The Parametric Measurement Handbook

Third Edition
March 2012

$$Vd = Vg - Vt$$
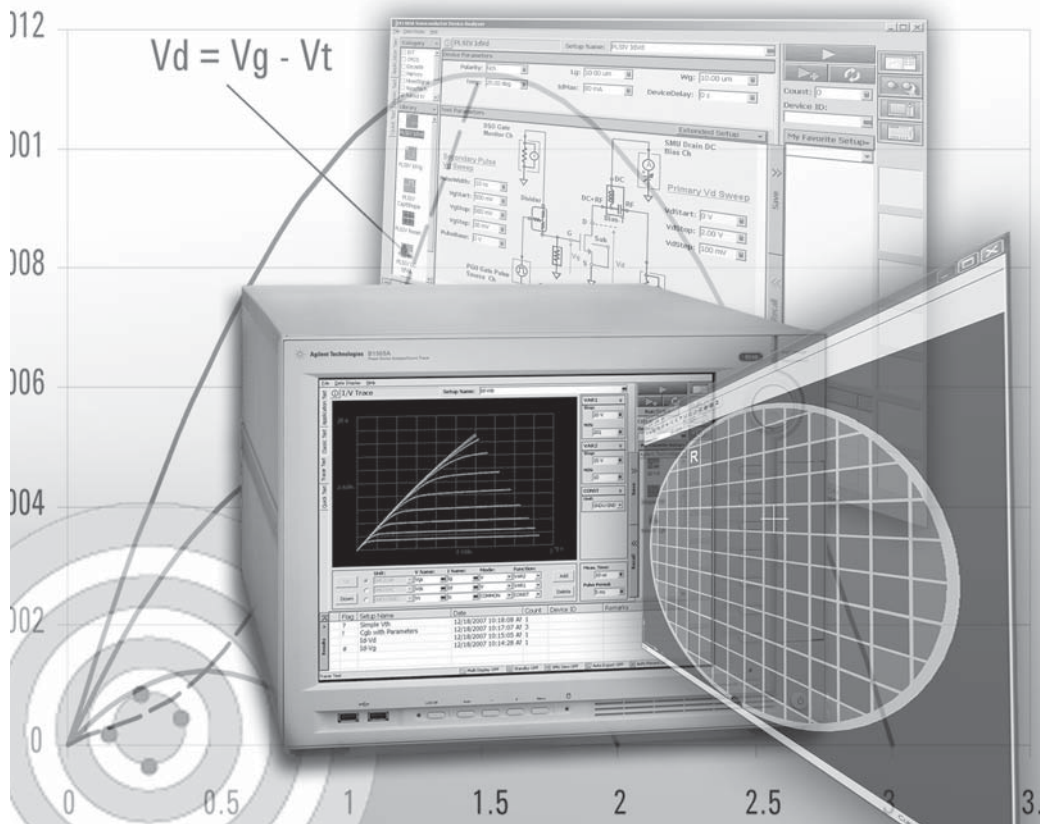
Agilent Technologies

Table of Contents

Table of Contents *continued...*

Table of Contents *continued...*

Preface

"An expert is a man who has made all the mistakes, which can be made, in a very narrow field" — Niels Bohr

When I first received the request to create this handbook I cringed at the scope involved. How could I possibly encapsulate all of parametric test into a single text? Fortunately I have had a great deal of help and support from a number of people. I would especially like to acknowledge the contributions of the following individuals:

- *Tomoyuki Akiyama* (Application Specialist, Agilent Technologies,) for the numerous and informative application notes he generated over the years.
- *Terry Burcham* (Application Specialist, Cascade Microtech) for his material on on-wafer parametric measurement from which I freely borrowed.
- *Yasushi Okawa* (Application Specialist, Agilent Technologies) for the insight he supplied to me on high-speed and fast pulsed measurement.
- *Bill Verzi* (Application Engineer, Agilent Technologies) for the knowledge and help he supplied on the measurement of thin-gate dielectric capacitors.

I would also like to thank all those that helped to proof read this document. Although every effort has been made to eliminate errors, inconsistencies and awkward language, no doubt some still remain to amuse the careful reader.

It is often said that you really do not understand a subject until you have to teach it to someone else. The same could easily be said for writing a measurement handbook. Being forced to put the accumulated measurement experience of many people into a few choice words compelled me to reexamine my own knowledge and make sure that I really did understand everything that I was talking about. Although this process was a lot of work it also was a certain amount of fun. I hope that you enjoy reading the final product as much as I enjoyed creating it.

Alan Wadsworth
March 2012

Chapter 1: Parametric Test Basics

"The central activity of engineering, as distinguished from science, is the design of new devices, processes and systems." — Myron Tribus

What is parametric test?

The question as to what constitutes parametric test is an interesting one and is possibly open to some debate. Nevertheless, in general parametric test involves the electrical testing and characterization of four main types of semiconductor devices: resistors, diodes, transistors, and capacitors. This is not to say that parametric test never involves the testing of other device types; however, the vast majority of parametric test structures can be classified into one of these categories or considered to be a combination of these categories.

| Transistors | Diodes | Resistors | Capacitors |
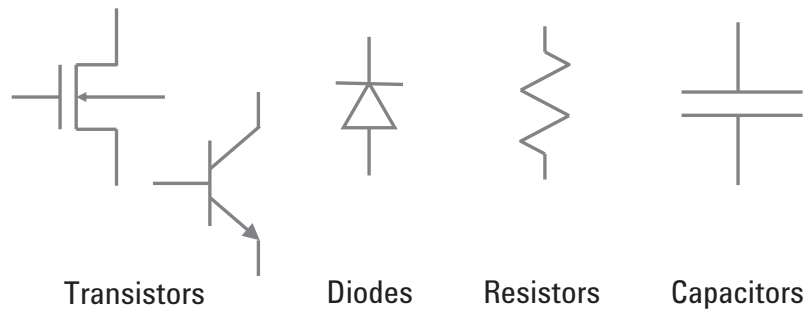
Figure 1.1. Parametric test involves the testing of these four basic device types.

The vast majority of parametric testing involves either current versus voltage (IV) or capacitance versus voltage (CV) measurements.

To many people parametric test means "DC" testing, but this is not an accurate description. Of course, it can take source/monitor units (SMUs) anywhere from milliseconds to seconds to make a measurement, which is certainly "slow" by the standards of functional testers (which typically perform measurements in the nanosecond or picosecond range). However, in recent years the need to perform extremely fast parametric measurements (1 μs spot measurements with data sampling rates in the nanosecond range) has greatly increased. This has required the creation of new measurement module types (such as the waveform generator/fast measurement unit or WGFMU) to meet this need. Extremely fast IV and pulsed IV measurements will continue to take on increased importance in the future, as transistor lithographies continue to shrink and more exotic materials are incorporated into semiconductor processes.

One major subcategory of parametric test is reliability testing. Reliability testing relies heavily upon the well-known Arrhenius equation that expresses the rate constant (k) of a chemical reaction as follows:

$$k = Ae^{-\frac{E_a}{RT}}$$

Where:
A is the pre-exponential factor
Ea is the activation energy
R is the ideal gas constant
T is the temperature in degrees Kelvin

In reliability testing devices are typically stressed by the application of large currents and/or voltages (larger than that experienced by the devices under normal operation) to lower the value of the activation energy and thereby increase the rate of the failure mechanism. Temperature is also often increased to achieve this same purpose. Once the failure mechanism occurs then the expected failure rate under normal operating conditions can be extrapolated using a variety of mathematical and statistical techniques.

Why is parametric test performed?

The purpose of parametric test is to determine the characteristics of a semiconductor manufacturing process. Broadly speaking, parametric test covers three main areas: process development, process modeling, and process production.

Process development   Device modeling   Process control
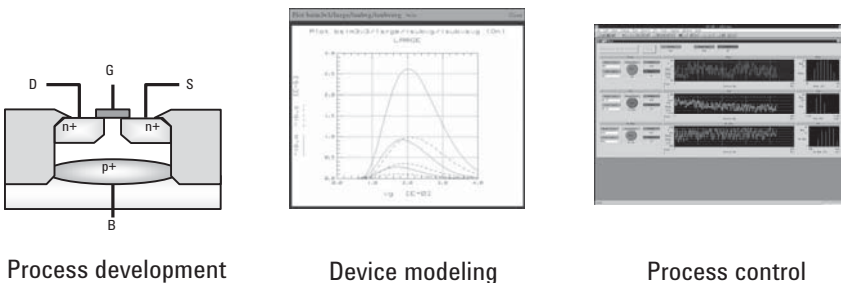
Figure 1.2. Parametric test focuses in three main areas: process development, device modeling, and process control.

The first two of these areas are performed in a laboratory, R&D or pre-production environment, while the last one is obviously performed in a manufacturing environment. The parametric equipment used in these disparate environments obviously has different requirements, with benchtop instruments being used for process development and process modeling and high-throughput testers being used for process production.

Figure 1.3. A production parametric tester is designed to optimize throughput

It is important to understand that parametric test is almost never performed on final products. Instead it is performed on special structures that are designed to yield information about the process itself. Parametric test is also generally performed directly on semiconductor wafers. In production test the parametric test structures are sometimes located in the scribe lanes or "streets" of the wafer to minimize the wafer area taken up by these devices. However, for process development and reliability testing entire wafers of nothing but parametric test structures are often fabricated.

Functional testers test product die
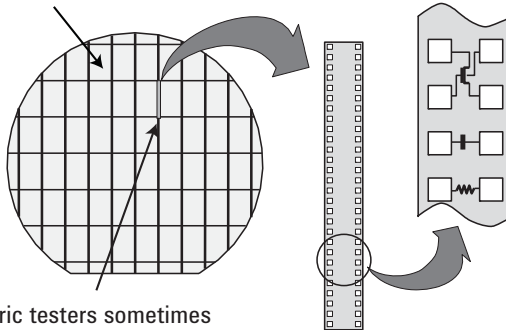
Parametric testers sometimes test special structures in the scribe lane

Figure 1.4. To conserve valuable wafer area, parametric test structures are sometimes placed in the wafer's scribe lines (or "streets").

Where is parametric test done?

In production, parametric test is typically performed on the wafers after they have completed the wafer fabrication process (i.e. after passivation has been applied) but before electrical sort (E-sort) on the functional product dice.
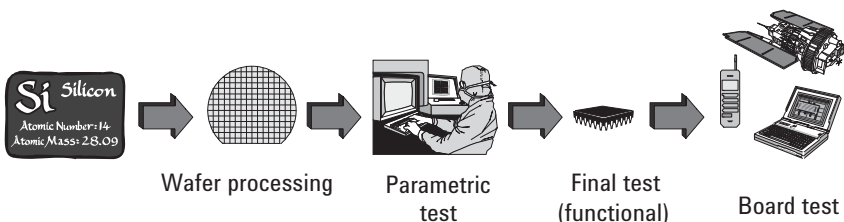
Figure 1.5. Parametric test performed after wafer fabrication is complete but before functional product verification.

Each wafer from every lot is tested and the data is stored into a database. Obviously, the amount of data is quite massive, and various software tools are employed to manipulate the data into a variety of different formats. One popular format is the wafer map, where a scalar quantity is plotted across a wafer using different colors for different ranges of the data value.
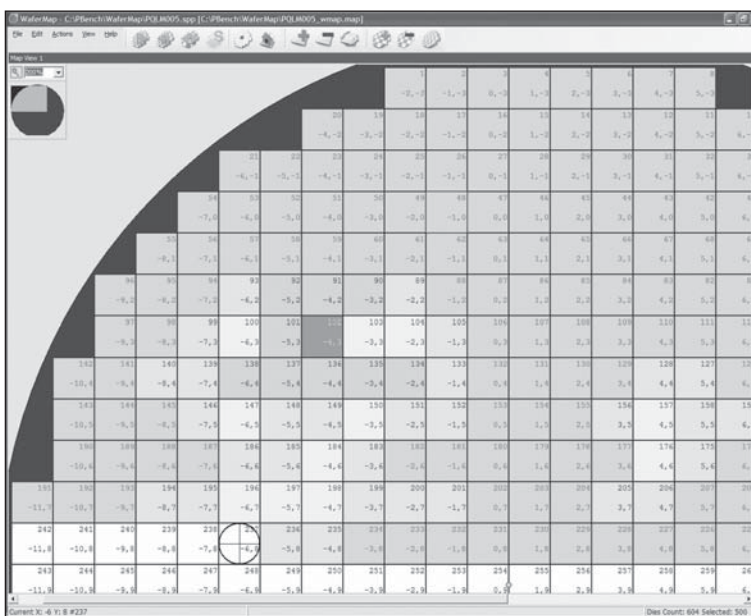
Figure 1.6. Wafer map example.

For advanced processes, conventional test structures placed in the scribe lines or even drop-in test die placed around the wafer may not be sufficient to adequately characterize the process. Advanced processes typically require much more testing due to their innate complexity, and it is sometimes difficult to fit all of the necessary test structures into the available area. Unfortunately, the tips of the probe card have physical limitations on how small they can be made and still maintain reasonable probe card lifetimes, which in-turn creates physical limits on the minimum size of the probe pads. This means that the pads cannot physically scale down as devices scale down with each new process generation.
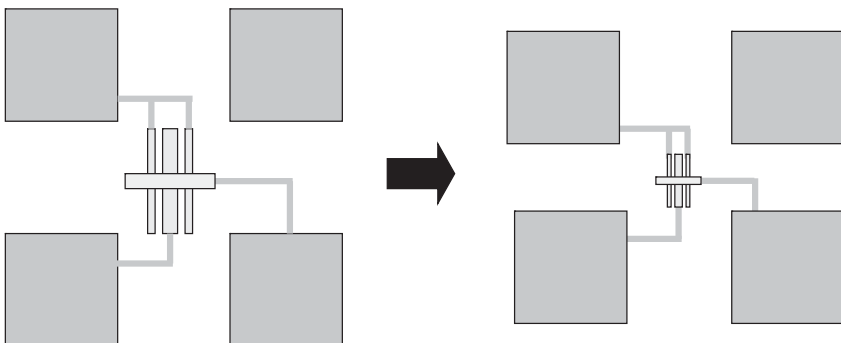
Figure 1.7. Pad size cannot scale with design rule changes, which limits the number of conventional test structures that can be placed on a wafer.

One solution to this issue is to use arrays, since an array allows test devices to share pads and thereby improve the test device to pad ratio. An example of this scheme is shown below.
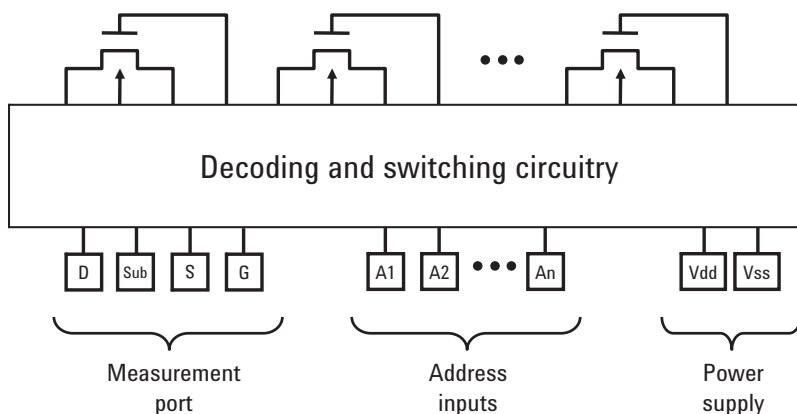
Figure 1.8. An example of an addressable array scheme used for production parametric test.

Addressable arrays can offer significant throughput improvement over conventional test device arrangements, although achieving substantial test time reduction requires a complete re-engineering of all of the parametric test processes.

Parametric instrument history

The first instruments that could perform anything resembling parametric measurement were analog curve tracers. However, these instruments had a number of limitations not the least of which was that the "data" was a display on a CRT screen. The only way to save this data was to take a Polaroid picture of the display, which still did not supply you with actual numeric information.

The 1980's

In the early 1980's Agilent Technologies (then part of Hewlett Packard) introduced the world's first digital parameter analyzer, the 4145A. The 4145A was the first instrument to incorporate four source/monitor units (to be discussed further in Chapter 3) in a single box along with the software necessary to integrate all of these resources. The 4145A produced plots similar to that of a curve tracer, but the 4145 plot consisted of discrete points of digital data that could be transferred into other software for storage and analysis. This product was something of a revolution for the semiconductor industry, and the 4145A quickly supplanted the curve tracer except for a few specialized applications. Later in the 1980's an improved version of this product, the 4145B, was also introduced.

Figure 1.9. The 4145B Semiconductor Parameter Analyzer

The 1990's

In the early 1990's Agilent Technologies (then part of Hewlett Packard) introduced the 4155 and 4156 semiconductor parameter analyzers. These instruments built upon the 4145A/B by adding a new capabilities including a color display and keyboard, automatic data analysis capability, a pulsed sweep capability, a thinned-out sampling mode, DC and AC stressing and a standby mode (to name just a few). There were several versions of these products (the "A", "B" and "C" models), with each subsequent version adding additional features and capabilities. In particular, the "C" version added quasi-static capacitance versus voltage (QSCV) measurement capability as well as vastly improved voltage monitor unit (VMU) measurement capability.

Figure 1.10. The 4156C Precision Semiconductor Parameter Analyzer

The 2000's

At the beginning of the 21st century it became increasingly clear that parametric test was becoming more complicated and that a complete solution needed to be able to perform more than just IV measurements. A solution that could perform both CV and IV measurements was needed, and this solution also needed the flexibility to add additional measurement resources in the future as testing needs evolved. To meet these challenges, Agilent Technologies introduced the B1500A semiconductor device analyzer with Agilent EasyEXPERT software in 2005. The B1500A supports all aspects of parametric test, from basic manual measurements to test automation across a wafer in conjunction with a semiautomatic wafer prober. Because the B1500A utilizes the Microsoft® Windows® XP Professional operating system, it integrates easily into PC-based work environments. In addition, the familiar Windows graphical user interface (GUI) and convenient online help menus minimize the need for instrument training. The B1500A has already supplanted the 4155C and 4156C as the tool of choice for state-of-the-art parametric measurement needs.

Figure 1.11. The B1500A Semiconductor Device Analyzer

The B1500A is a modular instrument, and it supports a variety of different module types. Modularity is an extremely important feature since parametric test continues to be an extremely dynamic area where new types of testing are constantly being required. Parametric test requires more than simple source/monitor unit resources, and increasingly complex tests requiring capacitance testing, high-speed testing and fast pulsed measurement are becoming common. Modularity insures that measurement equipment does not become outdated and that you have the ability to add new test capabilities to your parametric measurement equipment as your measurement needs change in the future.
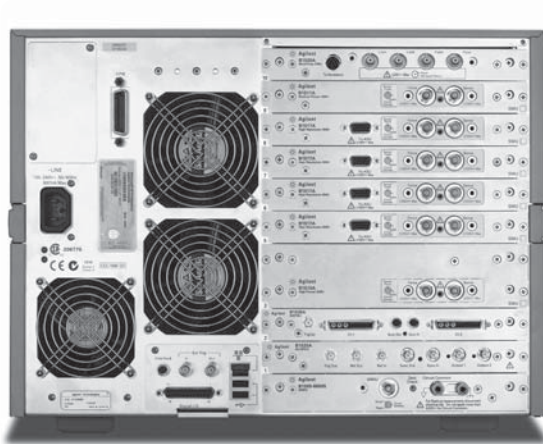
Figure 1.12. The ten-slot B1500A supports a variety of types of source/monitor units as well as a capacitance measurement module, a high-voltage pulse generator module and a waveform generator/fast measurement unit module.

To meet the needs of the expanding high-power market and to supply an alternative to conventional analog curve tracers (which are no longer being made by major manufacturers), Agilent introduced the B1505A Power Device Analyzer/Curve Tracer in 2009. The B1505A can supply up to ±20 A (pulsed) and ±3000 V. In addition, the B1505A has a tracer test mode that supports a knob sweep capability.
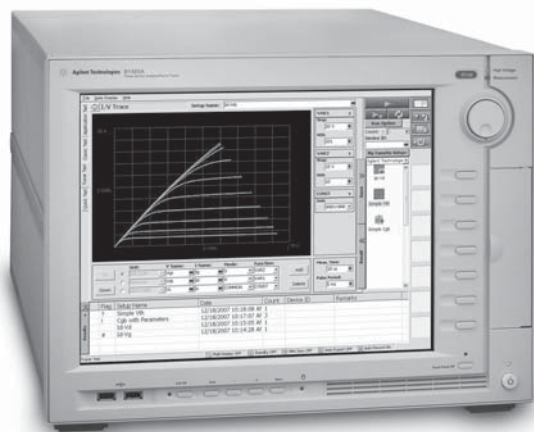
Figure 1.13. The B1505A Power Device Analyzer/Curve Tracer.

Unlike analog curve tracers, the B1505A is PC-based and uses the same user-friendly Agilent EasyEXPERT software as the B1500A.

Several new features in the B1505A's tracer test mode provide dramatic usability improvements over traditional curve tracers. A snapshot feature allows you to save and display multiple data traces so that you can easily compare them with data from the current measurement. A stoplight feature allows you to graphically define forbidden regions (either voltage or current based) such that the measurement immediately ceases if the trace enters the forbidden area. Best of all, an auto-record feature keeps a running record of the most recent trace changes so that you can replay and save measurement trace data even if your device is inadvertently damaged or destroyed. Taken together these improvements represent a truly revolutionary advance in curve tracer design that can significantly reduce device characterization cycle times.

