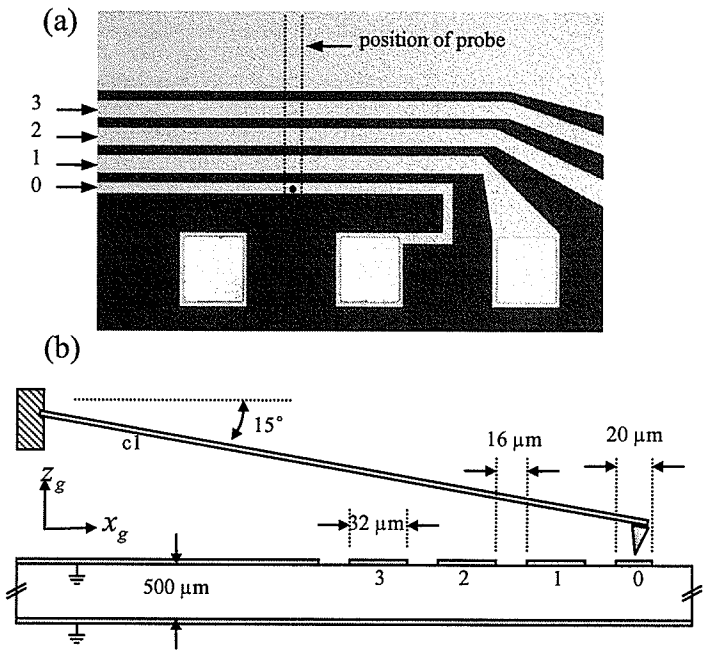

Fig. 7.15. (a) Top-view photo of the sample circuit traces. The cantilever is extended over the traces while the tip is located the middle of the trace 0. (b) The dimensions for the circuit traces.

The measured force and force gradient signals as a function of the tip-to-sample distance are shown in Fig. 7.16.

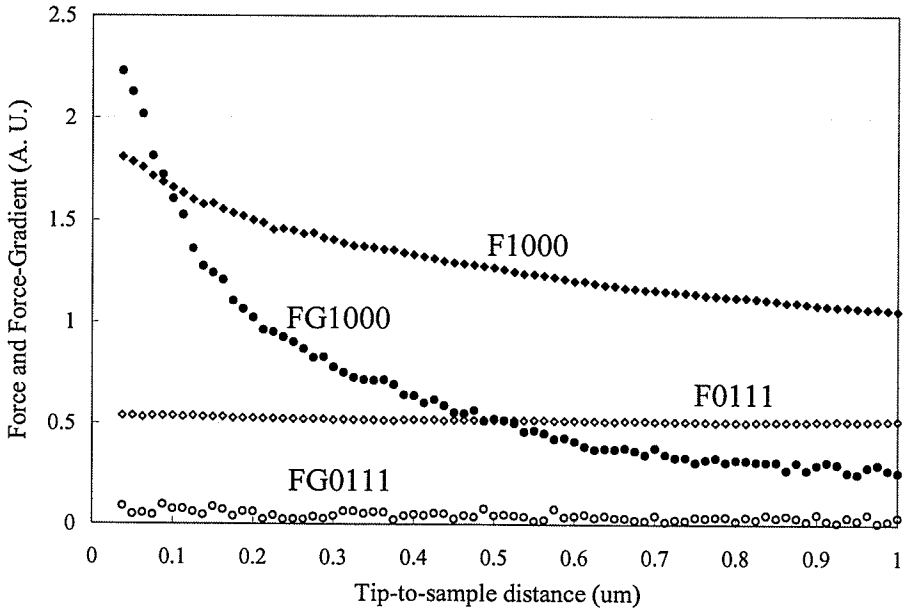

Fig. 7.16. Solid dots are measured force (F1000) and force gradient (FG1000) for V_{c0} = 1.3 V and $V_{c1} = V_{c2} = V_{c3} = 0$. Hollow dots are measured force (F0111) and force gradient (FG0111) for V_{c0} = 0 V and $V_{c1} = V_{c2} = V_{c3} = 1.3$ V.