Chapter 2: Parametric Measurement Basics

"I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it ..."

"But when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be."
— William Thomson (Lord Kelvin)

Measurement terminology

Before you can begin to collect measurement data you need to know basic measurement terminology. In particular, understanding the precise meaning of the terms accuracy, repeatability and resolution is essential to understanding parametric test and the measurement capabilities of a parametric measurement resource.

Accuracy and repeatability
Accuracy and repeatability are related but different quantities. Accuracy is the degree of conformity of a measured or calculated quantity to its actual (true) value. Repeatability (also known as precision) is the degree to which repeated measurements or calculations show the same or similar results. Accuracy and repeatability do not have any innate correlation. A measurement can have high accuracy and high repeatability, high accuracy and low repeatability, low accuracy and high repeatability, or both low accuracy and low repeatability.

The visual analogy that is often used to explain these terms is a bull's-eye target. A group of points that are close to the center of a target but spaced far apart from one another (as shown below) have high accuracy but low repeatability.
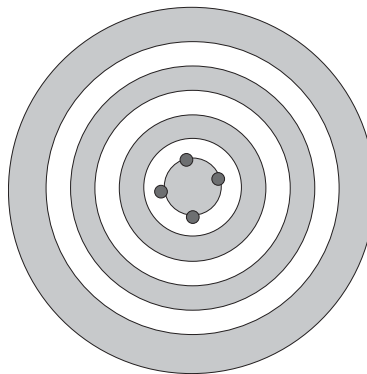
High accuracy, low repeatability

Figure 2.1. High accuracy and low repeatability.

A group of points that are far from the center of a target but spaced close together (as shown below) have low accuracy but high repeatability.
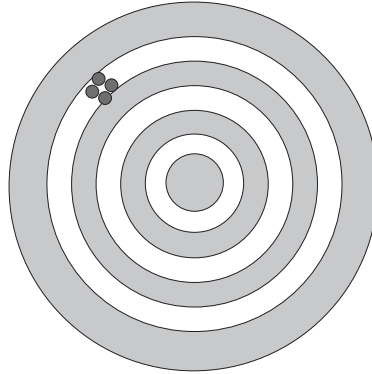
Low accuracy, high repeatability

Figure 2.2. Low accuracy and high repeatability.

The analogies for the two other cases should be fairly obvious. For production parametric test, repeatability is usually more important than absolute accuracy. This is because parametric test is concerned with monitoring trends, which makes having highly repeatable measurements extremely important.

Resolution
For parametric test, resolution is what allows us to gauge accuracy and repeatability. In this sense it is the most important specification for parametric test. Resolution is the lowest resolvable quantity of data that an instrument can accurately measure. The easiest way to think about this is in terms of an analog-to-digital converter (ADC) circuit. The figure below shows a simplified version of an ADC:
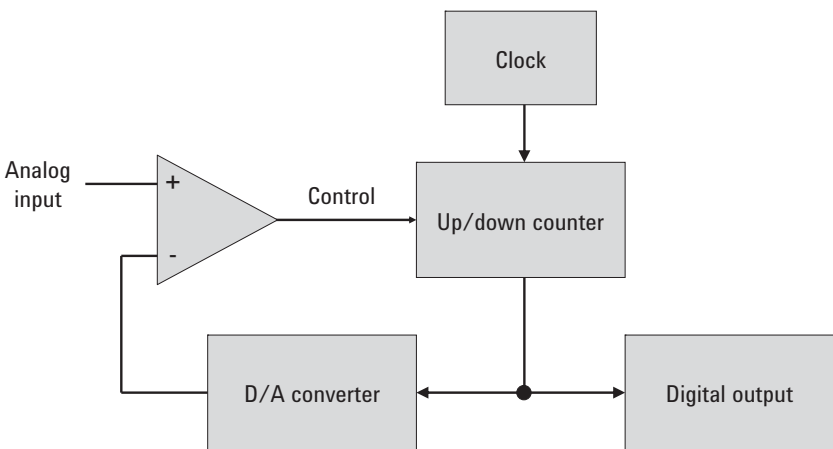
Figure 2.3. Simplified analog-to-digital converter (ADC) circuit.

As this illustrates, the number of bits available to the digital-to-analog converter (DAC) will determine the fineness of the measurement detail that can be distinguished, which is the resolution. For example, a resolution of 20 bits represents the ability to distinguish one part in 2^{20} or 1,048,576 (basically one part in a million).

To express the measurement resolution in terms of current or voltage, you need to know the measurement range in which you are operating. For example, in a 5 V measurement range the ADC measures from –5 V to +5 V, which (assuming a 20 bit ADC) yields a readable or resolvable measurement resolution of:

$$\frac{(5+5)}{1,048,576} \approx 10 \; \mu V$$

However, due to noise and other factors the readable resolution is not the same as the resolution specified in a data sheet. The resolution specified in a data sheet takes into account factors such as thermal noise and amplifier offset and noise, which introduce a stochastic element into the data sheet specification. Typically, the specified resolution is 1 to 2 orders of magnitude larger than the readable resolution.

Measurement traceability

Traceability is defined as the property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons, all having stated uncertainties. It is important for any company making measurement equipment to establish traceability to national and international standards to provide evidence that measurements are being made correctly.

There are a number of National Metrology Institutes (NMIs) that provide standard reference materials of the highest quality and metrological value. The intent of all NMIs is to realize SI units (such as volts, amperes, ohms, etc.) as closely as possible. A partial list of NMIs is shown below.

Name	Acronym	Country
National Institute of Standards and Technology	NIST	USA
National Metrology Institute of Japan	NMIJ	Japan
National Physical Laboratory	NPL	UK
Physikalisch Technische Bundesanstalt	PTB	Germany

Agilent Technologies obtains standards from these NMIs to use as calibration references. Agilent Technologies' measurement products include a certificate of calibration when they are originally shipped from the factory or after a product calibration has been performed. This document is in compliance with the traceability requirements from these NMIs.

Shielding and guarding

Beginners sometimes confuse shielding and guarding, although they are quite different and have totally different purposes. The purpose of shielding is to prevent electrostatic noise from interfering with a measurement. The purpose of guarding is to prevent leakage currents and enable low-current measurements (in the femtoamp range). A shield does not act as a guard, and a guard does not necessarily provide good shielding. As a general rule of thumb, when making measurements below 1 nanoamp you should use guarding; when making measurements below 1 picoamp you should use both guarding and shielding.

Shielding: Maintaining a low noise floor
Shielding involves surrounding the measurement environment with conductive material so as to eliminate electrostatic noise. Basic electromagnetic field theory shows that surrounding a volume with conductive material creates a barrier across which electromagnetic signals cannot cross. This is also often referred to as creating a "Faraday cage." Typically, for parametric test this is accomplished through the use of some sort of shielding box as shown below:
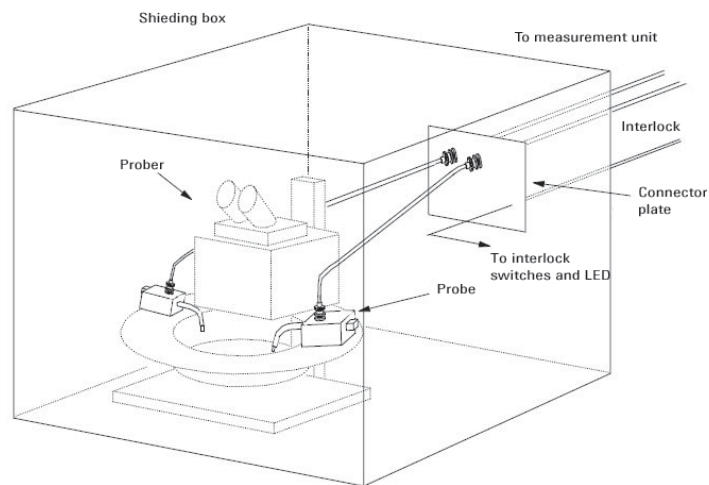
Figure 2.4. A well-shielded parametric measurement environment.

There are a number of best practices that one should keep in mind when trying to create a well-shielded measurement environment:

- Enclose the measurement area with conductive material and measurement cables with a metal shield or braid.
- Electrically connect shielding enclosure and cable shields to the test instrument common (shield) and/or to earth ground.
- To reduce capacitance & vibration effects, keep the shielding and test circuit as far apart as possible.
- Eliminate "light leaks" through gaps around doors and hinges, tubing and cable entry points, connectors and connector panels, and seams/joints between panels.
- Fill seams/gaps/joints with conductive caulking or gaskets.

A reasonable question to ask is: Since a proper shield completely surrounds the measurement environment, how do you connect anything to it? The solution is to use a connector plate that is specifically designed for this purpose.
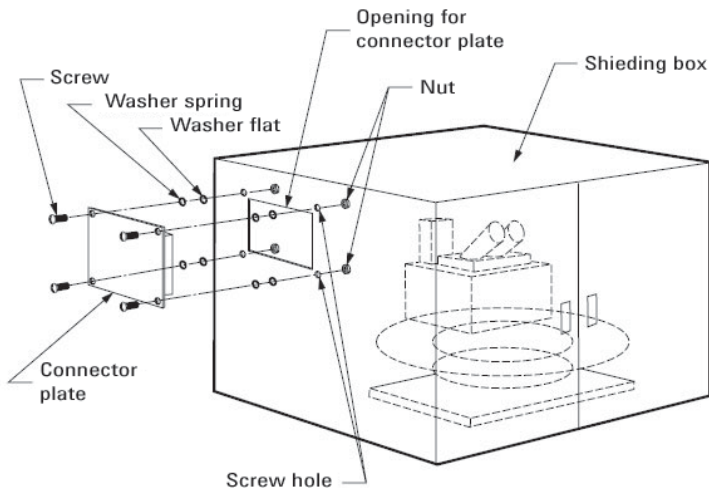
Figure 2.5. Using a connector plate to connect through a shielding box.

Agilent Technologies can supply a variety of connector plates with both triaxial and BNC connectors specifically for this purpose.

In summary, the following are the key points to keep in mind when constructing a well-shielded measurement environment:

- Keep all charged objects (including people) and conductors away from the measurement area.

- Use highly conductive materials instead of insulating materials near the test circuit.

- Avoid movement and vibration near the measurement area (air currents can cause movement and/or vibration).

- When measuring currents < 1 pA, shield the measurement area with a conductive (metal) enclosure and connect the enclosure to the test instrument common (shield) and/or to earth ground.

- Minimize the capacitance between the shielding enclosure and the test circuit.

Avoiding ground loops

One common problem when using shields that has not yet been mentioned is creating ground loops. While all shields are tied to ground, when you have multiple shielded instruments and cables in a given measurement environment it is almost a certainty that the grounds to which these are tied are not at the same potential. Consider the simple case of two conductive planes both tied to ground:
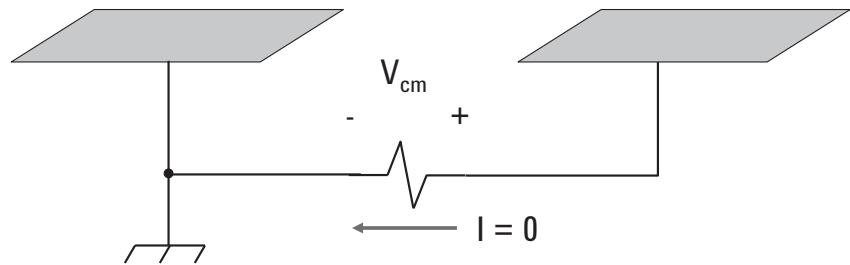
Figure 2.6. Conductive planes connected at only one point cannot have any current flow between them.

As long as these conductive planes are only connected together at one point, any difference in the common mode voltage (V_{cm}) of these planes is not an issue because no current can flow between the planes. However, if the ground planes are strapped together in more than one location then a loop path for current is created that can cause significant current to flow through the conductive planes. This common mode voltage can seriously impact parametric measurement results:
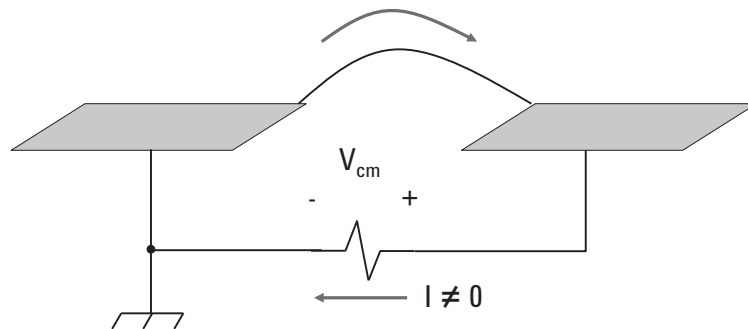
Figure 2.7. Conductive planes connected at multiple points creates a loop for current (a condition to be avoided).

Thus, although it may be counter-intuitive the best strategy when making parametric measurements is to avoid excessive ground connections and instead connect everything together through a single ground point. **Note:** Since instruments almost always have their ground tied to chassis ground, which is in-turn tied to earth ground, this is already taken care of for you. If you suspect that you are having a grounding issue, then the simplest solution is to make sure that the power cords of all of your equipment are tied together to a common power strip.

Guarding: isolation for low-current measurement

Guarding involves surrounding a signal line with an actively driven conductor maintained at the same voltage potential as that of the signal to eliminate leakage currents. Good guarding can only be achieved using triaxial connectors and cabling. In a triaxial cable the signal line is surrounded by the guard line (separated of course by insulating material), which is in-turn surrounded by a grounded shield line (also obviously separated by insulating material). The following illustration shows a cut-away view of a triaxial cable.
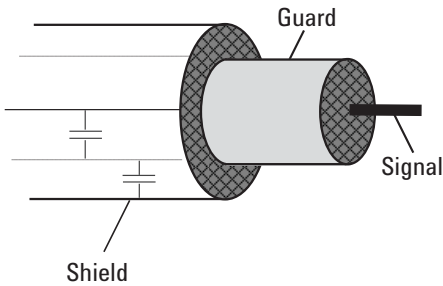
Figure 2.8. Cut-away view of a triaxial cable, showing parasitic capacitance between shield to guard and guard to signal.

Even though there is parasitic capacitance in this cable, the driven guard effectively eliminates the effects of this parasitic capacitance by isolating the signal line from the shield. The diagram below shows the structure of the circuitry driving the guarded (triaxial) output.
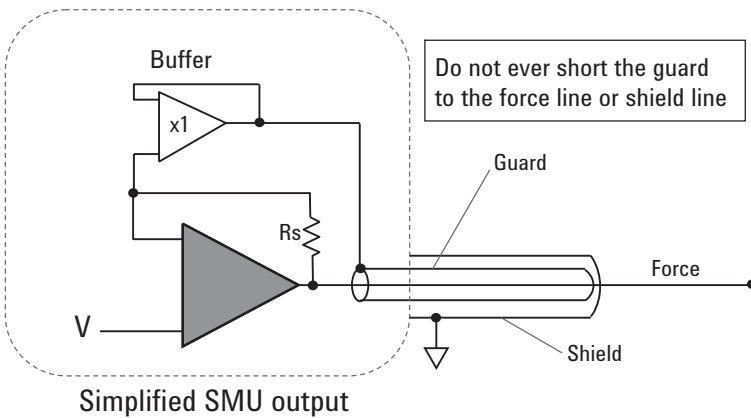
Figure 2.9. Circuitry to provide a guarded (triaxial) output.

The guard line is actively driven by a buffer circuit to maintain the same voltage potential as the center conductor (force) line. Obviously, if you short the guard line to either the force or shield lines then you risk damaging the output circuitry.

Inevitably the question gets asked: Why do we need triaxial cables to measure low current? After all, BNC cables use insulating material, so why should we expect any leakage current? Some simple calculations help to illustrate the need for triaxial cables when measuring ultra-low currents. Assume that the insulation material has a resistance of 1 GΩ and that we apply 100 V to the center conductor:

Outer ground shield

Insulation
$R = 1\ G\Omega = 1 \times 10^9\ \Omega$

Signal wire
$V_s = 100\ V$
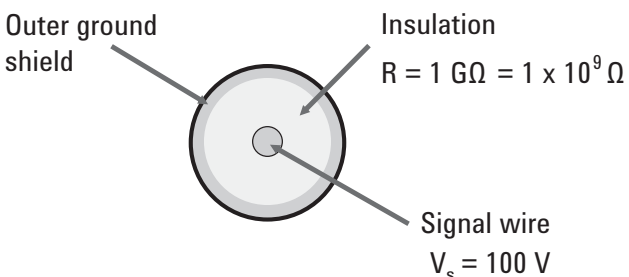
Figure 2.10. Leakage current when using a BNC (coaxial) cable.

A simple calculation shows that the leakage current will be:

$$i_{leak} = \frac{100\ V}{1 \times 10^9\ \Omega} = 1 \times 10^{-7}\ A = 100\ nA$$

Obviously, if we are trying to measure currents in the femtoamp (10^{-15} A) current range, we cannot use cables with 100 nA of leakage current!

Now let us examine the same case using a triaxial cable. In this case we assume that the driven guard can track the applied (signal) voltage to within one part in a million (which is a reasonable and conservative assumption):

Signal wire (V_s)
= 100 V

Conductive guard (V_g)
= 100.000001 V

Insulation
$R = 1\ G\Omega$
$= 1 \times 10^9\ \Omega$

Outer ground shield
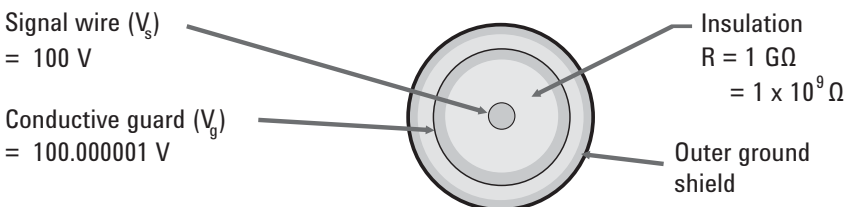
Figure 2.11. Leakage current when using a triaxial cable.

In this case the leakage current will be the difference between the voltages of the signal and guard divided by the effective resistance:

$$i_{leak} = \frac{(100.000001\ V - 100\ V)}{1 \times 10^9\ \Omega} = \frac{1 \times 10^{-6}\ V}{1 \times 10^9\ \Omega} = 1 \times 10^{-15}\ A = 1\ fA$$

By using triaxial cables with a driven guard we have reduced the leakage current to 1 fA, or by a factor of one-hundred million (10^8) as compared to using a BNC cable! It is for this reason that all low-current parametric measurements are made using triaxial cables.

Connecting triaxial and BNC cables
Inevitably you will be faced with the issue of connecting BNC and triaxial cables. The three key questions that require answers are:

1. How do I connect up the driven guard?
2. How does this affect my low-current measurement capability?
3. Where do I get the necessary connectors?

We will address each of these concerns in-turn.

The most important point to understand when connecting up triaxial and BNC connectors is how to connect the driven guard. Remembering that the driven guard exists to allow for sub-nanoamp measurements, we can state as a general rule of thumb that for current measurements above one nanoamp we can simply connect up the force/sense line to the center conductor of the BNC cable and the shield to the outer ground shield of the BNC cable as shown below.
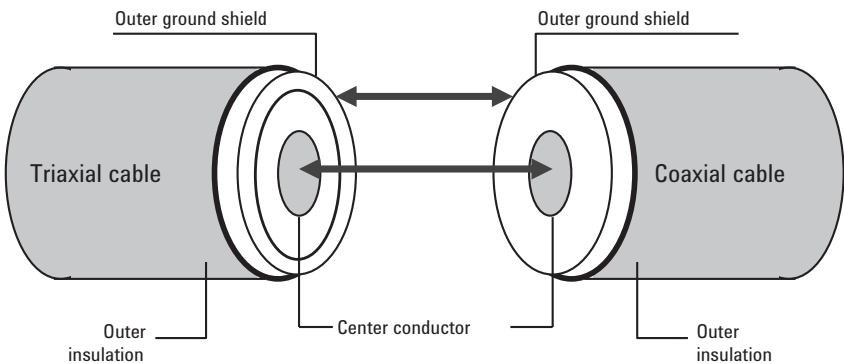
Figure 2.12. Triaxial to BNC connection when not making low-current measurements (1 nA and above).

The advantage of this configuration is its simplicity. We do not need to take any special precautions with the driven guard.

However, the only way that we can make a low-current (below one nanoamp) measurement with a BNC cable is to connect the driven guard to the outer shield of the BNC cable as shown below.
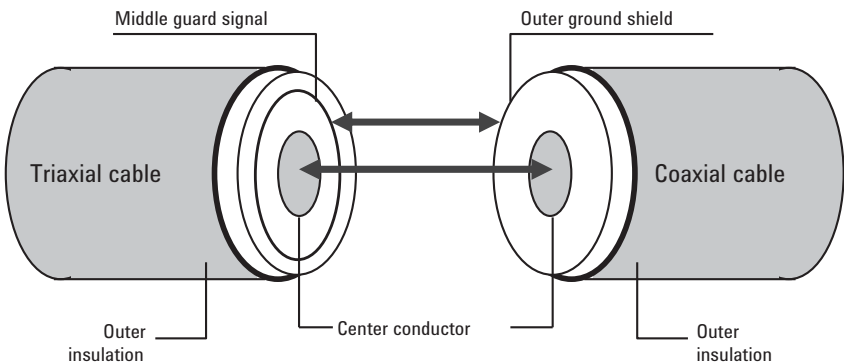
Figure 2.13. Triaxial to BNC connection when making low-current measurements.