Figure 7.16 shows that the large contribution from the cantilever in the force measurement (F0111/F1000) is significantly reduced in the force gradient measurement (FG0111/ FG1000). However the force gradient measurement has to be performed within a much smaller tip-to-sample distance to increase the measurement signal-to-noise-ratio.

Figure 7.17 shows the measured force and force gradient signals as a function of the circuit voltage (V_{c0}). The result shows that the force gradient is proportional to the measured circuit voltage. The nulling technique can also be used for obtaining the absolute voltage measurements and the spatial resolution can be significantly improved in comparison with the force measurement as indicated in Fig. 7.17. Further discussion will be presented in [101].

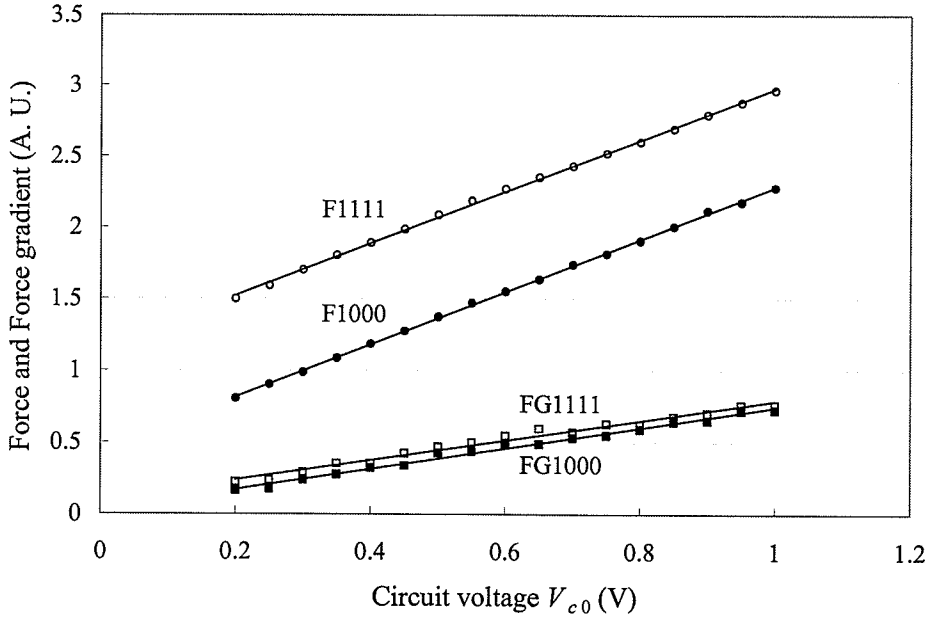

Fig. 7.17. Solid dots are measured force (F1000) and force gradient (FG1000) for $V_{c1} = V_{c2} = V_{c3} = 0$. Hollow dots are measured force (F1111) and force gradient (FG1111) for $V_{c1} = V_{c2} = V_{c3} = 2$ V.

In summary, the force gradient detection method can be applied to improve the measurement spatial resolution, and more importantly, to dramatically suppress the interference from signals on parallel interconnects.

Chapter 8

Conclusions and Future Considerations

8.1 Conclusions

Several electrostatic force microscopy techniques have been developed and shown to be useful tools for the sampling of high-speed internal circuit signal waveforms. In this work, frequency downconversion is mainly realized by driving the SPM probe with an amplitude modulated voltage pulse. Using ultrafast logic gates for generating high-speed sampling pulses, this EFM instrument has demonstrated a bandwidth of 8 GHz and a phase resolution of 10 ps. Up to 10 Gbit/s digital signals were measured on a wideband distributed amplifier to demonstrate the capability of the instrument. A technique for high-speed quantitative voltage measurement was realized by using a null-force technique. Quantitative voltage measurements of a wideband distributed amplifier's 250 Mbit/s internal signals was demonstrated with a voltage sensitivity of approximately 5 $mV/\sqrt{Hz}$. The spatial resolution of the standard EFM technique, using a commercial micro-fabricated probe, was estimated to be about 1.4 μm (used in ambient conditions). In addition, large interference from signals on adjacent interconnect traces were observed. Two dynamic force-gradient detection methods were proposed and implemented to improve the probe spatial resolution. A submicron spatial resolution and a significant reduction of interference from signals on parallel interconnects was achieved in the force-gradient measurements of signals up to 500 Mb/s. The cantilever loading was found to be about 4 fF for a tip-to-sample spacing of about 300 nm.

8.2 Future considerations

Due to the increasing circuit operation speed, further improvement of the EFM probe bandwidth is needed for the near future. The Picosecond edge compressor's 4 ps voltage

edge and a CPW transmission line cantilever could be applied to allow for 10 Gbit/s measurements.

DEFM force gradient detection has shown a significant improvement of the measurement spatial resolution. However, its sensitivity is usually much worse than the force measurement. To improve the sensitivity of the force gradient detection, a vacuum system should be applied.

References

[1] Semiconductor Industry Association, "2001 International Technology Roadmap for Semiconductors", http://www. semichips.org.

[2] "Probe Tutorial", Application Notes, Tektronix. (http://www.tek.com).

[3] G. Rabjohn, J. Wolczanski, and R. Surridge, "High-frequency wafer-probing techniques," Canadian J. Phys. **65**, 850, (1987).

[4] GGB Industries Inc., Naples, Florida, USA. (http://www.picoprobe.com).

[5] K. Krieg, R. Qi, D. Thomson, and G. Bridges, "Electrical Probing of Deep Sub-Micron ICs Using Scanning Probes," 38th Annual IEEE International Reliability Physics Symposium, Santa Jose, California, (2000).

[6] MikroMasch, USA. http://www.spmtips.com/.

[7] K. Yang, L. P. B. Katehi, and J. F. Whitaker, "Electro-optic field mapping system utilizing external gallium arsenide probes," Appl. Phys. Lett. **77**, 486, (2000).

[8] Electron Beam Testing Technology, edited by J. T. L. Thong (Plenum, New York, 1993).

[9] S. Kasapi, C. C. Tsao, K. Wilsher, W. Lo, and S. Somani, "Laser beam backside probing of CMOS integrated circuits," Microelectron. Reliab. **39**, 957 (1999).

[10] J. C. Tsang, J. A. Kash and D. P. Vallett, "Optical tools for measuring timing and switching in silicon ICs: status and future challenges," LEOS Newsletter, **15** (2), 3, (2001).

[11] Y. Martin and H. K. Wickramasinghe, "Magnetic imaging by 'Force Microscopy' with 1000 Angstrom resolution," Appl. Phys. Lett. **50**, 1455, (1987).

[12] A. S. Hou, F. Ho, and D. M. Bloom, "Picosecond electrical sampling using a scanning force microscope," Electron. Lett. **28**, 2302, (1992).

[13] G. E. Bridges, R. A. Said, and D. J. Thomson, "Heterodyne electrostatic force microscopy for non-contact high frequency integrated circuit measurement," Electron. Lett. **29**, 1448, (1993).

[14] C. Bohm, F. Saurenbach, P. Taschner, C. Roths, and E. Kubalek, "Voltage Contrast in Integrated Circuits with 100 nm Spatial Resolution by Scanning Force Microscopy," J. Phys. D.:Appl. Phys., **26**, 1801, (1993).

[15] G. M. Steeves, A. Y. Elezzabi, and M. R. Freeman, "Advances in picosecond scanning tunneling microscopy via junction mixing," Appl. Phys. Lett. **70**, 1909 (1997).

[16] J. Kim, S. Williamson, J. Nees, S. Wakana, and J. Whitaker, "Photoconductive sampling probe with 2.3-ps temporal resolution and 4-μV sensitivity," Appl. Phys. Lett. **62**, 2268, (1993).