This configuration presents a couple of significant challenges:

1. The outer shield of the BNC must be isolated from ground.
2. The outer shield of the BNC can reach dangerous voltages (> 40 V), so precautions must be taken to ensure that the user is properly protected from receiving a fatal electrical shock.

Some analytical wafer prober companies have ready-made solutions to isolate the BNC shield from ground to prevent these types of issues from occurring.

You must be very careful when selecting triaxial to BNC adapters to make sure that you understand which of the above two schemes the adapter is using. If you are uncertain about your connector then it is strongly advised that you use a multi-meter to check the connections. You can purchase triaxial to BNC connectors from Agilent as well as a variety of third-party suppliers.
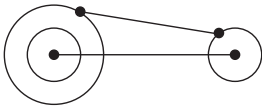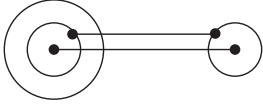
Agilent part number	Description		
1250-2652	Triaxial (F) -BNC(M)		Safe. Not suitable for low-current measurements
1250-2653	Triaxial (M) - BNC(F)		
Agilent part number	Description		
1250-2650	Triaxial (M) - BNC(F)		**WARNING!!!! Shock Hazard!!!!** Required for low-current measurements.
1250-2651	Triaxial (F) - BNC(M)		
1250-1830	Triaxial (F) - BNC(F)		

Figure 2.14. Table showing the different types of triaxial to BNC connectors supplied by Agilent Technologies.

Kelvin (4-wire) measurements

No wire or cable is a perfect conductor. Everything that you use to connect your instruments to the device under test (DUT) has some innate resistance (even though the resistance may be very small). The important question is: When do I need to be concerned about this cable resistance? The corollary to this question is: How do I eliminate the cable resistance effects? In this section we will answer both of these questions.

Every cable can also be thought of as a resistor. The resistance of a cable or wire typically used in parametric test is very small (on the order of milliohms). Obviously, if you are trying to measure a test structure with a resistance that is many orders of magnitude greater than $0.1\ \Omega$ you can safely ignore the cable resistance effects. However, suppose you are trying to measure a structure with only a few Ohms of resistance. In this case the error introduced by the cable resistance can significantly skew your measurement results.

In order to remove the effects of the cable resistance we need to eliminate the resistive voltage drop caused by the current flow through the cable and only measure the resistive effects associated with the test structure. We can do this by making what is known as a 4-wire measurement. This technique requires two separate lines for each terminal on the structure that we want to measure. One pair of lines is used to force current to the DUT, and the other pair of lines is used to sense the voltage measurement. The key point of this measurement is that we separate out the lines used to supply current to the DUT from the lines used to measure the voltage drop across the device. Since the sense lines making the voltage measurement are not conducting any current, there is no voltage drop due to cable resistance. Therefore, the cable resistance effects are eliminated. This technique is often described as a Kelvin measurement, after the famous British mathematician and physicist Lord Kelvin who invented the technique (and who is quoted at the beginning of this chapter).
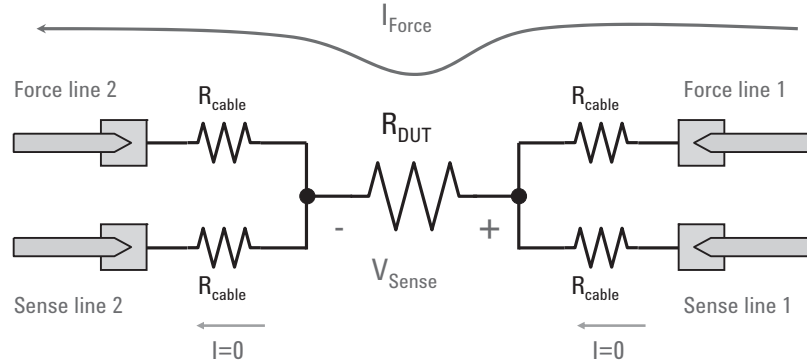
Figure 2.15. A Kelvin (or 4-wire) voltage measurement eliminates the effects of cable resistance by separating out the lines carrying the force current from the lines sensing the voltage.

To be effective, a Kelvin measurement requires that the force and sense lines be shorted together as close to the DUT as possible. Many analytical wafer probing companies make Kelvin triaxial positioners that short the force and sense lines together right at the test pad.
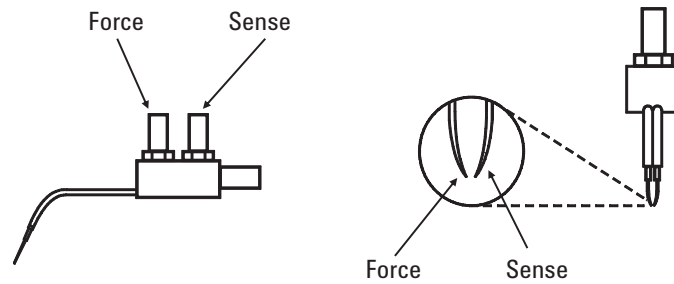
Figure 2.16. Kelvin triaxial positioners (shown above) are ideal since they terminate the force and sense lines at the DUT pad.

Kelvin triaxial cables

Agilent Technologies can supply Kelvin triaxial cables. Each Kelvin triaxial cable can replace two standard triaxial cables. The internal configuration of the Kelvin triaxial cable is as shown below:
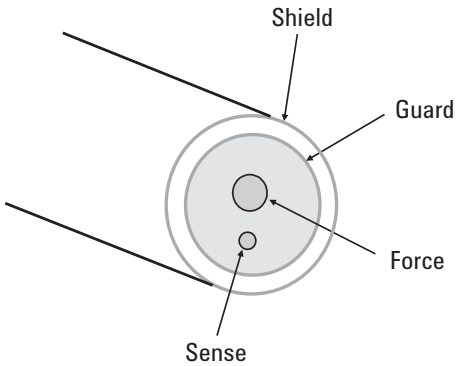
Figure 2.17. The Kelvin triaxial cable combines both the force and sense lines into a single cable.

The Kelvin triaxial cable generally provides better low-noise performance than does using two standard triaxial cables. In addition, the connector on the Kelvin triaxial cable makes it impossible to connect the force and sense lines incorrectly. ***Note:*** You should only order this cable if you are certain that whatever equipment you plan on using it with (wafer prober, switching matrix, etc.) supports this connector.

Figure 2.18. The Kelvin triaxial cable mates with Agilent SMUs and many analytical wafer prober connector plates.

Noise in electrical measurements

Noise is an inescapable part of electrical measurement that arises from the fact that all electronic charge is discrete. It is impossible to give adequate coverage to all of the possible types of noise that can be encountered in parametric test in this short section. Instead, we will give a brief overview of the most common types of electrical noise.

Shot noise

Shot noise is the noise associated with electrical current. It is the result of the fact that an electrical current consists of a flow of electrons and that the electrons have a discrete charge. The value that we define as the electrical current is the average value of the number of electrons passing through a given cross section per second. This is given by:

$$I = nq$$

Where q is the value of the electron charge and n is the *average* number of electrons passing over the given cross section per unit time. The true current, *i(t)*, is time dependent and fluctuates around this average value depending upon the actual number of electrons passing over the given cross section at any instant in time. The shot noise can therefore be written as:

$$i_{Shot}(t) = i(t) - I = i(t) - nq$$

By making certain idealized assumptions we can show that the spectral density function of the shot noise is proportional to only to the electronic charge and the value of the average current:

$$S_i(f) \propto qI$$

The most important feature of this function is that the noise spectrum is constant at all frequencies. Spectra with this type of frequency independence are referred to as *white noise*. Shot noise can have a particularly strong impact on parametric measurements when you are trying to measure a very small signal in a DUT that is being biased by a much larger DC current.

It should be pointed out that the above results are only an approximation. In reality one can easily see that the spectral density function cannot have a constant value across all frequencies. If we were to calculate the mean square current by integrating the spectral density function across all frequencies (from zero to infinity), then we would end up with an infinite value. This does not make any physical sense, since no real current can supply an infinite noise power. There has to be an upper frequency limit at which the spectral density function drops off. However, for most situation of interest in parametric test this frequency is so high that we can make the simplifying assumption that shot noise is constant across frequency.

Thermal noise

Resistors that are not carrying any current can exhibit voltage fluctuations due to the random thermal motion of electrons. This thermal noise (also known as *Johnson noise*) results in voltage fluctuations that have a zero time average value and a spectral density function described by the equation shown below:

$$S_v(f) = 4kTR$$

In this equation k is Boltzmann's constant, T is the temperature in degrees Kelvin and R is the resistance. Since this spectral density function is independent of frequency, it is also a form of white noise. It should be noted that the actual noise observed across a resistor will be larger than that predicted by the above equation due to other thermal processes (such as the continuous generation and recombination of carriers) that are occurring in addition to electron thermal motion. Sometimes this additional noise is accounted for by substituting a "noise temperature" into the above equation, and this noise temperature can be much higher than the actual temperature of the resistor.

Flicker or 1/f noise

Many semiconductor devices exhibit a type of noise that is much greater at low frequencies than would be expected by shot or thermal noise. This type of noise has a spectral density function that can be described by the equation shown below:

$$S_i(f) \propto \frac{1}{f^n} \text{, where } 0 < n < 2$$

In most cases, n is very close to 1. Various types of noise with a 1/f dependency occur in nature, and they are actively studied in many scientific fields. In electronics, 1/f noise is typically referred to as flicker noise, and it is caused by the presence of a direct electric current. You may also see 1/f noise referred to as *pink noise*, to emphasize its intermediate position between white noise (which has no frequency dependence) and red or Brownian noise (which has $1/f^2$ frequency dependence).

Power line cycle noise

One of the most common sources of noise in parametric test is power line cycle noise. The AC current that measurement instrumentation uses for power typically has a frequency of either 50 or 60 Hz (depending on a given country's standards), and this AC current creates powerful electromagnetic fields that can superimpose noise on the quantity under measurement. The effects of power line cycle noise can be mitigated through proper shielding; however, as the measurement range decreases the effects of power line cycle noise become more pronounced. In this case, the only effective means to eliminate power line cycle noise is to average the measurement over more than one power line cycle. Virtually all parametric measurement instruments have this capability. Careful reading of most instrument specifications reveals that their lowest level of measurement resolution actually requires integration over 16 power line cycles. Power line cycle integration will be discussed further in Chapter 3.

Chapter 3: Source/Monitor Unit (SMU) Fundamentals

"You do not really understand something unless you can explain it to your grandmother." — Albert Einstein

SMU overview

Introduction

The primary measurement resource for parametric test is the source/monitor unit or SMU. This is also sometimes also referred to as a source/measurement unit, although the acronym ends up the same. The SMU can force voltage or current and simultaneously measure voltage and/or current. The diagram below shows a simplified SMU equivalent circuit:
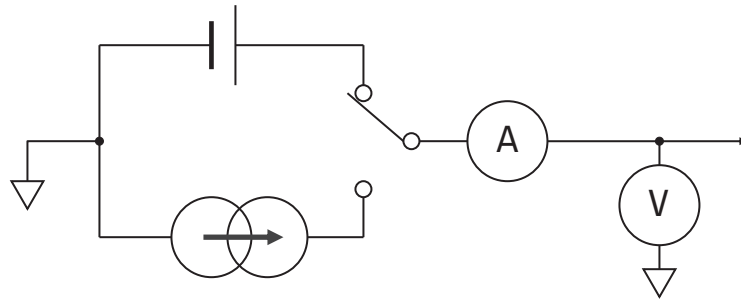
Figure 3.1. Simplified SMU schematic.

Since SMUs must measure very low currents (1 fA or less), they always have triaxial outputs for the reasons discussed in the previous chapter.

There are a variety of different types of SMUs. The most common SMU is the medium-power SMU (MPSMU); as the name implies this SMU can supply moderate levels of voltage and current (±100 V and ±100 mA) and current measurement resolution down to 10 fA. For precision measurements there is the high-resolution SMU (HRSMU); this SMU can supply the same current and voltage as the MPSMU, but it provides current measurement resolution of 1 fA or less and voltage measurement resolution of 0.5 μV. The high-power SMU (HPSMU) is for situations requiring larger currents and voltages than can be supplied by the MPSMU or HRSMU; this SMU can supply current up to ±1 A and voltages up to ±200 V, and the measurement resolution capability is similar to that of a MPSMU.

In addition to the basic SMU types just mentioned, some SMUs support an additional module that enables them to achieve current measurement resolutions of 0.1 femtoamps (100 attoamps). In order to achieve this level of current measurement resolution the actual measurement unit has to be placed in close proximity to the device under test (DUT). This means that this module has to be mounted onto the wafer prober and connected back to an SMU installed in the parameter analyzer mainframe via some sort of cabling arrangement. In the case of the Agilent B1500A, this module is called the atto-sense and switch unit (ASU). In addition to providing 0.1 fA current measurement resolution, the ASU also has some switching capabilities that will be discussed further in Chapter 4.

Module	HPSMU	MPSMU	HRSMU	ASU
Maximum Force Voltage	±200 V	±100 V	±100 V	±100 V
Maximum Force Current	±1 A	±100 mA	±100 mA	±100 mA
Voltage Measurement Resolution	2 μV	0.5 μV	0.5 μV	0.5 μV
Current Measurement Resolution	10 fA	10 fA	1 fA	0.1 fA

Figure 3.2. The key specifications of the basic SMU module types.

The B1505A supports two special types of SMUs. The high-current SMU (HCSMU) operates in two modes: DC and pulsed. In DC mode the HCSMU can source up to ±1 A at 40 V; in pulsed mode the HCSMU can source up to ±20 A at 20 V. The HCSMU has a unique output configuration and cabling requirements that will be discussed later in this chapter. The high-voltage SMU (HVSMU) can source up to 3000 V at 4 mA. Due to its high output voltage capability, the HVSMU requires a special high-voltage triaxial cable that has a special screw-on triaxial connector. This insures that a standard triaxial cable cannot accidentally be used with this module.

Module	HCSMU	HVSMU
Maximum force voltage	±40 V (DC) ±20 V (Pulsed)	±3000 V
Maximum force current	±1 A (DC) ±20 A (Pulsed)	±8 mA at ±1500 V ±4 mA at ±3000 V
Voltage measurement resolution	200 nV	200 mV
Current measurement resolution	10 pA	10 fA

Figure 3.3. The key specifications of the HCSMU and HVSMU (B1505A only).
***Note:** For the HVSMU module the output voltage and current must be the same polarity.*

Except for the HCSMU, SMUs are single-ended devices with one end always tied to a common internal reference point. The SMU reference level is normally tied to chassis ground via an external shorting bar, but this shorting bar can be removed and the SMU reference can be tied to an external voltage (up to 42 V maximum) using various types of connectors.
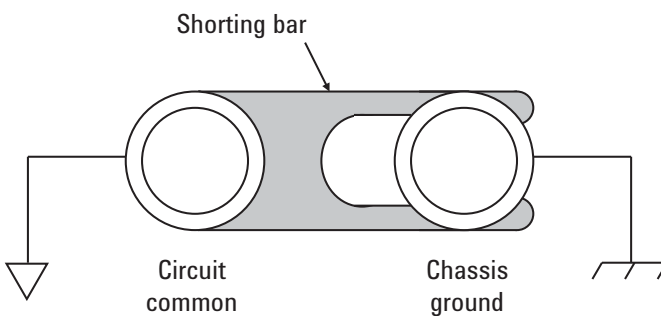
Figure 3.4. You can remove the shorting bar from the chassis ground to float the SMU circuit common.

All SMUs have some pulsing capability during sweep measurements (used to prevent heating on thermally sensitive devices), and SMUs can also make time sampling measurements. However, the pulsing and time sampling capabilities of SMUs are relatively slow (in the microsecond range). It is important to understand when you can use an SMU to make pulsed measurements and when you need the pulsing capabilities of a semiconductor pulse generator unit (SPGU) or waveform generator/fast measurement unit (WGFMU) (which possess nanosecond pulsing capabilities). These different solutions will be covered in detail when we discuss making high-speed measurements in Chapter 5. Because of their relatively slow pulsing capabilities, triaxial cables and DC probes can be used with SMUs when performing pulsed and time sampling measurements. However, as will be discussed in subsequent chapters pulsing and fast measurements made with other types of modules (such as the HV-SPGU and WGFMU modules) require specialized cables and RF probes for optimal measurement results.
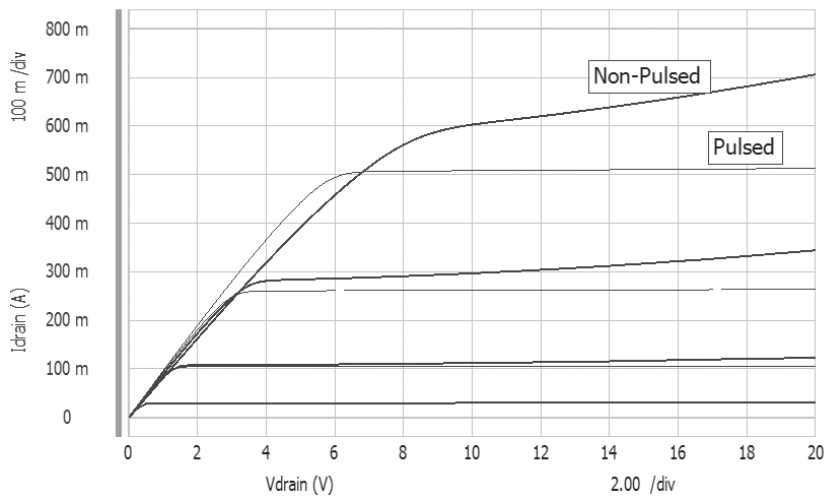
Figure 3.5. SMUs can be pulsed during sweep measurements to eliminate device self-heating effects that can distort measurement results as shown in the above MOSFET Id versus Vd plots.

SMU operation modes and settings

SMUs have three basic modes of operation: voltage source, current source and common. In common mode the SMU acts as a 0 V voltage source, it cannot perform any measurement, and the current compliance is automatically set to the SMU's maximum value. In addition, for sweep measurements you can have one SMU in voltage pulse or current pulse mode to prevent device self-heating on thermally sensitive devices.

SMUs have the ability to specify a compliance setting. The compliance setting is always opposite to that of the source setting of the SMU (that is, current compliance when the SMU is in voltage source mode and voltage compliance when the SMU is in current source mode). When an SMU reaches compliance, it acts as a constant voltage or current source. The compliance feature prevents inadvertent device damage by not allowing the measured quantity to exceed the specified compliance value. In addition, on swept sources it is also possible to specify power compliance. The power compliance prevents the total power output by the SMU from exceeding the specified power compliance value. If both standard and power compliance are specified, then the SMU will never exceed whatever is the lower of these two settings.

On the "Measurement Setup" tab in Agilent EasyEXPERT software you can specify the various compliance settings for each SMU. In addition, there is a "Sweep status" selection menu that allows you to select either "CONTINUE AT ANY" or "STOP AT ANY ABNORMAL". The "continue at any setting" will continue making measurements regardless of any abnormal conditions that may occur (such as a measurement error, reaching compliance, etc.). Conversely, the "stop at any abnormal" setting will immediately halt testing when any of these conditions occurs. Especially in the case of automated measurements (where you are not monitoring the status of the instrument as it measures) the stop at any abnormal feature can be valuable in reducing needless measurement time. An example of this is shown below.
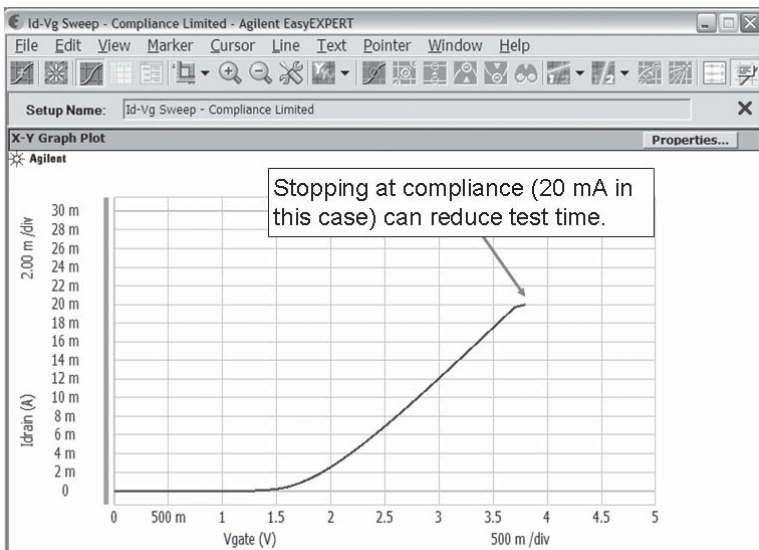
Figure 3.6. By using the stop at any abnormal setting you can automatically halt testing when compliance is reached and avoid needlessly continuing a measurement.