[17] J. A. Valdmanis and G. Mourou, "Subpicosecond electro-optic sampling: principles and applications," IEEE J. Quantum Electron. **22**, 69, (1986).

[18] Hironori Takahashi, Shin-ichiro Aoshima and Yutaka Tsuchiya, "Sampling and real-time methods in electro-optic probing systems", IEEE transactions on instrumentation and measurement, vol. 44, no. 5, Oct 1995.

[19] W. Mertin, "Two-dimensional field mapping of monolithic microwave integrated circuits using electro-optic sampling techniques," Opt. and Quantum Electron. **28**, 801, (1996).

[20] K. Yang, D. David, S. Robertson, J. F. Whitaker and L. P. B. Katehi, "High resolution electro-optic mapping of near-field distributions in integrated microwave circuits", IEEE Microwave Theory and Techniques Digest, 949, (1998).

[21] W. Mertin, A. Leyk, U. Behnke, and V. Wittpahl, "Contactless gigahertz testing," Proceedings of the 1998 IEEE International Test Conference, 843, (1998).

[22] IDS 10000da, E-beam probe system for IC design analysis.

[23] J. Giesler, G. O'Malley, M. Williams, and S. Machuga, "Flip chip on board connection technology: Process characterization and reliability," IEEE Trans. Components, Packaging, and Manufacturing Technology B, **17**, 256, (1994).

[24] H. Heinrich, " Picosecond non-invasive optical detection of internal electrical signals in flip-chip-mounted silicon integrated circuits," IBM J. Res. Dev. **34**, 162, (1990).

[25] W. Lo, S. Kasapi, and K. Wilsher, "Comparison of laser and emission based optical probe techniques," Proceedings for the 27th International Symposium for Testing and Failure Analysis, **33** (2001).

[26] IDS PICA, Advanced optical imaging for analysis of 0.13-micron and SOI devices, http://www.idspica.com.

[27] D. Sarid, Scanning Force Microscopy with Applications to Electric, Magnetic and Atomic Forces, Oxford University Press, (1991).

[28] G. Binnig, C. F. Quate, and C. Gerber, "Surface Studies by Scanning Tunnelling Microscopy," Phys. Rev. Lett. **49**, 57, (1982).

[29] B. T. Rosner, D. W. van der Weide, "High-frequency near-field microscopy", Rev. Sci. Instrum., **73 (7)**, 2505, (2002).

[30] N. A. Burnham, R. J. Colton, H. M. Pollock, "Interpretation of Force Curves in Force Microscopy," Nanotechnology, **4**, 64, (1993).

[31] W. Mertin, R. Weber, F. Seifert, E. Kubalek, G. Zimmermann, and C. Boit, "Contactless failure analysis of integrated circuits via current contrast imaging with magnetic force microscopy," Proceedings from the 27th International Symposium for Testing and Failure Analysis, 199, November, (2001).

[32] Digital Instruments, Santa Barbara, USA. (http://www.di.com).

[33] M. Abea and Y. Tanaka, "High frequency write head measurement with the phase detection magnetic force microscope," J. App. Phys. **89** (11), 6766, (2001).

[34] R. Proksch, P. Neilson, S. Austvold, and J. J. Schmidt, "Measuring the gigahertz response of recording heads with the magnetic force microscope," Appl. Phys. Lett. **74**, 1308, (1999).

[35] G. M. Steeves, A. Y. Elezzabi, R. Teshima, R. A. Said, and M. R. Freeman, "Circuit analysis of an ultrafast junction mixing scanning tunnelling microscope," IEEE J. Quantum Electron **34**, 1415 (1998).

[36] J. M. R. Weaver and D. W. Abraham, "High Resolution Atomic Force Microscopy Potentiometry," J. Vac. Sci. Technol. B, **9**, 1559, (1991).

[37] M. Nonnenmacher, M. P. O'Boyle, and H. K. Wickramasinghe, "Surface investigation with a Kelvin Probe Force Microscope," Ultramicroscopy, **42-44**, 268, (1992).