SMU force and sense outputs

The situations requiring Kelvin measurements and the basic theory of Kelvin measurement have already been discussed. By separating the force and sense lines you can eliminate the effect of cable resistance from your parametric measurement. To make this task easier, modern SMUs are designed with both force and sense outputs. The following illustration shows the output configuration of a Kelvin SMU:
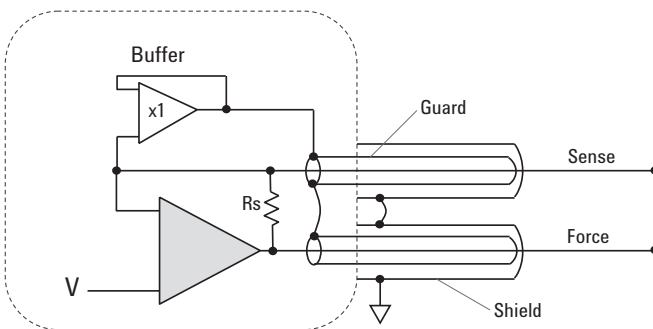
Figure 3.7. Simplified Kelvin SMU output circuit.

The great benefit of a Kelvin SMU configuration is that you only need two SMUs to perform a Kelvin measurement as shown below:
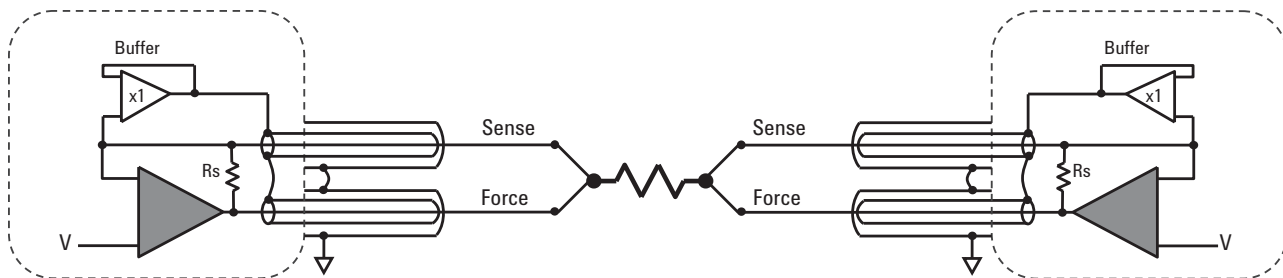
Figure 3.8. Making a Kelvin measurement on a resistor with two Kelvin SMUs.

It is important to point out that the force and sense lines should be shorted together as close to the DUT as possible (for example by using Kelvin triaxial positioners), since the effects of any additional resistance beyond the shorting point will not be eliminated from the measurement.

Many situations do not require Kelvin measurements so it is important to understand which SMU output to use when NOT making Kelvin measurements. If not making Kelvin measurements and using only one output of the SMU then you must always use the <u>force</u> output. The force and sense lines are internally connected via a resistor and this internal sense point is a high-impedance node, so the SMU has no problems with monitoring the current and voltage on the force output when only using the force line. However, consider what happens if you use only the sense line. In this case the entire current flowing out (or in) through the sense output has to pass through the internal resistor as shown below:
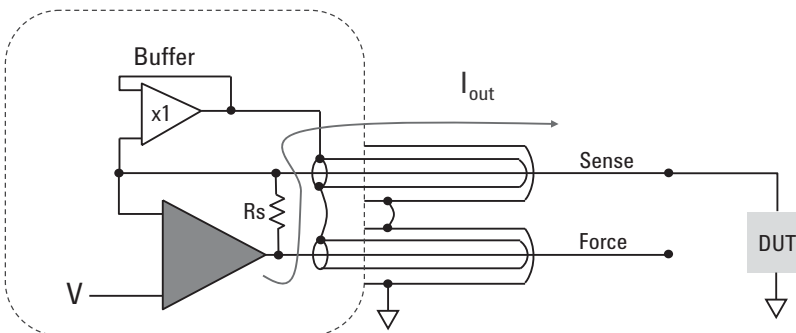
Figure 3.9. The problem with using only the sense line on a Kelvin SMU.

This will distort all of you measurement results and give completely incorrect measurement data! **Never connect only the sense line of a Kelvin SMU to the DUT**.

It is possible to use the SMU sense line as a buffered voltage monitor of the SMU force line in non-Kelvin measurement situations. The most common case is when you are driving the gate of a MOSFET with the force line of an SMU and you want to monitor the gate voltage. Using a triaxial to coaxial adapter (guard floating) you can connect the SMU sense line directly to the oscilloscope input as shown below.
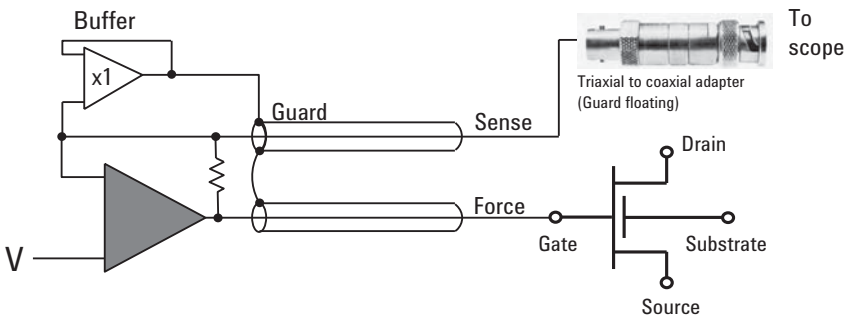
Figure 3.10. If the sense line of the SMU is not used, then you can connect it to an oscilloscope to monitor the SMU output.

Note: When placing any loading onto the output of an SMU you have to be careful, as there is a specified limit on the capacitance. Attaching large capacitive loads to SMU outputs can result in oscillation, so you need to make sure that your oscilloscope input capacitance does not exceed the specified limit of SMU load capacitance.

Understanding the ground unit

The ground unit (GNDU) is a special type of SMU that does not have any measurement capabilities. Its purpose is to provide an active ground to use with the other instrument resources. The ground unit will always maintain a voltage of zero volts as long as you do not exceed the maximum specified value that it can sink or source (for example ±4.2 Amps in the case of the Agilent B1500A and B1505A). The benefit of using an active ground versus a passive ground should be obvious: you do not need to worry about the Ohmic drop from large currents distorting your measurement results (at least if you maintain a Kelvin environment).

The configuration of the ground unit is a source of confusion for many users. This confusion is understandable given that, while the ground unit looks like a standard triaxial connection, it is not. The ground unit has the configuration that it does for historical reasons; in the past there simply was not enough room on some instruments for a separate force and sense line for the ground unit. Therefore, the force and sense lines were merged into a single connector. A comparison of the connections on a standard triaxial cable and the ground unit is shown below:

Standard triaxial connection:

Ground shield

Driven guard

Force/sense line

Ground unit connection:

Ground shield

Force line
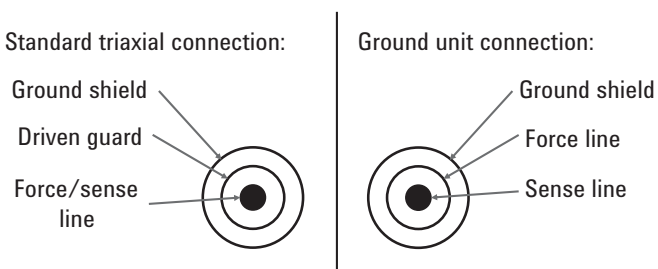
Sense line

Figure 3.11. Comparison of standard triaxial connection and the ground unit.

The ground unit can have the configuration that it does because the potential of the force and sense lines is always at zero volts, so there is no need to isolate them from the outer ground shield to prevent leakage currents. It should be noted that the ground unit is the only case in which this scheme can work.

The illustration shown below highlights the problem with connecting up the ground unit like a standard triaxial output.
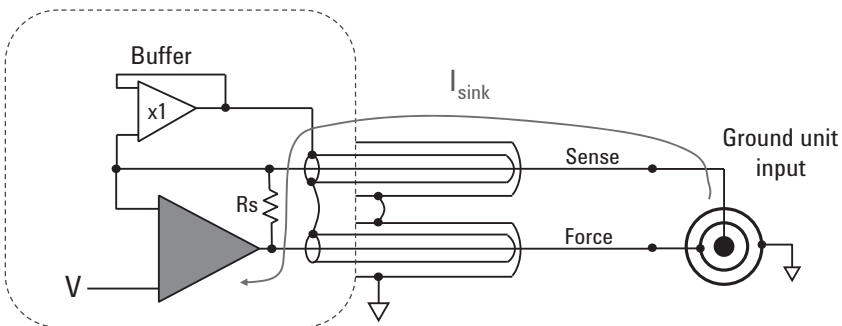
Figure 3.12. The problem with connecting up the ground unit like a standard triaxial connection.

As you can see, connecting up the ground unit like a standard triaxial connection is equivalent to connecting up only to the sense line on a Kelvin SMU. Obviously, this will lead to erroneous measurement results.

The following illustration shows the proper way to connect the ground unit to standard Kelvin SMU connections. Agilent can supply a ground unit to Kelvin adapter, the N1254A-100, which converts the ground unit output into the force and sense outputs as shown below.
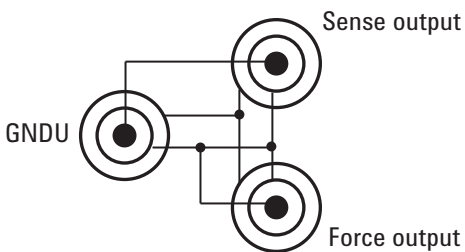
Figure 3.13. The proper way to connect a ground unit to standard triaxial connections.

Of course, the ground unit adapter can be used with only the force output con-
nected just like a standard SMU. However, you need to be careful when doing
this. Assumedly one reason that you are using the ground unit is to sink large
currents, and large currents innately require a Kelvin connection (using both the
force and sense outputs). Therefore, especially if you are using the ground unit
to sink current from one or more high-power SMUs, it is highly recommended
that you connect the ground unit in a Kelvin configuration as shown below:
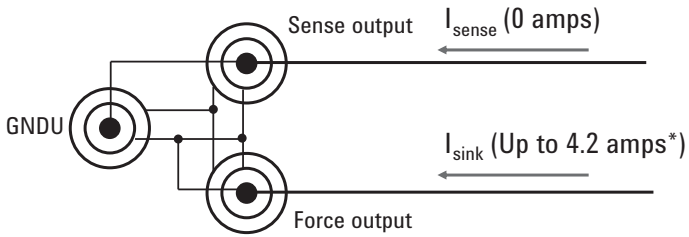
*B1500A & B1505A

Figure 3.14. The ground unit should be used in a Kelvin configuration when sinking large currents.

Note: Agilent makes a special triaxial cable designed to be used with the force
output of the ground unit. This cable can handle the 4.2 amps maximum current.
The part numbers are 16493L-001 (1.5 m), 16493L-002 (3 m) and 16493L-003 (5 m).

High current SMU connections (B1505A)

As was mentioned earlier, the high current SMU (HCSMU) is a special module available only for the B1505A. Its structure is similar to that of an SMU except that it can source up to 20 A of current (pulsed). Because no other module can sink this much current, the HCSMU also has to have the ability to sink its own current. This gives the HCSMU a unique output configuration.
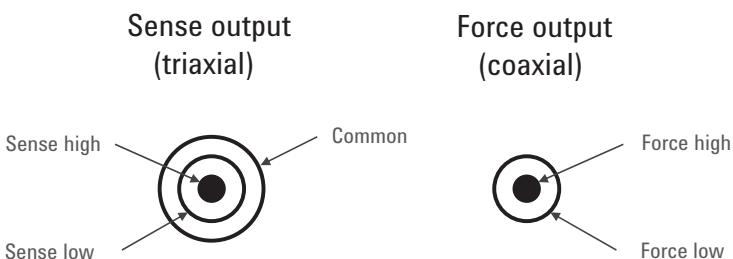
Figure 3.15. The output connections on the HCSMU.

The force lines of the HCSMU do not perform any measurement so they do not require any shielding and they can be coaxial. On the other hand, the sense lines of the HCSMU do perform measurement so they require shielding and they need to be triaxial. Fortunately, this arrangement makes it impossible to confuse the two outputs.

The HCSMU module is floating and is not tied internally to the instrument ground. This means that its low force and sense outputs must be tied to a reference level (normally the ground unit) when making a measurement. An example MOSFET measurement using the HCSMU is shown below.
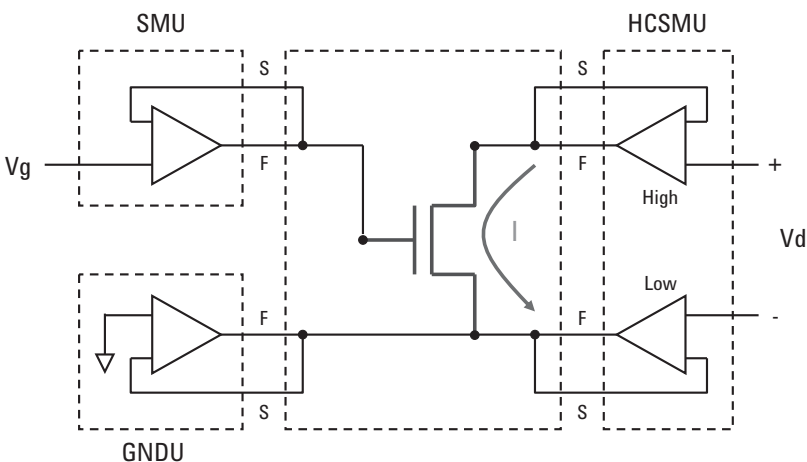
Figure 3.16. The HCSMU needs to have its low outputs (force low and sense low) tied to a known voltage reference when making a measurement.

When you are using the HCSMU with the N1259A high-power test fixture the N1259A takes care of correctly separating out the HCSMU connections. However, you need to make sure that you use the correct cables and adapters when connecting the HCSMU up to a wafer prober. Agilent supplies a variety of adapters and connectors for this purpose. This will be covered in greater detail when we discuss making on-wafer measurements in Chapter 4.

Measurement ranging

Measurement ranging basics

Measurement ranging is intimately interrelated with measurement accuracy and resolution. However, before proceeding it is it important to understand why SMUs have a range setting in the first place. The SMU circuitry has to switch in and out (using relays) different resistor values in order to handle the maximum expected current or voltage value based upon the initial compliance setting specified by the user. An example of this circuitry is shown below:
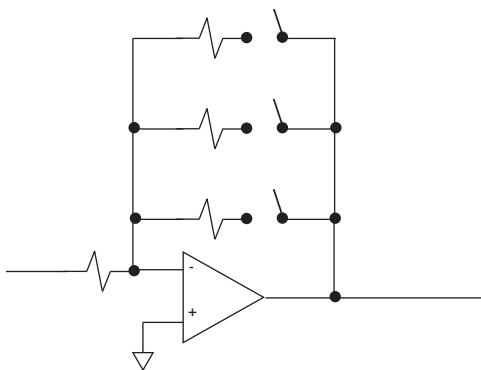
Figure 3.17. SMU ranging requires that resistors of different values be switched in and out of the circuit depending upon which particular range you are in.

Selecting one or more resistors via these relay switches places the SMU into a given measurement range. Obviously, it takes some time to switch these relays and move from one range to the next. While it would be possible to always have the SMU start at the highest possible measurement range and work its way down to the lowest measurement range that contained the quantity being measured, this would result in extremely slow measurements. By allowing flexibility as to how a measurement range is selected, the user gains the ability to trade off measurement speed versus accuracy.

There are typically three types of ranging selectable on an SMU: fixed, limited, and auto. The illustration below shows how each of these choices impacts the measurement range used by the SMU:
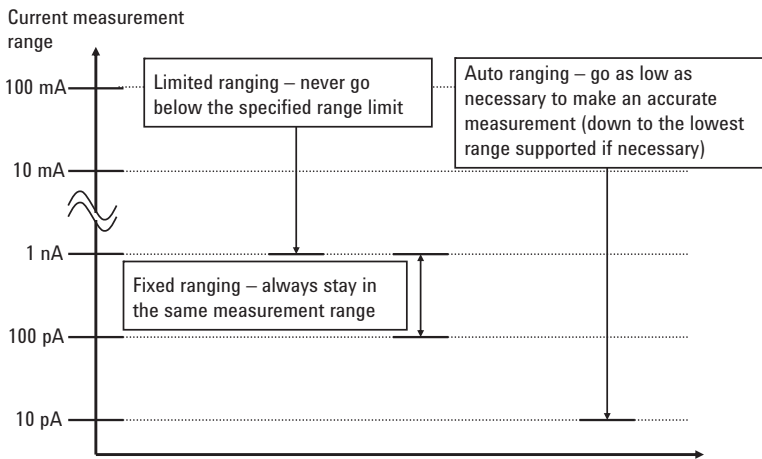
Figure 3.18. An explanation of the three types of measurement ranging: fixed, limited and auto.

Ranking the measurement ranges in terms of fastest to slowest, the order would be: fixed (fastest), limited (next fastest), and auto (slowest). Although fixed measurement ranging yields the fastest measurement results, it has the limitation that the SMU will not go into a lower measurement range to improve measurement accuracy. Also, if you attempt to measure a current or voltage in fixed measurement range that exceeds the maximum value of the fixed measurement range you will get a measurement error. Limited ranging and auto ranging are similar in that they both start in the highest measurement range that contains the user-specified compliance value, and they both work their way down to find the optimal range in which to make the measurement. The difference between the two ranging choices is that limited ranging will never go below the user-specified range limit. Thus, limited ranging is useful when you are uncertain of the value of current or voltage that you will be measuring and you do not care about ultra-precise measurement. However, if you absolutely have to have the best measurement accuracy and measurement time is not a concern then auto ranging is the correct choice.

There is no hard and fast rule to specify the measurement resolution achievable within a given measurement range. The Chapter 2 discussion of measurement resolution pointed out that the primary determining factor is the number of bits in the SMU analog-to-digital converter (ADC). However, the actual measurement resolution achievable has to take into account other factors such as noise, drift, etc. This introduces a stochastic element to the measurement that requires averaging. The net result is that in most cases the minimum measurement resolution is 4-5 decades below the measurement range. Note that on some instruments the SMUs may have more than one ADC available to them; in this case you need to check carefully to make sure that you understand the measurement resolution associated with the ADC that you are using.

There is one important point to understand regarding the use of fixed measurement ranging that was mentioned previously but is worth repeating again here. In fixed measurement ranging you can measure currents or voltages smaller than the measurement range that you have chosen, since as mentioned above the measurement resolution is 4-5 decades below the measurement range. Even if the actual measured value is more than 4-5 decades below the selected measurement range, the instrument will still return a result (albeit with reduced measurement resolution). However, if you try to measure a current or voltage that exceeds the specified fixed measurement range then you will get a measurement error. Therefore, finding the optimal fixed measurement range is always a balancing act between selecting a range low enough to provide sufficient measurement resolution and high enough to always contain the quantity under measurement.

Measurement range management

In order to understand the range management feature and why it is sometimes necessary we must first understand the issue of range compliance. We previously explained that the SMU switches various resistors in its internal circuitry in and out as it moves between different ranges. This process takes some time and it is part of the normal operation of the SMU. However, what was not mentioned was that as the SMU (for example) moves into lower current measurement ranges, its output sourcing capability is also temporarily reduced to the value of the measurement range plus a guardband factor (around 10%). This is called "range compliance" to differentiate it from the SMU compliance that is specified by the user. Obviously, range compliance is only a transient issue in that if the SMU needs to source more current then it will move up into the next measurement range until it either can supply the required current or it reaches the user-specified SMU compliance.
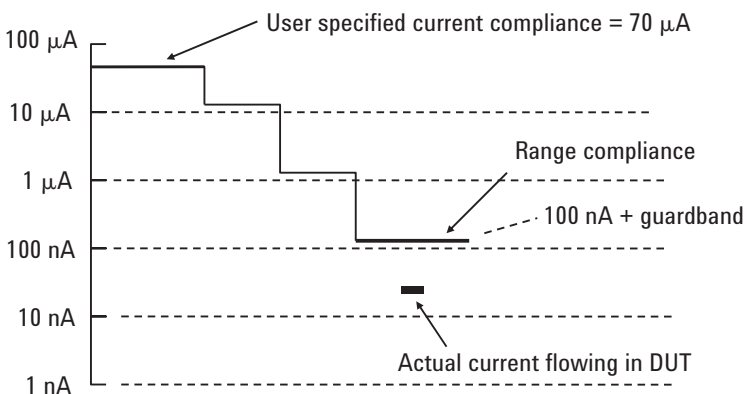
Figure 3.19. Range compliance limits the amount of current that an SMU can supply in a given measurement range.

In order to understand why range compliance can sometimes be an important issue, please refer to the figure shown below.
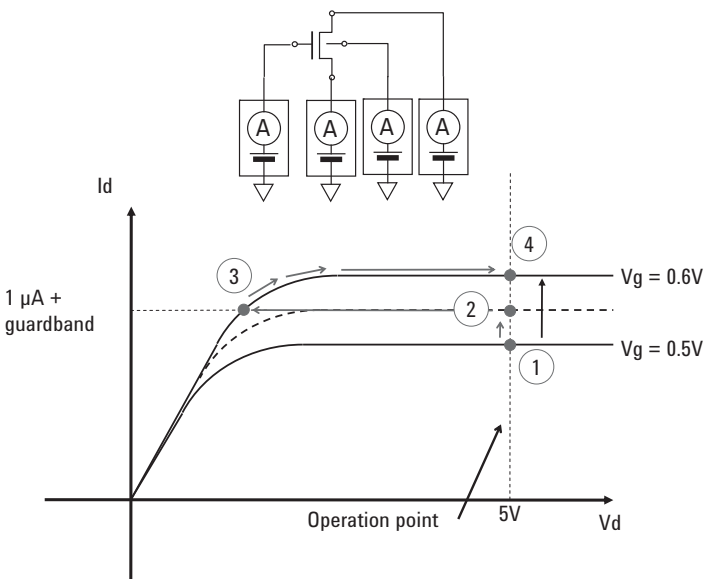
Figure 3.20. The effect of range compliance on SMU voltage output.

41

Suppose that we are measuring a device with the Id-Vd characteristic shown in the above figure and that we are starting at the operating point denoted by "1" (Vg = 0.5 V). We want to move to the operating point denoted by "4" by changing the applied gate voltage to 0.6 V. However, this action requires us to move up into the next measurement range. Since this measurement range change is not instantaneous, as soon as the operating point moves to the position denoted by "2" the SMU cannot supply all of the current that the DUT wants. This means that the operating point must move to the position denoted by "3". When the SMU yfinally moves up into the next measurement range it can supply the current requested by the DUT and the operating point moves to position "4". When viewed on an oscilloscope this would give the appearance of an SMU voltage glitch, even though what is really happening is that the DUT is causing the voltage droop because the SMU is starving it for current.
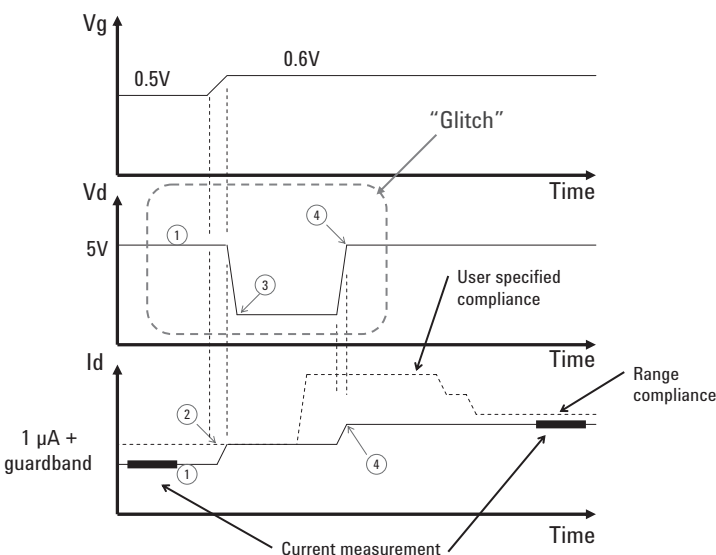
Figure 3.21. Voltage "glitch" on a MOSFET drain caused by temporary range compliance.

For the vast majority of parametric measurements, range compliance never causes any measurement issues. In fact, if we had not just discussed this issue you probably would never have suspected that it existed. However, in a certain small percentage of cases the voltage swings caused by the DUT due to current starvation can impact parametric measurements and even cause device damage (if the voltage swing is on the substrate and it causes the device to latch up). Until the development of the range management feature, the only solution to this issue was to replace the sweep measurement being performed with a series of spot measurements. Replacing the sweep measurement with a series of spot measurements always fixes this issue, since a spot measurement always starts at the measurement range containing the SMU compliance value and works its way down to the correct measurement range (thus avoiding any range compliance issues). However, performing a series of spot measurements in this fashion takes much more measurement time than an equivalent sweep measurement, and this was not acceptable to most users. To solve this issue Agilent Technologies developed (and patented) the range management feature.

Range management can be viewed as a sort of "range look-ahead" feature. The range management feature allows you to set trigger points within a measurement range (from 11% to 99% of the range) that force the SMU to go up (or optionally down) to a new measurement range BEFORE the measurement range limit is reached. The figure shown below helps to illustrate how the range management feature works.

"*current1*" is the up-ranging trigger point set via Range Management; this is set as a percentage of the total range.
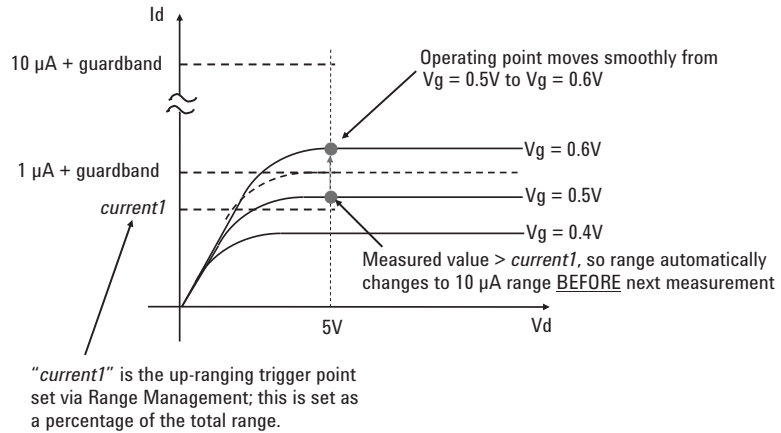
Figure 3.22. Illustration of how the range management feature causes the SMU to move up in range before reaching the range limit.

Applying the range management feature to the case previously discussed, we can see that the voltage droop seen at the MOSFET drain can be eliminated completely.
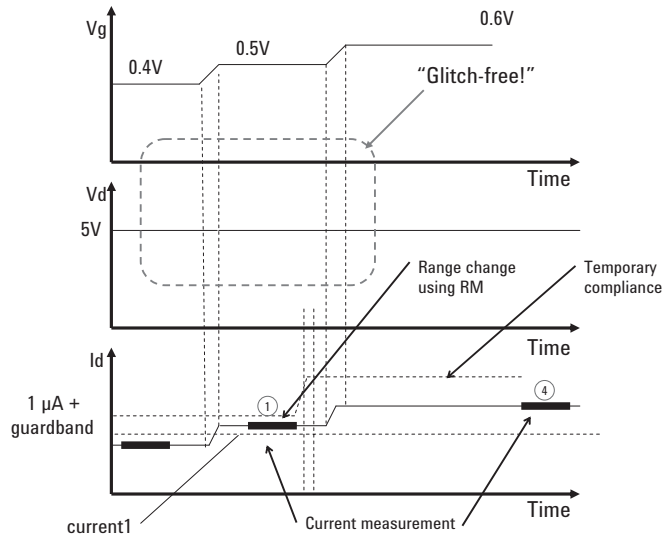
Figure 3.23. Eliminating the effects of range compliance through the use of the range management feature.

It is a worthwhile question to ask if there are any disadvantages to using the range management feature. The answer is "yes" in the sense that if the range limits are set too low then you might uprange unnecessarily and lose some measurement accuracy. However, this is a small potential price to pay if range compliance is causing repeated device damage. It should be obvious that the phenomena of SMU voltage spikes due to DUT current starvation can be

affected by many factors, including the sweep range, the selected sweep step and variations in the DUT characteristics. Therefore, some trial and error is inevitable in order to find the optimal point at which to set the range trigger ("current1"). Finally, it also needs to be pointed out that not all device damage issues and SMU glitching are due to range compliance issues, so you need to be careful and look at all possible causes if you are experiencing these types of problems.

Agilent EasyEXPERT software supports the range management feature for the B1500A and B1505A as part of its built-in GUI. By opening up the range setup window in Classic Test mode you can modify the range change rule as shown below.
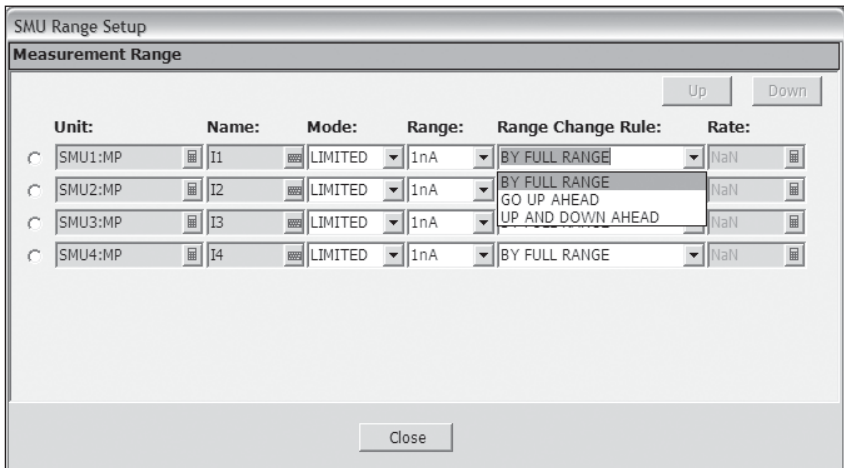
Figure 3.24. Agilent EasyEXPERT software allows you to activate the range management feature using its built-in GUI.

You can select an option to only uprange ("Go Up Ahead") or both uprange and downrange ("Up And Down Ahead"). Once you have specified a range change rule you can then select the rate or percentage of the total range at which you want to have the SMU uprange (or downrange too if "Up And Down Ahead" is selected).
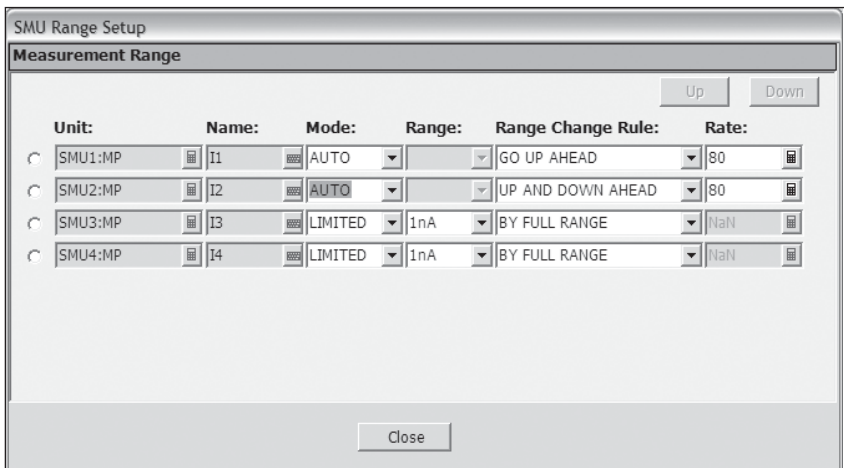
Figure 3.25. The "Rate" parameter sets the percentage of the total range at which the SMU will uprange (or optionally downrange).

Eliminating measurement noise and signal transients

Integration time

Inexperienced users sometimes confuse the purposes of measurement ranging and integration time. It is important to understand that the purpose of integrating a measurement over time is to eliminate noise. Increasing the integration time does not have the same effect as using a lower measurement range. Before increasing integration time to improve your measurement results you should first determine if you have chosen the correct measurement range for the level of current or voltage that you are trying to measure. For example, it makes no sense to try and make a femtoamp current measurement using limited 1 nA ranging on an SMU, since the SMU needs to get down into the 10 pA range in order to make good femtoamp measurements. As a general rule, lower measurement ranges require longer integration times in order to obtain a satisfactory measurement because noise becomes more of an issue as you try to measure small currents and voltages.

The terminology used for integration time is not uniform across all products. In some cases, the term "Short" is used to refer to any integration time that occurs in a time period that is less than one power line cycle (PLC); "Medium" is then used to refer to integration over exactly one PLC, and "Long" is used to refer to integration over multiple PLCs. In other cases, the terms "Auto" and "Manual" are used for integration times of less than one PLC, and "PLC" is simply used to refer to integration over one or more PLCs. The documentation included with your instrument explains the exact terminology used for your equipment. The important point to understand is not the particular terminology used but rather when to use each type of integration time.
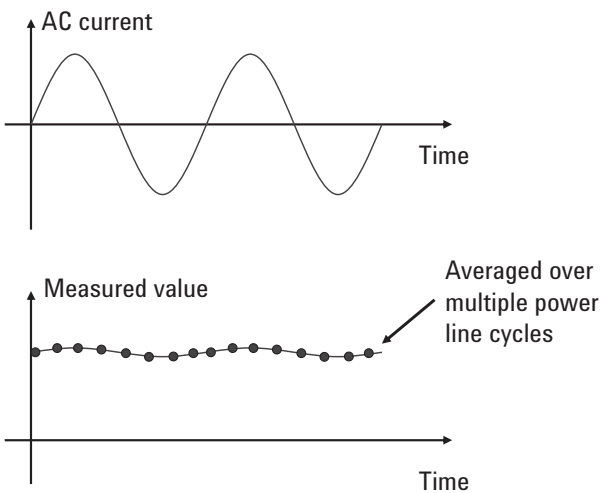
Figure 3.26. Power line cycle (PLC) integration eliminates measurement error caused by noise from the AC supply current by sampling over multiple power line cycles and averaging the samples.

Hold time and delay time

In addition to noise, capacitance on the SMU outputs can cause ringing and other transient phenomena each time the SMU applies a new voltage or current. To insure that the applied voltage or current is stable before making a measurement, you can specify both a measurement hold time and a measurement delay time as shown below.
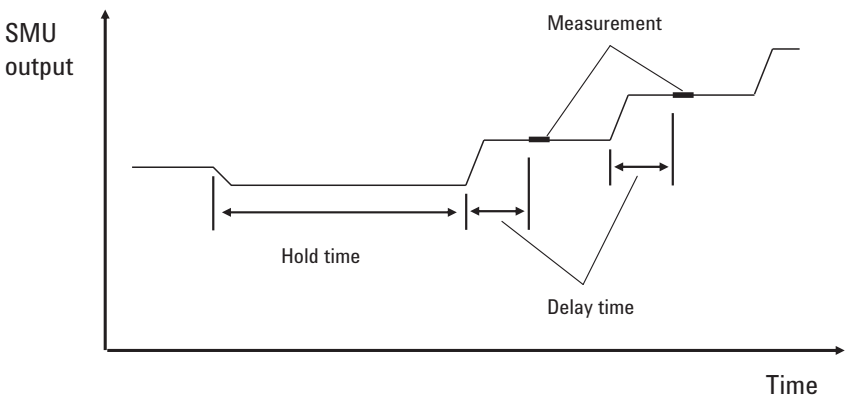
Figure 3.27. The hold time and delay time settings allow you to specify how long to wait before starting a measurement after the SMU applies voltage or current.

The hold time insures that the SMU outputs are stable before the start of a measurement, and the delay time insures that the SMU outputs are stable during a measurement if the value of the SMU output is being changed.

Low current measurement

Making low current measurements with SMUs (in the femtoamp or even sub-femtoamp range) presents a variety of measurement challenges. However, none of these challenges are insurmountable, and all of them can be met using well-established measurement techniques. A summary of these techniques is shown below:

- Create a Low-noise Probing Environment
    - Maintain a shielded measurement environment (Faraday cage)
    - Get rid of electrical noise sources

- Eliminate Stray Leakage Currents
    - Use fully guarded cabling and probes (down to the probe tips)
    - Use a guarded wafer chuck

- Choose Appropriate Measurement Setup Parameters
    - Select correct measurement range, integration time and hold time

- Perform Calibration and SMU Zero Offset

Techniques for guarding and shielding have already been discussed in the previous chapter, and techniques for making low-current measurements on-wafer will be discussed in the next chapter. However, there is one SMU feature that is critical for making successful low current measurement that has not yet been discussed: the zero cancel function.

SMU zero cancel function

Offset currents can create some significant measurement challenges when making current measurements in the femtoamp range. There are a variety of sources for offset currents, but the important point is that as long as they are consistent they can be eliminated. The 4155C, 4156C, B1500A and B1505A all have a built-in feature known as "SMU Zero Offset Cancel". With the outputs open the SMU zero cancel function measures the SMU offset current and then automatically subtract out this offset when the SMU performs subsequent measurements.
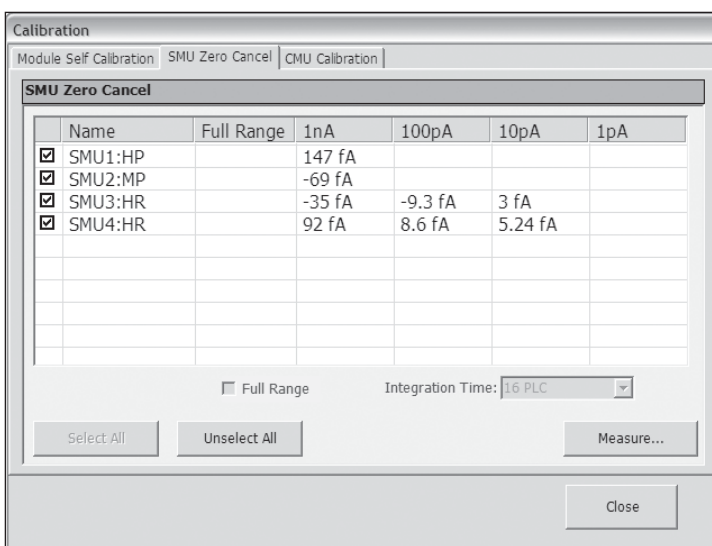
Figure 3.28. The results of performing an SMU zero cancel calibration.

On the B1500A and B1505A the SMU zero offset cancel function status appears in the bottom menu bar as shown below.
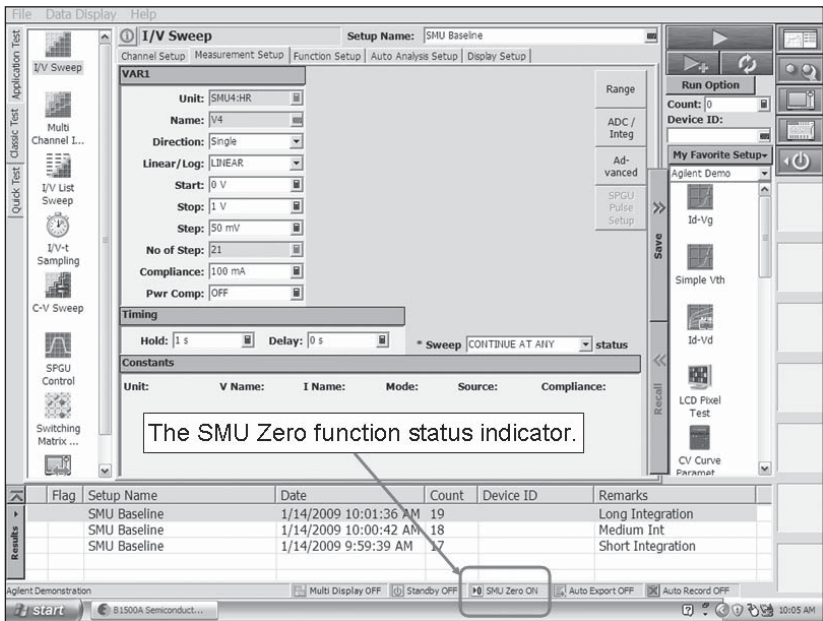
Figure 3.29. The location of the "SMU Zero" function status in the main EasyEXPERT window.

After performing a zero offset cancel calibration, you should be able to obtain a very good (±3 fA) low-current base line measurement on a HRSMU with the outputs open. The following plot shows the results of such a measurement using short integration on the B1500A.
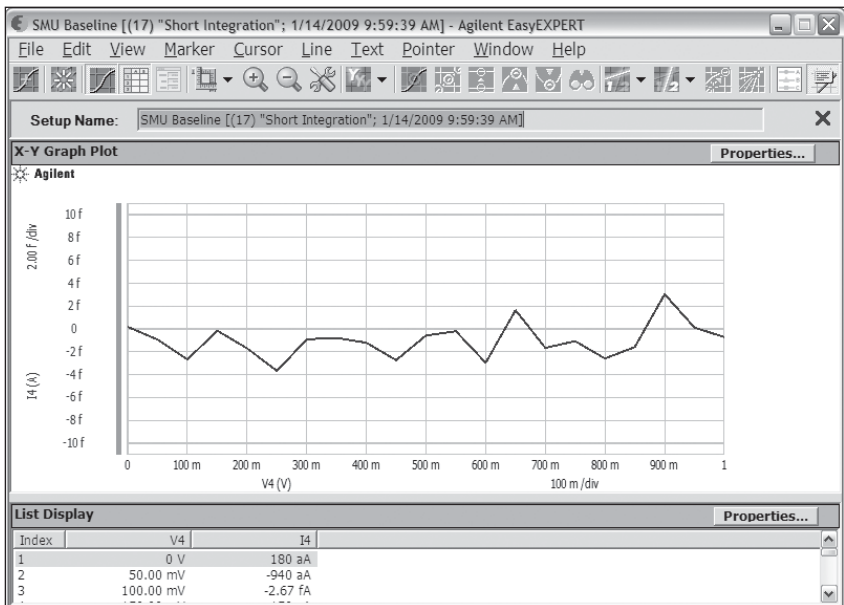
Figure 3.30. The low current base line for a high-resolution SMU (HRSMU) using short integration. As can be seen, this baseline only varies by about ±3 fA.