[38] R. A. Said, G. E. Bridges, and D. J. Thomson, "Scanned Electrostatic Force Microscopy for non-invasive high frequency potential measurement," Appl. Phys. Lett., **64**, 1442, (1994).

[39] A. S. Hou, B. A. Nechay, F. Ho, and D.M. Bloom, "Scanning Probe Microscopy for Testing Ultrafast Electronic Devices," Opt. Quantum Electron. **28**, 819, (1996).

[40] G. M. Amer, "Novel optical approach to Atomic Force Microscopy," Appl. Phys. Lett. **53**, 2400, (1988).

[41] NT-MDT Co., CSCS12/W2C data sheet.

[42] S. Rast, C. Wattinger, U. Gysin, and E. Meyer, "Dynamics of damped cantilevers," Rev. Sci. Instru. **71**, 2772, (2000).

[43] U. Rabe, K. Janser, and W. Arnold, "Vibrations of free and surface-coupled atomic force microscope cantilevers: Theory and experiment," Rev. Sci. Instrum. **67**, 3281, (1996).

[44] R. A. Said, "Scanning Force Potentiometry Techniques for Semiconductor Circuit Characterization," Ph. D. Thesis, University of Manitoba, 1995.

[45] S. P. S. Cheung, "Electrostatic Force Sampling of Digital Waveforms using Synchronous Time Domain Gating," Master Thesis, University of Manitoba, (2000).

[46] C. A. J. Putman, B. G. De Grooth, and N. F. Van Hulst J. Greve. "A detailed analysis of the optical beam deflection technique for use in atomic force microscopy," J. Appl. Phys. **72**, 6, (1992).

[47] Micron-Force Instruments, San Jose, California, U.S.A.

[48] SR510 Lock-in Amplifier manual, Stanford Research System.

[49] C. Falkingham, "Vector Voltage Measurement of RFICs using Electrostatic Force Microscopy," Master Thesis, University of Manitoba, 2001.

[50] G. E. Bridges, "Vector-voltage Scanning Force Probe for Non-contact MMIC Measurement," Electron. Lett. , vol. 35, pp. 1724-1725, 1999.

[51] Z. Weng, G. Bridges, and D. Thomson, "Internal probing of 10 Gbit/s ICs," Micronet Annual Workshop, (2002).

[52] Tektronix, http://www.tek.com; Agilent, http://www.agilent.com.

[53] G.E. Bridges, R. A. Said, M. Mittal, and D.J. Thomson, "Sampled waveform measurement in integrated circuits using heterodyne electrostatic force microscopy," Rev. Sci. Instrum., **65**, 3378, (1994).

[54] R. A. Said, S. P. Cheung, and G. E. Bridges, "High resolution sampling electrostatic force microscopy using pulse width modulation technique," J. Vac. Sci. Technol. B, **18**, 626, (2000).

[55] R. A. Said, "Enhanced pulse-sampling force probing for picosecond measurement of integrated circuits," J. Phys. D: Appl. Phys. **34**, L26, (2001).

[56] Hewlett-Packard, Germany, HP80000 Data Generation System Manual, (1993).

[57] NTT Electronics Technology Corporation, Tokyo, Japan, AND/NAND Gate data sheet. (www.nel-work.com)

[58] D.W. van der Weide, "Delta-doped Schottky diode nonlinear transmission lines for 480-fs, 3.5-V transients," Applied Physics Letters, **65**, 881, (1994).

[59] M. R. T. Tan, S. Y. Wang, D. E. Mars, and J. L. Moll, "A 12 psec GaAs double heterostructure step recovery diode," Hewlett-Packard Lab., HPL-91-187, (1991).

[60] Picosecond Pulse Labs, Colorado, USA. http://www.picosecond.com/.

[61] G. E. Bridges and D. J. Thomson, "High-frequency Circuit Characterization using the AFM as a Reactive Near-field Probe," Ultramicroscopy, **42-44**, 321, (1992).

[62] TDS 8000 Digital Sampling Oscilloscope Manual.