Debugging low current measurements

Debugging low current measurement issues is not a difficult process as long as you follow a logical procedure. The following debug flow assumes that high-resolution SMUs (HRSMUs) are being used. As mentioned earlier, it makes no sense to try making femtoamp measurements unless at least a 10 pA measurement range is available. If you are having trouble making low-current measurements then the following debug flow should allow you to isolate the problem area.

1. Sweep 0 V to 1 V with no cables attached (after 30 minute warm-up)
   a. Check to make sure that all SMU modules passed self-calibration
   b. Perform SMU zero offset cancel
   c. Verify that there is a ±3 fA base line
   d. **Note:** Success indicates that the SMUs are functioning correctly.

2. Connect prober cables but do not connect the cables to the wafer prober
   a. Wait several minutes to allow piezoelectric effects to dissipate
   b. Perform SMU zero offset cancel
   c. Check base line
   d. **Note:** Success indicates that the cables are not damaged (leaky).

3. Connect the prober cables to the prober connector plate but do not connect them to the wafer probes
   a. Wait several minutes to allow piezoelectric effects to dissipate
   b. Perform SMU zero offset cancel
   c. Check base line
   d. **Note:** Success indicates that the connector plates are not leaky.

4. Connect wafer probes to connector plate
   a. Wait several minutes to allow piezoelectric effects to dissipate
   b. With probes up (not contacting wafer), perform SMU zero offset cancel
   c. With probes up (not contacting wafer), check baseline
   d. **Note:** Success indicates that the wafer probes are functioning correctly.

Power line cycle noise tends to become more prominent at lower current measurement ranges. This means that you may need to integrate low-current measurements over multiple power line cycles in order to get noise-free results. For sub-femtoamp measurement made using the ASU, this can become even more of an issue. In fact, one experimenter once reported that the only way he could obtain decent sub-femtoamp measurements was to come in during the weekend when most of the electrical equipment in the building in which he was working was shut-off.

Spot and sweep measurements

Spot versus sweep measurements

Two basic types of measurements are made during parametric test: spot and sweep. A spot measurement is a single-point measurement that produces a scalar result. A sweep measurement is a series of measurements that produces a vector result. During a sweep measurement one independent variable is changed and the dependent variable is then measured and plotted. At first it might seem that a sweep measurement is simply a series of spot measurements; however, this is actually not the case. In order to make the sweep measurement complete quickly, the instrumentation makes certain assumptions as it moves from one measurement point to the next in the sweep. The basic assumption made is that the measurement range (if in limited or auto ranging) does not need to change from one measurement point to the next. Of course, if in moving from one point in the sweep to the next the measurement range needs to change, then the measurement circuitry will make this change. The important point is that, unlike a spot measurement (which always starts at compliance and works its way down to the correct measurement range), a sweep measurement tries to minimize unnecessary range changes. By making this assumption, a sweep measurement can complete in significantly less time than could a similar collection of spot measurements.

Primary sweep measurement

For a primary sweep you must specify three parameters: start, stop and step (the amount the sweep variable should be incremented between points). The instrument will first measure with the sweep variable at the "start" value, and then it will increment the sweep variable by the amount specified by the "step" value and measure again. This process of incrementing by the step value and measuring continues until the sweep variable reaches the "stop" value. Note: The 4145A/B referred to the primary sweep variable as VAR1 (for variable #1), and the 4155C, 4156C and B1500A/B1505A (in classic mode) also use this same terminology for consistency.
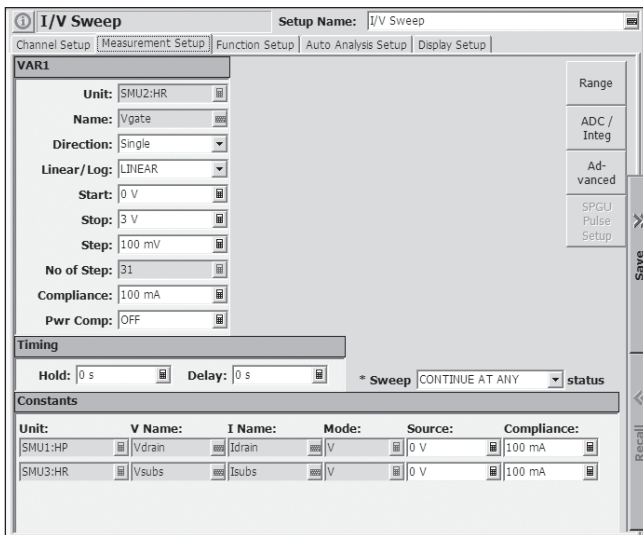
Figure 3.31. Setting up a sweep measurement in classic test mode.

In addition to the aforementioned sweep parameters, the other basic sweep setup parameters are direction and linear/log. The direction parameter allows you to have the sweep only perform a single sweep (from start to stop), or to perform a double sweep (from start to stop and then repeat backwards from the stop to the start value). The linear/log parameter allows you to specify whether or not the sweep steps are spaced linearly or logarithmically. In the case of a logarithmic sweep you can select 10, 25 or 50 points per decade.

Pulsed sweep measurement

As was mentioned in an earlier section, SMUs have the ability to do pulsed measurement.

For the B1500A the minimum pulse width is 500 μs and the maximum pulse period is 5 s (for a minimum duty cycle of 0.01%). In general, this is more than adequate for the testing of most high-power devices (at least up to the ±200 V and ±1 A sourcing capabilities of a high-power SMU). The B1505A's HCSMU has a minimum pulse width of 50 μs (at 20 A) and the maximum pulse period is also 5 s (for a minimum duty cycle of 0.001%). A comparison of a standard sweep measurement versus a pulsed sweep measurement is shown below.
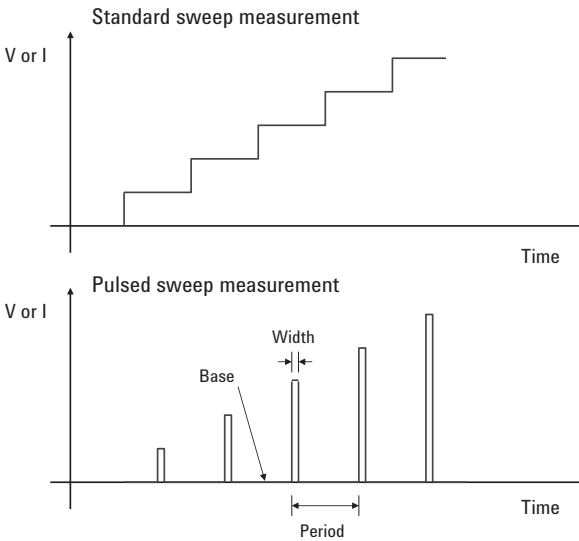
Figure 3.32. By making a pulsed sweep measurement with an SMU, the effects of device heating can be eliminated.

Note: SMUs do not have a small enough duty cycle to prevent self-heating in SOI transistors, as these can require pulse widths in the nanosecond range. Techniques for measuring these types of devices will be discussed in Chapter 5.

It is important to remember that when making pulsed measurements you need to make sure that you are not specifying any other measurement conditions that might conflict with the pulse settings. Although this issue will be discussed in greater depth in Chapter 5 when we discuss time dependent measurements, the basic rule for SMUs is that accuracy trumps all other settings. This means that you must specify short integration times and use fixed ranging when making very small pulsed measurements or you will not get the pulse width that you specify. In other words, the SMU will make the pulse width however long is necessary to accommodate the integration time and/or range changes that you have specified. If you are using the B1500A or B1505A to make measurements then this is taken care of for you automatically; however, if you are using an older instrument (4155 or 4156) then you need to take care of this manually.

Subordinate sweep measurement

The subordinate sweep function causes the primary sweep variable to repeat each time that the secondary sweep variable is incremented. For a subordinate sweep measurement you must specify three additional parameters: start, step and the number of steps.
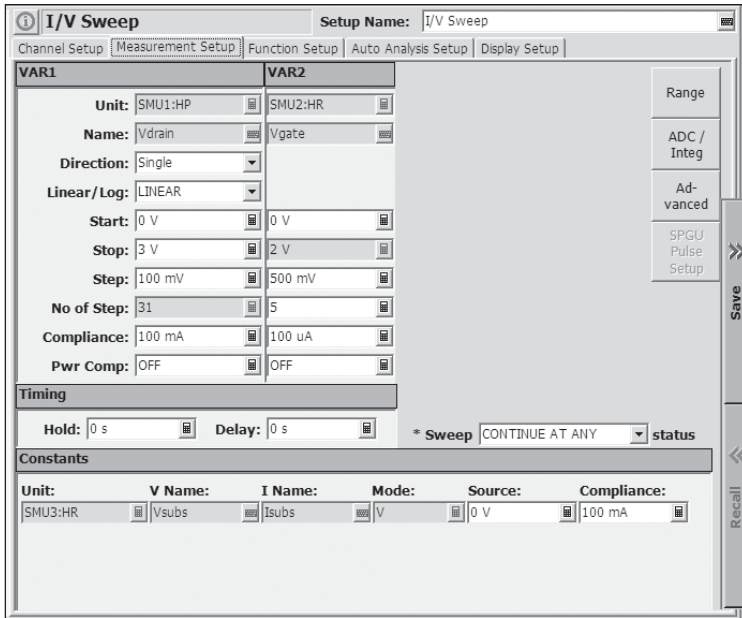
Figure 3.33. Setting up a subordinate sweep in classic test mode.

The instrument will perform a sweep measurement with the subordinate sweep variable at the "start" value, and then it will increment the subordinate sweep variable by the "step" value and repeat the sweep measurement. This process of incrementing by the step value and repeating the sweep measurement continues until the specified number of steps has been reached. **Note:** The 4145A/B referred to this variable as VAR2 (for variable #2), and the 4155C, 4156C and B1500A/B1505A (in classic mode) also use this same terminology for consistency.
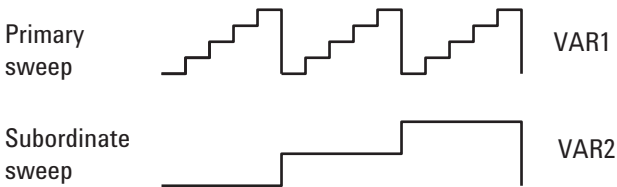

Figure 3.34. The subordinate sweep function causes the primary sweep SMU to repeat each time that the subordinate sweep variable is incremented.

The subordinate sweep function is useful for obtaining a family of curves, such as the Id-Vd characteristics of a MOSFET as the gate voltage is varied over a range as shown below.
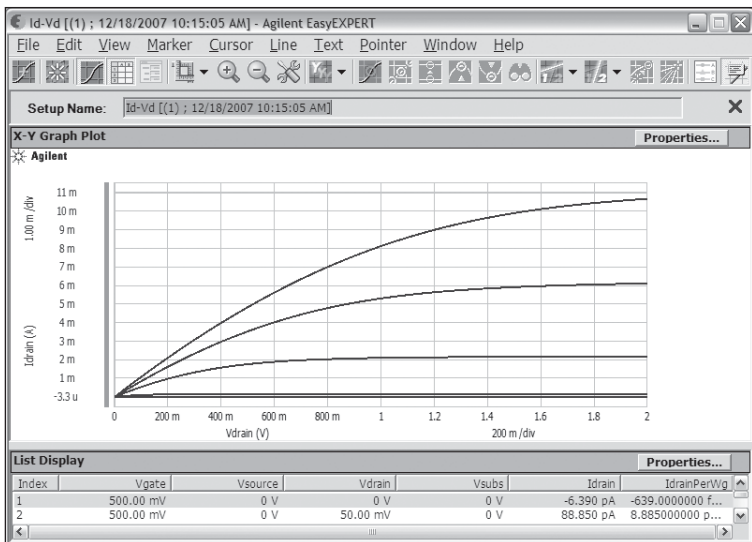
Figure 3.35. The subordinate sweep function can be used to create a family of Id-Vd curves.

Synchronized sweep measurement

The synchronized sweep function allows you to set up an SMU to track the primary sweep source by a user-specified ratio. For a synchronized sweep measurement you must specify two additional parameters: a ratio and an offset value as shown below.
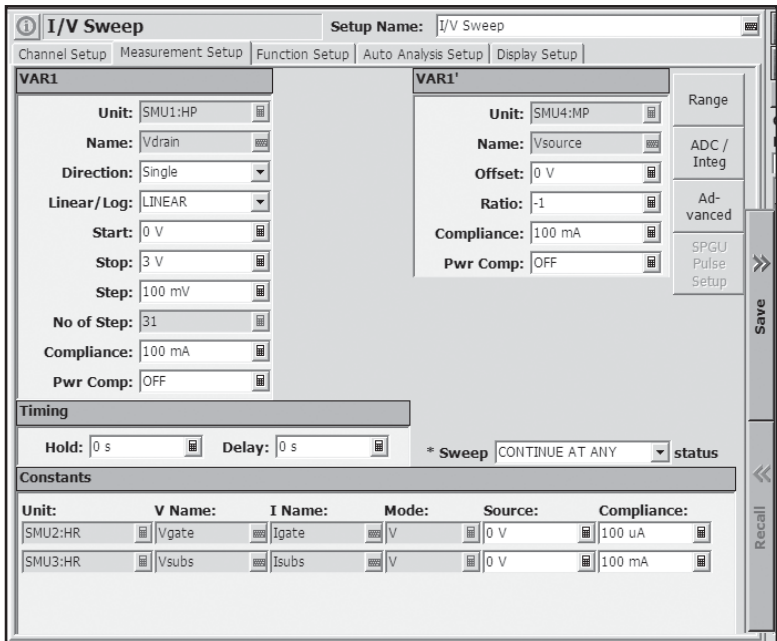
Figure 3.36. The synchronized sweep function allows one SMU to track the primary sweep source by a user-defined ratio.

The instrument will perform a sweep measurement with the synchronized sweep variable tracking the primary sweep variable by an amount determined by the "ratio". The "offset" parameter also allows for the synchronized sweep variable to have a DC offset from the primary sweep variable. The 4145A/B referred to this variable as VAR1′ (the "prime" of the first variable), and the 4155C, 4156C and B1500A/B1505A (in classic mode) also use this same terminology for consistency.

Primary sweep   VAR1

Synchronous sweep   VAR1'
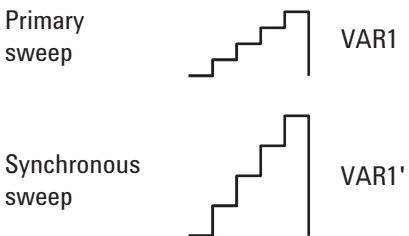
Figure 3.37. The synchronized sweep function allows one SMU to track the primary sweep source by a user-defined ratio.

Although the synchronized sweep function has many uses, two important uses are placing SMUs in series and parallel (as will be discussed shortly).

Multi-channel sweep measurement

Some instruments (such as the B1500A and B1505A) support a true multi-channel measurement sweep capability. An example of the B1500A multi-channel sweep measurement feature is shown below.
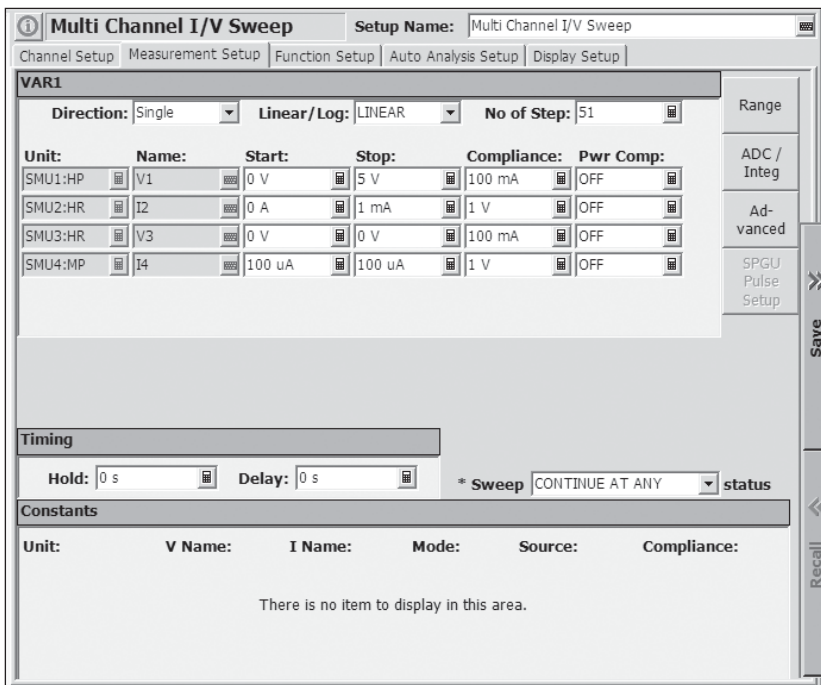
Figure 3.38. The multi-channel sweep function allows all of the available SMUs to be swept simultaneously.

The key points of the multi-channel sweep feature are:

1. You can independently select each swept source to be in either voltage or current force mode.

2. The start and stop values are independent for each swept source, but the number of points in the sweep has to be the same for each source.

3. You can specify the start and stop values to be the same, which effectively makes the swept source output a constant value (the use of this will be explained later).

4. As with the standard sweep function, you can specify power compliance for each swept source.

5. The subordinate sweep (VAR2) feature is supported, but the synchronous sweep (VAR1') feature is not supported.

List sweep measurement

The Agilent B1500A and B1505A support a list sweep capability. This feature allows the measurement points in a sweep to be defined by an arbitrary vector list created in a spreadsheet type of format. An example of the B1500A/B1505A list sweep function is shown below.
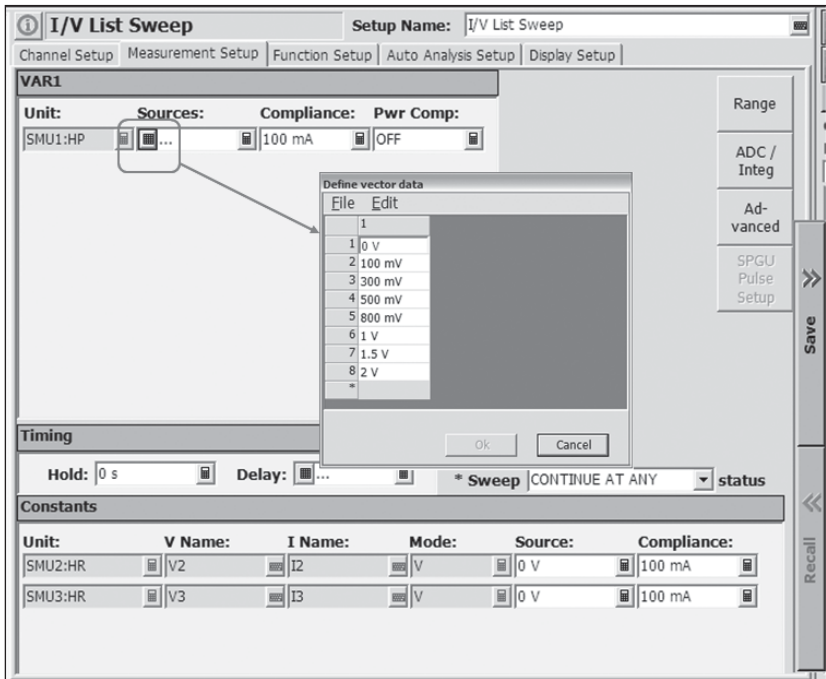
Figure 3.39. The list sweep feature allows the swept variable to be defined via a spreadsheet style format rather than in a linear fashion.

The list sweep function adds a great deal of flexibility to the basic sweep function. Instead of being forced to use constant sweep steps (either linear or logarithmic) as in the case of the standard sweep function, list sweep permits you to specify the points completely arbitrarily. One possible use of this feature would be a MOSFET subthreshold leakage measurement, where you could specify the initial values for the gate voltage sweep very close together and then spread them further apart once you have exited the subthreshold leakage region. You can create the vector list within applications such as MS Excel and cut-and-paste them into the vector data within EasyEXPERT. When used programmatically within an EasyEXPERT application test, you can also insert the name of a vector variable as the "source" for the list sweep.

Combining SMUs in series and parallel

It is possible to combine SMUs in both series and parallel to obtain more voltage and current than is possible using a single SMU. However, the methodology for doing this is not always simple or obvious.

Combining SMUs in series
Although SMUs are single-ended devices (one end is always tied to ground), by having one SMU track another SMU via the synchronized sweep function (VAR1′) it is possible to double the voltage applied to a DUT (just as if the SMUs were actually in series). The following figure illustrates how this can be achieved
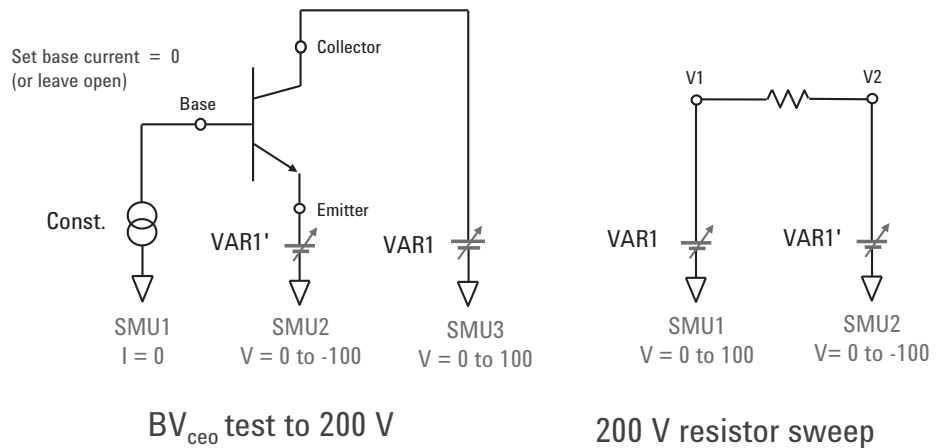
Figure 3.40. By using the VAR1′ function SMUs can effectively be placed in series.

Although the above example is for medium power SMUs and high resolution SMUs (which both have a maximum voltage output of $\pm$100 V), high-power SMUs (which have a maximum voltage output of $\pm$200 V) can achieve differential voltages of up to 400 V.

Combining SMUs in parallel

Connecting SMUs in parallel in current force mode is relatively trivial. However, the usefulness of this procedure is limited. In most cases we want to force a voltage, and we want to place SMUs in parallel in order to improve the current sourcing capability of our voltage source as shown below.
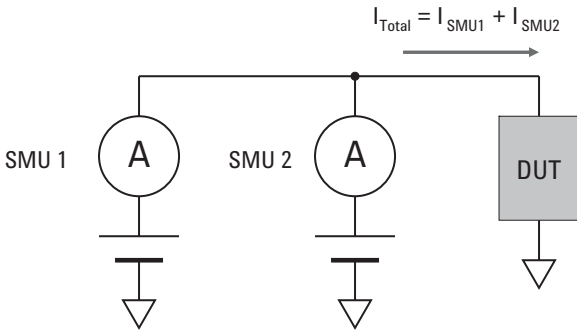
Figure 3.41. By placing two SMUs in voltage force mode in parallel, we can increase the total amount of current available.

Let us examine the general case of placing two SMUs in parallel in voltage force mode with a Kelvin configuration.
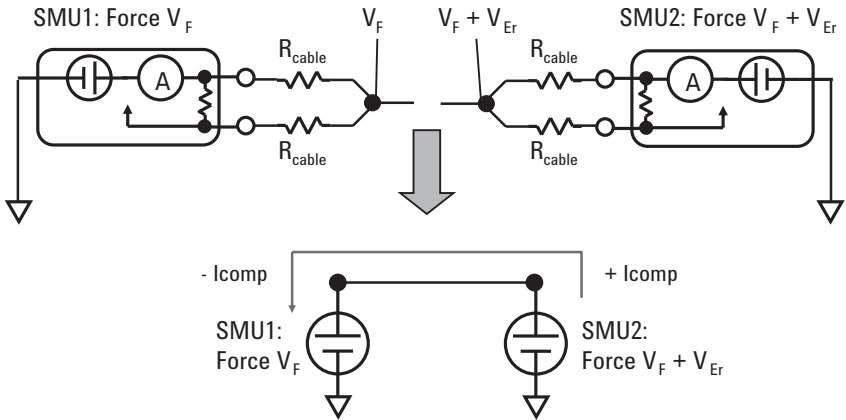
Figure 3.42. When placing SMUs in parallel, even a slight forcing voltage difference (error) can cause one or both SMUs to hit their current compliance limit.

The problem with this approach is that, even if you specify the exact same voltage for both SMUs, in practice there will be some voltage force error between the SMUs. This will cause one SMU to source current into the other SMU and very quickly one or both SMUs will hit their current compliance limit.

To prevent this situation, we can create a "quasi-Kelvin" configuration using two small resistors. These resistors limit the current flow to keep the SMUs from hitting compliance.

Insert small resistors (R_{Q1} & R_{Q2}) between the load and the point where the sense and force lines are connected.
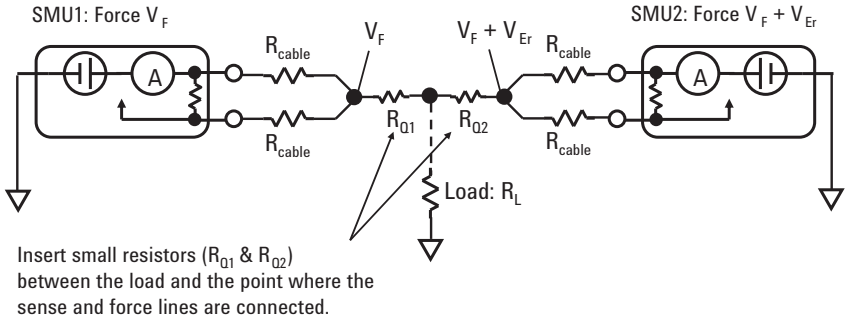
Figure 3.43. By inserting small resistors between the load and the point where the force and sense line are connected, we can prevent the SMUs from hitting current compliance.

A worst-case scenario for this situation is two high-power SMUs placed in parallel each outputting 1 A of current. We know that the 20 V range is the maximum measurement range in which we can output 1 A of current, and that in this measurement range the maximum voltage force error is 10 mV.

Maximum force voltage error in 20 V range: $\pm 0.03\% \times 20\ V + 4\ mV = \pm 10\ mV$
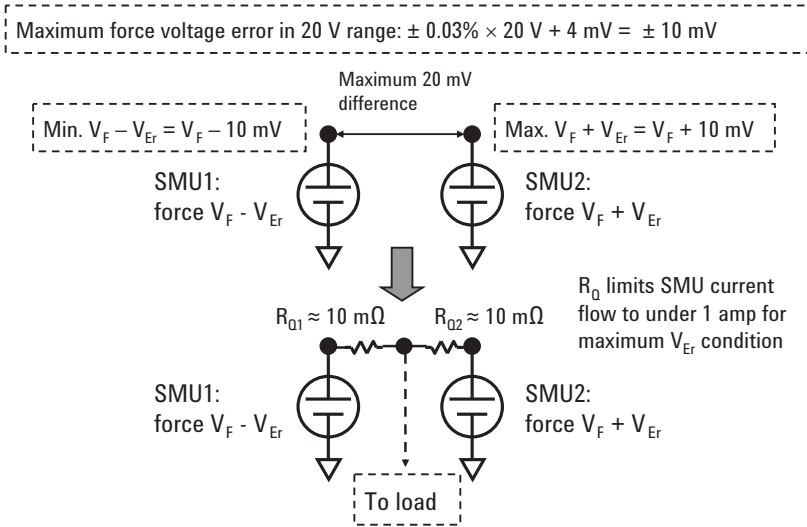
Figure 3.44. Calculating the required value of the resistors for this quasi-Kelvin connection from SMU specifications.

From these calculations we can see that a 10 mΩ value for R_Q is sufficient to compensate for the 10 mV error for each SMU. In many cases a small piece of wire can be used to create the 10 mΩ resistor.

10 cm wire $\approx$ 10 mΩ
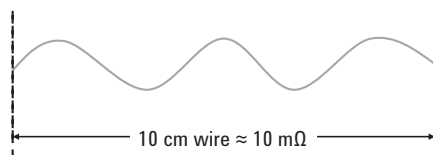
Figure 3.45. One simple way to create a 10 mΩ resistor is to use a small piece of wire.

Safety issues

Even medium power SMUs can output voltages and currents that can be lethal. Therefore, it is important to understand the precautions that must be observed to insure a safe parametric measurement environment.

Interlock

To prevent accidental electrocution, all Agilent parametric instruments have an interlock connection located in the rear of the instrument. The interlock prevents the instrument from sourcing more than 42 V unless a safety interlock is in-place. The location of this interlock on the B1500A (and B1505A) is shown below.
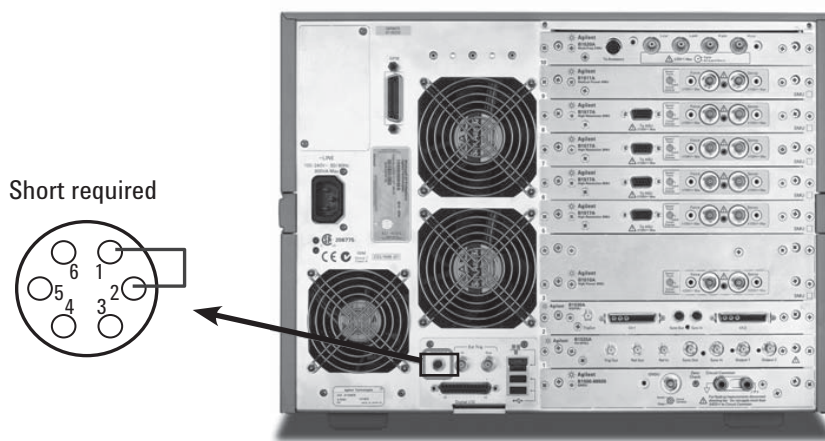
Figure 3.46. Rear view of the B1500A showing the location of the safety interlock.

In order for any of the SMUs to supply more than 42 V, a short between pins 1 and 2 must be made. DO NOT ATTEMPT TO DEFEAT THIS SCHEME BY SHORTING THESE PINS TOGETHER AS AN EXTREME SAFETY HAZARD WILL BE CREATED. The rear interlock should be used only with an approved interlock cable and equipment designed to work with an interlock cable.

If you plan on using your parametric instrument with a semiautomatic wafer prober, then you should consult with your wafer prober supplier for the proper accessories to support safety interlock scheme. All major semiautomatic wafer prober companies can supply solutions that are compatible with this scheme. In addition, if you want to test packaged devices at up to 1 A and 200 V then Agilent Technologies can supply the 16442B test fixture for this purpose. The 16442B supports the Agilent 16493J interlock cable, which requires that the lid of the 16442B be closed in order for voltages greater than 42 V to be sourced. A picture of this test fixture is shown below.

Figure 3.47. The Agilent 16442B Test Fixture supports a variety of socket types for testing packaged devices and also has a built-in interlock.

Agilent can also supply a high current and high voltage test fixture for use with the B1505A Power Device Analyzer/Curve Tracer. This standard test fixture supports component testing at up to 20 A and 3000 V.

Figure 3.48. The Agilent N1259A High-Power Test Fixture provides a safe and supported solution for testing power devices at up to 20 A and 3000 V.

The N1259A has a modular configuration that supports a variety of different package types, including universal and PTFE board socket modules that can be used to create solutions for custom package types. The N1259A also has a safety interlock that prevents measurements at dangerous voltage levels unless the lid is closed.

Module selector unit (B1505A)

Overview

The B1505A supports both a high-voltage SMU (HVSMU) and up to two high-current SMUs (HCSMUs). The HVSMU module has a high-voltage triaxial output and the HCSMU module has a unique configuration consisting of triaxial and BNC outputs (which was explained earlier in this chapter). Of course, the high-power SMU (HPSMU) has two standard triaxial outputs (force and sense). The incompatibility between these three modules can make it tedious to switch between high-voltage, high-current and low-level measurements due to all of the recabling involved. To solve this issue and make it easy to switch between these various modules, Agilent can supply a module selector unit. A schematic of the module selector unit is shown below.

Module selector unit
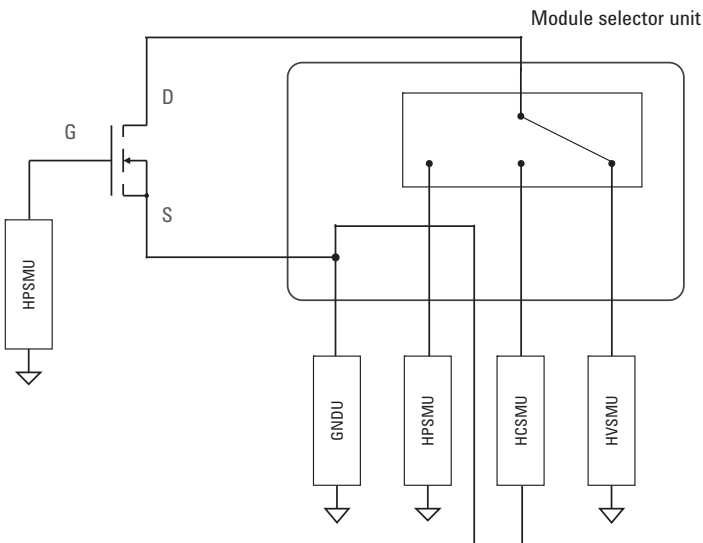
Figure 3.49. The module selector unit allows you to switch between standard, high-voltage and high-current measurements without having to change any cabling.

The module selector unit is connected to the B1505A via its digital I/O interface, and the B1505A's EasyEXPERT software automatically takes care of connecting the correct measurement resources.

The module selector unit comes in
two versions: one for use with the
N1259A test fixture and one for use
with a wafer prober. The packaged
device version of the module selector
unit (N1259A-300) mounts under-
neath the N1259A high power test
fixture as shown below.

*Figure 3.50. The N1259A high power test
fixture mounts on top of the N1259A-300
module selector unit for packaged device
testing.*

This configuration allows you to
switch between high-voltage and
high-current measurements on
packaged devices without having to
change any of the device connections
or cabling. The module selector unit
for use with a wafer prober is quite
different, and it has a different part
number (N1258A). The wafer prober
version of the module selector unit
has two high-voltage triaxial outputs
(force and sense high) and two BNC
outputs (force and sense low). The
correct cabling and adapters neces-
sary to support the N1258A on a high
power wafer prober will be discussed
when we talk about on-wafer para-
metric measurement in Chapter 4.

Chapter 4: On-Wafer Parametric Measurement

"One must learn by doing the thing; for though you think you know it, you have no certainty until you try." — Sophocles

Introduction

The vast majority of parametric measurements are performed on-wafer. Obviously, this implies that you need some sort of a wafer prober to make these measurements. What is often missed by many engineers performing on-wafer parametric measurements are the profound effects that various factors such as the wafer chuck, cables, wafer probes and switching matrix (if used) can have upon their measurement results.

Fully automatic versus analytical wafer probers
Fully automatic wafer probers are designed to be used in production environments in conjunction with either parametric or functional test equipment. Fully automatic wafer probers can test an entire cassette or FOUP (front-opening unified pod) of wafers at a time, and they have the ability to automatically load and align a wafer for testing. Fully automatic wafer probers are virtually always used with some sort of a probe card and switching matrix. Due to their size, cost, and complexity, fully automatic wafer probers are used almost exclusively in production test environments with high-volume production testers. Since this handbook is primarily focused on parametric test instruments, we will not spend any additional space discussing them.

Analytical wafer probers are designed to be used in laboratory environments, and they come in both manual and semi-automatic versions. Manual wafer probers do not possess any electrical motors to automate any part of the wafer alignment and stepping process, but must instead (as their name implies) be manually adjusted each time that you want to move to a new location on the wafer. In contrast, most semi-automatic wafer probers possess the capability to automatically align an individual wafer; however, they typically do not support any sort of automated wafer loading and they require the manual loading of each wafer. Semi-automatic wafer probers can be used with probe cards and switching matrices, but it is more common to see them used with multiple individual positioners (probes) for maximum measurement flexibility. Most semi-automatic wafer probers also do support automated testing across an entire wafer in conjunction with some sort of parametric instrumentation.

Figure 4.1. Some examples of semi-automatic wafer probers.

Wafer prober measurement concerns

It should be clear upon some reflection that the wafer prober measurement environment has a significant impact on the quality of parametric measurements that you can obtain. Unfortunately, many times these effects are not adequately accounted for by users trying to make parametric measurements. Some of the key issues are as follows.

1. Limited current noise floor – An inadequately shielded wafer prober will greatly increase the noise floor of the measurement environment.

2. Slow measurements – Capacitance and piezoelectric effects can significantly impact the speed at which parametric measurements can be made.

3. Low and high temperature measurements – Low and high temperature measurements both present issues in terms of moisture and noise.

Chuck isolation
The basic rules for proper guarding and shielding discussed in Chapter 2 take on great significance when applied to the wafer probing environment. It is important to understand that a wafer chuck is an extremely large capacitor and that it can also act as a large antenna (collecting random noise). Therefore, optimal measurement performance is obtained through the use of a guarded Kelvin wafer chuck as shown below.
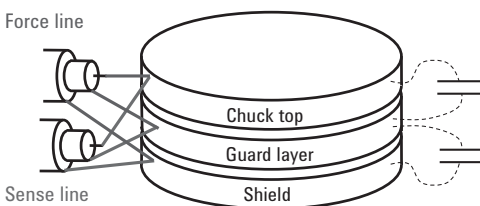
Figure 4.2. A guarded Kelvin wafer chuck can eliminate the effects of the parasitic chuck capacitance and greatly reduce the leakage from the chuck to the outside environment.

By using a guard layer in the wafer chuck, the effects of the parasitic chuck capacitance and the leakage currents through the wafer chuck can be virtually eliminated. The following plot compares the settling time for both guarded and unguarded wafer chucks, and it is clear that for the case of the guarded chuck the necessary wait time before a measurement can begin is greatly reduced.
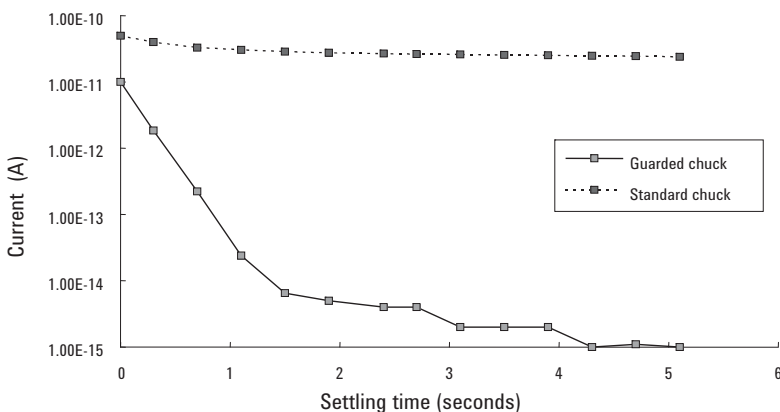
Figure 4.3. Plot comparing the settling time for both guarded and unguarded (standard) wafer chucks.

In addition to guarding the wafer chuck, if the guard can also be placed above the wafer being measured then noise can be reduced further and an optimal low-noise measurement environment is guaranteed. Such a measurement scheme is shown below.
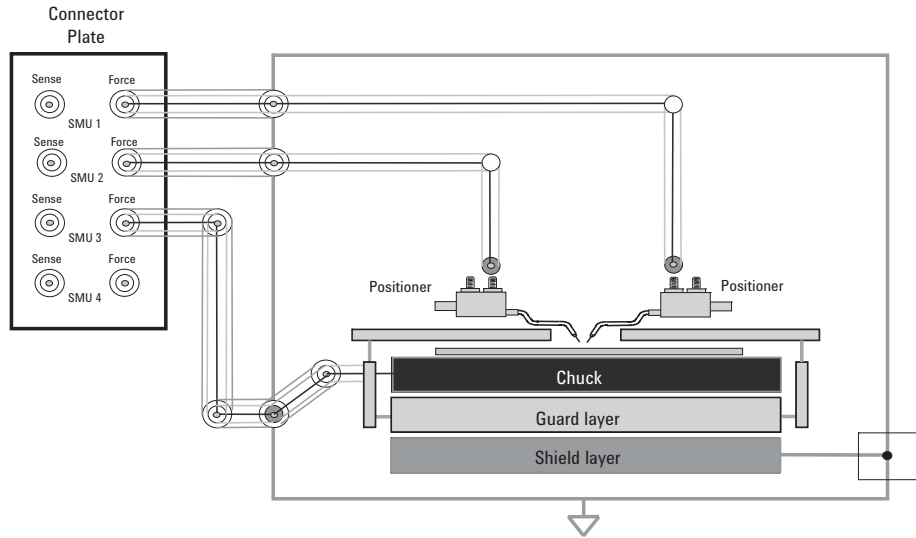
Figure 4.4. A completely guarded and shielded wafer prober environment (Note: Scheme shown is Cascade Microtech's patented AttoGuard® technology).

Low and high temperature measurement issues
Performing parametric measurements at low and high temperatures on-wafer presents the following measurement challenges.

1. High electrical noise caused by the thermal control circuitry.
2. Slow measurement times due to parasitic capacitance.
3. Large transient noise after wafer chuck moves.
4. Frost induced moisture leakage.

We will proceed to examine each one of these in-turn.

The electrical noise generated by the thermal control circuitry is probably the single greatest source of error when making low-level measurements on-wafer at temperature. While it is not possible to completely eliminate all of the effects of this circuitry, many analytical wafer prober companies have mitigated them to a great extent. The graph shown below shows the noise current associated with a standard triaxial wafer chuck as the chuck temperature varies over time.