[63] B.Q. Liang, Z. Weng, G.E. Bridges, C. Shafai and D.J. Thomson, "Matched Impedance Cantilever for High Frequency Scanning Probe Microscopy," CMC MEMS workshop, 2004.

[64] B. Liang, "Stressed MEMS CPW Cantilevers for High Frequency Scanning Probe Microscopy," Master Thesis, University of Manitoba, 2004.

[65] LabVIEW, National Instruments, Austin, Texas, USA. (www.ni.com).

[66] Z. Weng, G. E. Bridges and D. J. Thomson, "Probe Position Control for Enhanced Resolution of Electrostatic Force Microscopy," Proceedings of the 2002 IEEE Canadian Conference on Electrical & Computer Engineering, 378, (2002).

[67] G. E. Bridges, D. J. Thomson, and R. Qi, "Non-contact probing of integrated circuits using electrostatic force sampling," Proceedings from the 24th international Symposium for Testing and Failure Analysis, Dallas, (1998).

[68] Z. Weng and G. Bridges, "Non-contact electrostatic force probing of Nortel NT GB98 wideband distributed amplifier," test report, (2002).

[69] C. Mittermayer and A. Steininger, "On the Determination of Dynamic Errors for Rise Time Measurement with an Oscilloscope," IEEE Trans. on Instru. and Meas., **48**, 1103, 1999.

[70] S. Roy and M. Souders, "Noniterative Waveform Deconvolution Using Analytic Reconstruction Filters with Time-Domain Weighting," IEEE Trans. Instrum. Meas., **50**, 899, (2001).

[71] N. S. Nahman and M. E. Guillaume, "Deconvolution of time domain waveforms in the presence of noise," NBS Tech. Note 1047, (1981).

[72] LeCroy Cooporation, Chestnut Ridge, New York, USA, Application Notes. (http://www.lecroy.com).

[73] S. Rast, C. Wattinger, U. Gysin, and E. Meyer, "The Noise of Cantilevers," Nanotechnology, **11**, 169, (2000).

[74] Nyquist's theorem : an analog signal waveform may be uniquely reconstructed, with error, from samples taken at equal time intervals. The sampling rate must be equal to, or greater than, twice the highest frequency component in the analog signal.

[75] F. Du, G. Bridges, D. J. Thomson, R. R. Goruganthu, S. Mcbride, and M. Santana Jr, "Enhancements of Non-contact Measurements of electrical waveforms on the proximity of a signal surface using groups of pulses," to be published.

[76] G. E. Bridges, "Vector voltage scanning force probe for non-contact MMIC measurement," Electron. Lett. **35**, 1724, (1999).

[77] G. E. Bridges, R. A. Said, M. Mittal, and D. J. Thomson, "High-frequency pattern extraction in digital integrated circuits using scanning electrostatic force microscopy," J. Vac. Sci. Technol. A **13**, 1375, (1995).

[78] J. Bangert and E. Kubalek, "Absolute quantitative time resolved voltage measurements within today's integrated circuits via Electric Force Microscopy-(EFM)-testing," Microelectron. Reliab. **37**, 1579, (1997).

[79] Z. Weng, C. J. Falkingham, G. E. Bridges, and D. J. Thomson, "Quantitative Voltage Measurement of High-frequency Internal Integrated Circuit Signals by Scanning Probe Microscopy," J. Vac. Sci. Technol. A, **20**, 999, (2002).

[80] S. Gomez-Monivas, L. S. Froufe, R. Carminati, J. J. Greffet and J. J. Saenz, "Tip-shape effects on electrostatic force microscopy resolution," Nanotechnology, **12**, 496, (2001).

[81] J. Colchero, A. Gil, and A. M. Baro, "Resolution enhancement and improved data interpretation in electrostatic force microscopy," Phys. Rev. B **64**, 245403-1, (2001).

[82] S. Belaidi, P. Girard, and G. Leveque, "Electrostatic forces acting on the tip in atomic force microscopy: Modelization and comparison with analytic expressions," J. Appl. Phys. **81**, 1023 (1997).

[83] O. Cherniavskaya, L. Chen, V. Weng, L. Yuditsky, and L. E. Brus, "Quantitative noncontact electrostatic force imaging of nanocrystal polarizablility," J. Phys. Chem. B **107**, 1525, (2003).

[84] G. Koley, M. G. Spencer and H. R. Bhangale, "Cantilever effects on the measurement of electrostatic potentials by scanning Kelvin probe microscopy," Appl. Phys. Lett. **79**, 545, (2001).

[85] V. Wittpahl, U. Behnke, B. Wand and W. Mertin, "Cantilever influence suppression of contactless IC-testing by electric force microscopy," Microelectron. Reliab. **38**, 981, (1998).

[86] H. Richter, "Test chip for the characterization of electron beam measurement systems," Microelectron. Eng. **16**, 225 (1992).

[87] S. Belaidi, F. Lebon, P. Girard, G. Leveque and S. Pagano, "Finite element simulations of the resolution in electrostatic force microscopy," Appl. Phys. A **66**, S239, (1999).

[88] T. Kaminski, "Coupled Electrostatic-Mechanical Analysis of Probes used in Dynamic EFM," Master Thesis, University of Manitoba, (2004).

[89] R. F. Harrington, "Field Computation by Moment Methods," The Macmillan Company, London, 1968.

[90] Z. Weng, T. Kaminski, G. E. Bridges, and D. J. Thomson, "Resolution enhancement in probing of high-speed integrated circuits using dynamic electrostatic force-gradient microscopy," J. Vac. Sci. Technol. A **22**, 1, (2004).

[91] P. Girard, M. Ramonda, and D. Saluel, "Electrical contrast observations and voltage measurements by Kelvin probe force gradient microscopy," J. Vac. Sci. Technol. B **20**(4), 1348, (2002).

[92] A. Gil, J. Colchero, J. Gomez-Herrero and A. M. Baro, "Electrostatic force gradient signal: resolution enhancement in electrostatic force microscopy and improved Kelvin probe microscopy," Nanotechnology **14**, 332, (2003).

[93] C. H. Lei, A. Das, M. Elliott and J. E. Macdonald, "Quantitative electrostatic force microscopy-phase measurements," Nanotechnology **15**, 627, (2004).

[94] Y. Martin, C. C. Williams and H. K. Wickramasinghe, "Atomic Force Microscope-Force Mapping and Profiling on a Sub 100-A scale," J. Appl. Phys. **61**, 4723, (1987).

[95] S. Watanabe, K. Hane, M. Ito, and T. Goto, "Dynamic mode force microscopy for the detection of lateral and vertical electrostatic forces," Appl. Phys. Lett. **63**, 2573, (1993).

[96] T. R. Albrecht, P. Grutter, D. Horne, D. Rugar, "Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity," J. Appl. Phys., **69**, 668, (1991).

[97] R. A. Said, "Perturbation detection of electric force gradients using the phase shift method," J. Phys. D: Appl. Phys. **34**, L7, (2001).

[98] A. Kikukawa, S. Hosaka, and R. Imura, "Silicon *pn* junction imaging and characterizations using sensitivity enhanced Kelvin probe force microscopy," Appl. Phys. Lett. **66**(25), 3510, (1995).

[99] H. Yokoyama, T. Inoue and J. Itoh, "Nonresonant detection of electric force gradients by dynamic force microscopy," Appl. Phys. Lett. **65**(24), 3143, (1994).

[100] THORLABS INC, piezo-electric multi-layer actuator, AE0203D04 data sheet.

[101] Z. Weng, T. Kaminski, G. E. Bridges, and D. J. Thomson, "Dynamic electrostatic force-gradient microscopy employing mechanoelectric cross-modulation," submitted to J. Vac. Sci. Technol. A (published at 12th Canadian Semiconductor Technology Conference, Ottawa, 2005, TH4.2).