Standard triaxial chuck
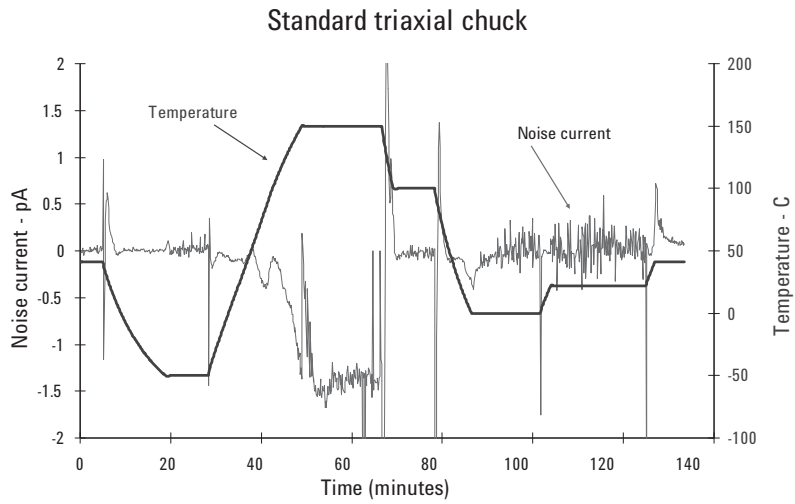
Figure 4.5. Noise current for a standard triaxial wafer chuck as the temperature varies over time.

In contrast, the following graph shows the noise current for a well-designed thermal wafer chuck.

Low noise thermal chuck
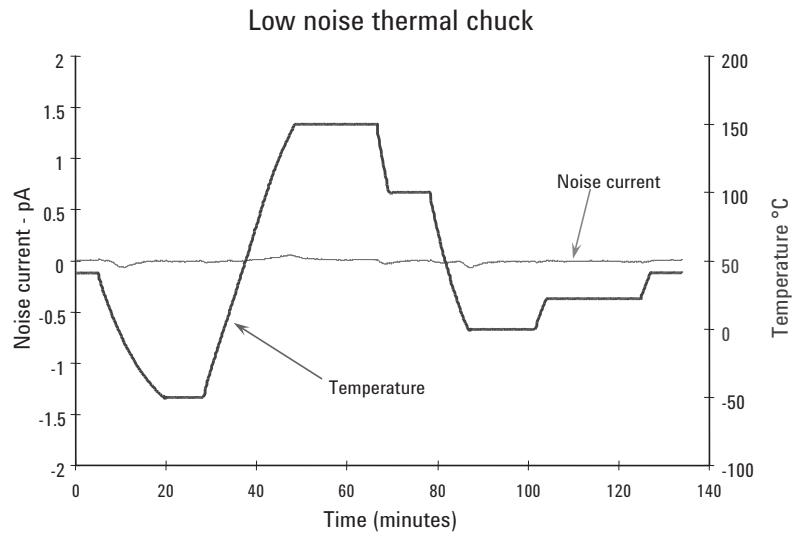
Figure 4.6. Noise current for a well-designed low-noise thermal wafer chuck.

In the end you get what you pay for, so it does not make sense to try and make low-level measurements at high or low temperatures on anything other than a wafer chuck specifically designed for that purpose. Before purchasing any analytical wafer prober with a thermal chuck option you should very carefully evaluate its low-current noise performance to make sure that it meets the requirements of your parametric measurements.

In addition to reducing noise on the wafer chuck, it is important to reduce noise and capacitance effects caused by the probes. Ceramic bladed probe tips are the best solution for making on-wafer measurements at high temperatures. Bladed probe tips can easily handle temperatures of up to 300 degrees Celsius, and they have minimal residual capacitance. A picture of a ceramic bladed probe tip for positioner-based wafer probing is shown below.
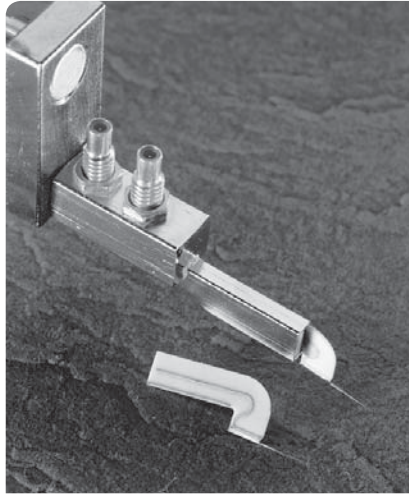
Figure 4.7. Ceramic bladed probe tips are the best solution for on-wafer high-temperature measurement. [Note: Photo courtesy of Cascade Microtech]

Note that these types of probe tips are also easily replaced if they become damaged.

The same type of methodology should be implemented when using probe card schemes (which will be discussed later in this chapter). However, in the case of probe cards the juxtaposition of many signal lines close together adds some design complexity. The key is to shield the entire probe card within a metal enclosure to isolate the signal lines as much as possible from thermal noise. An example of a fully guarded ceramic probe card for high temperature measurement is shown below.

Figure 4.8. A ceramic bladed probe card with full shielding that works from −65 °C to +300 °C. Note: This is a picture of Cascade Microtech's Attofast™ probe card.

On a semiautomatic wafer prober the electronics and cable connections used to move the wafer chuck and to control its temperature have a strong impact on low-noise measurements, and these are exacerbated at higher temperatures. There is no single "magic bullet" that can completely eliminate these noise sources, but sound design principles can mitigate them. Some key points to look for in a wafer prober are listed below.

1. Low triboelectric effect materials to minimize the electronic charge generated via friction.
2. Low dielectric absorption materials to minimize the residual capacitive charge.
3. Low noise design through proper shielding.

The final issue to consider when making on-wafer low temperature parametric measurements is frost, which is related to the dew point. Moisture degrades parametric measurement significantly, and can cause high residual capacitance, elevated leakage currents, and increased noise. For this reason, it is often recommended to "bake" a wafer chuck at +200 °C for twenty-four hours prior to making any at temperature parametric measurements. In any case, low temperature measurements require an atmospheric flow (using either clean dry air or nitrogen) to prevent frost from forming. However, as the graph shown below indicates, increasing the air flow also increases the noise.
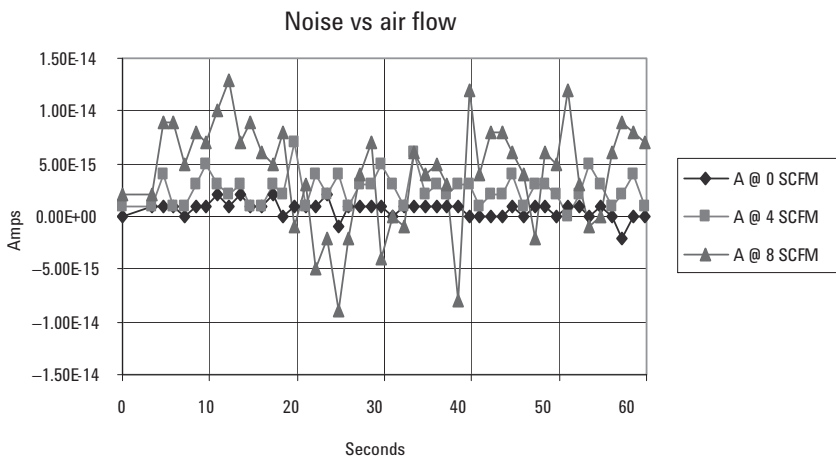
Figure 4.9. The effect of air flow on noise.

Therefore, the best probing systems are those that can achieve (for example) frost-free operation at –55 °C with minimal air flow. ***Note:*** An SCFM flow of less than 1 is probably sufficient for most measurements.

DC and RF wafer probes

For analytical wafer probers, there are two distinct types of positioners: DC and RF. Although many people might assume that RF probes are only required for true RF measurements made in the GHz range (such as S-parameter extraction), this belief is not correct. Capacitance measurements made above 5 MHz may require RF probes in order to achieve satisfactory measurement results, and many high-speed pulsed measurements require RF probes to avoid creating ringing on the pulse edges. Therefore, it is important to understand the differences and limitations of both DC and RF probes.

There are several key points to keep in mind when designing test structures for RF wafer probes.

1. RF wafer probes are typically mounted in fixed positions on the wafer prober (180 degrees opposite to each other), so whatever structures you place on the wafer must adhere to this limitation.
2. RF wafer probes primarily come in ground-signal (GS) or ground-signal-ground (GSG) configurations, and they are available in only certain pad pitches (typically ranging from 50 μm to 250 μm).
3. The probes themselves should be separated by at least 200 μm to avoid crosstalk.
4. All grounds should be connected together.
5. Any biasing or control pads should be completely out of the main signal path.
6. RF probes require BNC cabling and usually use an SMA style connector.

A generic example of a test structure adhering to these guidelines is shown below.
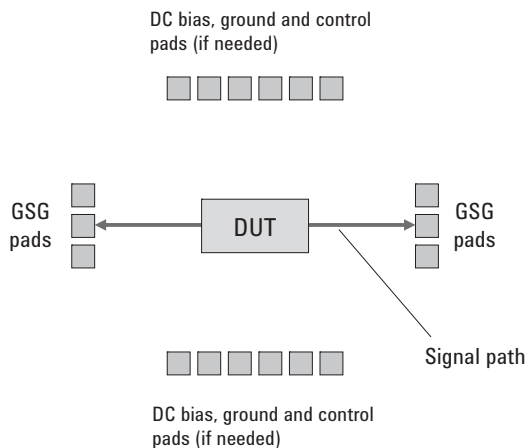
Figure 4.10. The proper way to layout a structure for use with RF positioners.

A picture of RF probes used for on-wafer measurement is shown below.
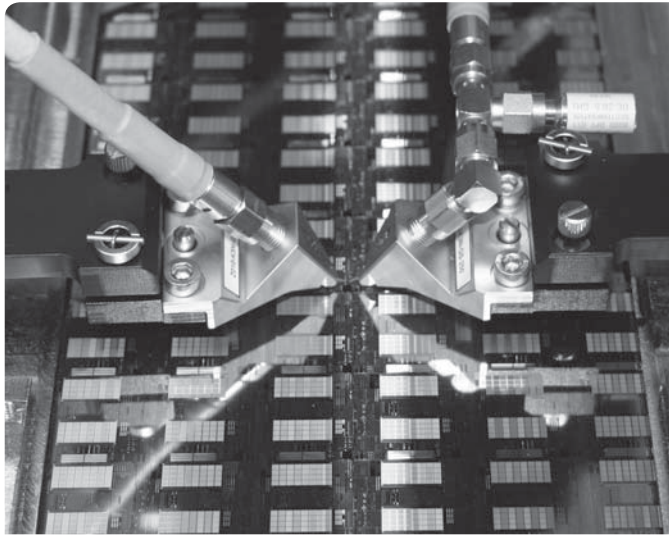
Figure 4.11. An example of on-wafer measurement using RF wafer probes.
[Note: Photo courtesy of SUSS Microtech]

RF positioner device layout and design will be discussed further when we cover high-speed pulsed measurement in Chapter 5 and high-frequency capacitance measurement in Chapter 8. Note that the RF probes use SMA connectors, so BNC to SMA connector adapters will be required if you are trying to connect up instruments such as capacitance meters.

Switching matrices

Introduction

Although switching matrices do not necessarily have to be used in conjunction with on-wafer semiconductor measurements, it is very rare to see them used outside of this type of a measurement environment. When making on-wafer measurements, a switching matrix can both facilitate faster measurement and simplify the connection environment (for example switching between IV and CV measurements). A photo of a switching matrix for use in parametric test is shown below.

Figure 4.12. The Agilent B2200A femtoamp low leakage switch.

The basic concept of a switching matrix is not hard to understand. A switching matrix consists of a series of inputs and outputs; connections between the inputs and outputs can be made by closing one or more relays. Since virtually all connections in parametric test are made with triaxial cables, almost all switching matrices convert some number of coaxial inputs into triaxial outputs. While the number of inputs on a switching matrix is fixed, the number of outputs is usually determined by the number of cards installed in that matrix mainframe. Each card will typically add some number of outputs to the configuration. An example of a switching matrix card is shown below.
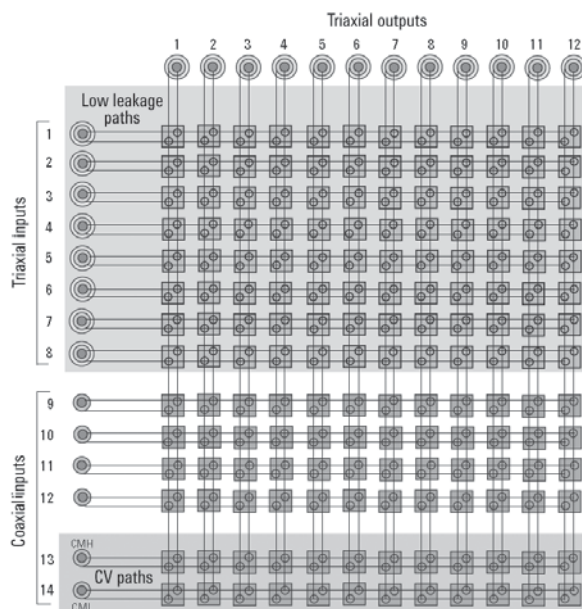
Figure 4.13. An example of a switching matrix card with both triaxial and coaxial inputs (as well as a dedicated path for CV measurement)

There are some key takeaways from the schematic just shown. The first of these is that not all triaxial switching matrix paths are necessarily "low-current". In the above schematic this is true, but this is not always the case for all switching matrix cards. The second observation is that a switching matrix converts all inputs (triaxial and BNC) into triaxial outputs. This is extremely convenient since eliminates the need to worry about triaxial to BNC adapters. The third (and final) observation is that some matrix cards have a dedicated path for capacitance measurement. However, we will defer discussion of making CV measurements through a switching matrix until we reach Chapter 8 and discuss capacitance measurements.

Probe cards and module testing

In order to understand the benefits of a switching matrix when making on-wafer measurements one must first understand how parametric test structures are organized on the wafer. Parametric test structures are typically organized into modules with some sort of regular pad organization (such as two rows of twelve pads). When used in conjunction with a probe card designed to fit the module pad configuration, a switching matrix permits the testing of all of the devices in that module without the need to physically move the probes. By using the switching matrix to connect and disconnect the DUTs from the parametric measurement resources, the time associated with physically moving the probes from one device to the next is eliminated. The following figure shows a photo of probe card tips contacting a test module on a wafer.
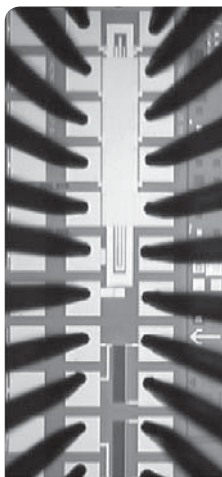
Figure 4.14. A photo showing a probe card contacting a module on-wafer.

Even in a semi-automatic wafer prober environment, the use of a probe card can significantly improve test throughput.

It should be obvious that in order to take full advantage of a switching matrix and probe card solution, some sort of software test shell is necessary to control everything. The key requirements of this software are shown below:

1. The ability to coordinate testing with the wafer map stored on the semi-automatic wafer prober.
2. The ability to specify sub-die moves so that different modules within a die can be tested.
3. The ability to store and recall switching matrix settings to connect to a particular device to be tested.
4. The ability to store measurement results for each device tested along with information as to the exact location (both die and sub-die) of that device on the wafer.

Agilent Technologies' EasyEXPERT software possesses all of these capabilities and can be used to automate the testing across a semiautomatic wafer prober.

In order to perform automated wafer testing using Agilent EasyEXPERT you must first set up a wafer map on the semiautomatic analytical wafer prober that you plan on using. An example of this is shown below.
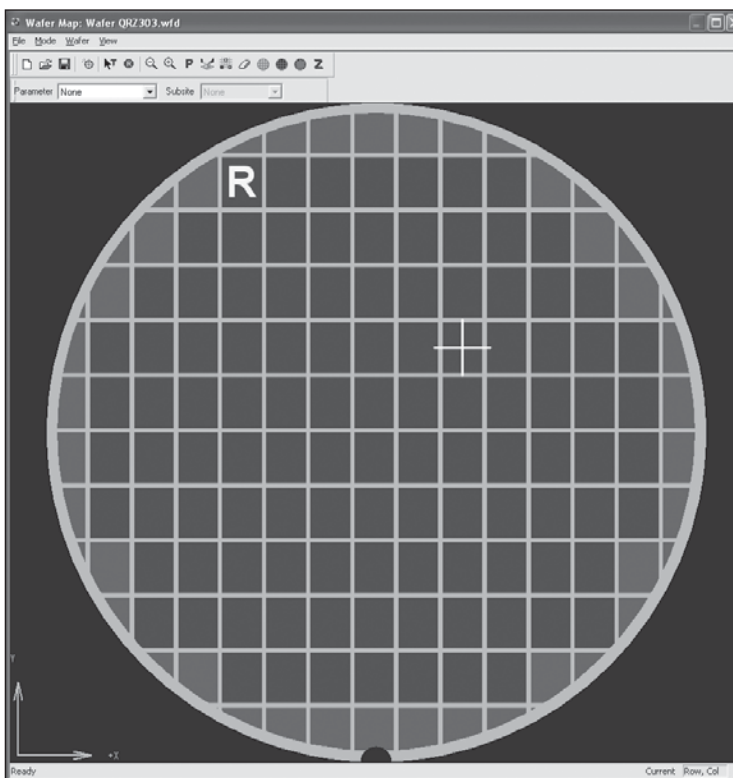
Figure 4.15. Setting up a wafer map on a semiautomatic wafer prober in preparation for automated wafer probing.

Once the wafer map is set up you can start the automated testing by calling up the "Repeat Measurement Setup" window by clicking on the repeat button in the upper right corner of the main EasyEXPERT screen.
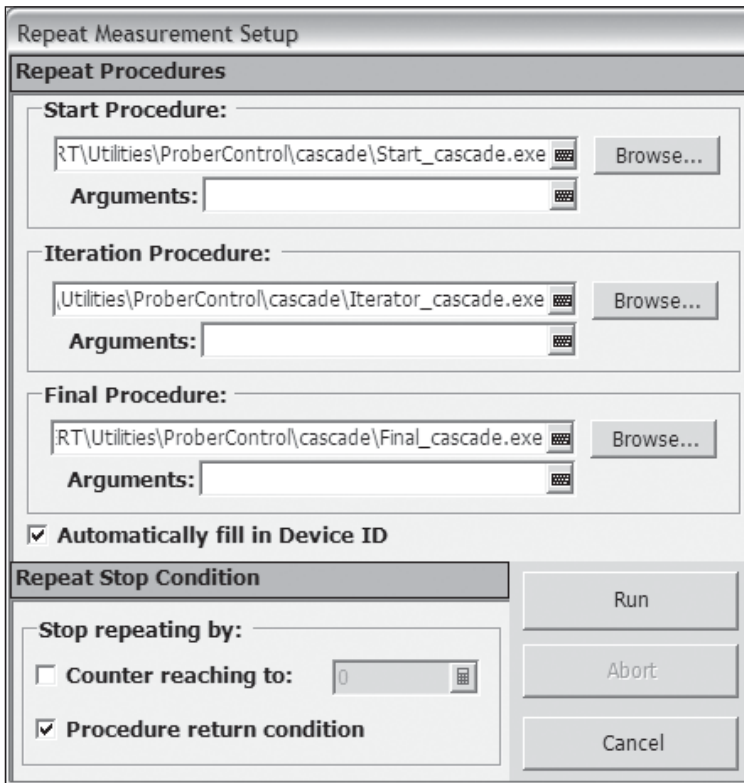
Figure 4.16. The Repeat Measurement Setup screen allows you to set up the conditions to automatically step and repeat one or more measurements across a wafer.

Note: You must have the data auto-record feature turned on in order use this feature.

EasyEXPERT has built-in wafer prober drivers for Cascade, SUSS and Vector semiautomatic wafer probers. In addition, sample source code is available so that you can create your own wafer prober drivers if necessary. The prober driver files can be found in the following location: Program Files\Agilent\ B1500A\EasyEXPERT\Utilities\Prober Control. Each supported prober has a file folder, and there are several executable files located in these file folders. In the "Repeat Procedures" section of the "Repeat Measurement Setup" window the correct executable file can be selected using the "Browse" button. For each of the three lines, Start, Iteration and Final, there is a corresponding executable file (as can be seen in Figure 4.16). By default the "Automatically fill in Device ID" box is checked, which will automatically create a device ID label for each EasyEXPERT test record to keep track of which test or tests are associated with which die on the wafer. In addition, you can select a "Repeat Stop Condition" if the "Procedure return condition" is checked then the testing will continue until the wafer map reaches the last die; if the "Counter reaching to:" condition is checked then the testing will continue until the specified number of die have been tested. If both conditions are checked then testing will stop at whichever condition is reached first. After filling in the above information, simply click on the "Run" button to begin coordinated testing of the B1500A and your semi-automatic wafer prober.

Positioner based switching solutions

Introduction

Switching matrices can be a very cost-effective solution when probe cards are used. However, the vast majority of laboratory-based analytical wafer probing involves the use of individual positioners rather than probe cards. Although it is certainly possible to use a switching matrix with positioners, the cost and complexity of this type of solution may not be optimal. Fortunately, Agilent Technologies has several options for positioner-based wafer probing switching that do not require a switching matrix. The following sections will explain these available options. Note: Since RF positioners should NEVER be used with any sort of a switching solution, all further discussions in this section refer to DC positioners.

Atto-sense and switch unit

In the earlier discussion of SMUs in Chapter 3 it was mentioned that an atto-sense and switch unit (ASU) was available, and that this unit works with the HRSMU to achieve current measurement resolution down to 0.1 fA. In addition, to possessing this impressive low-current measurement capability, the ASU (as its name implies) also has a built-in switching capability. Each ASU has two BNC inputs that are nominally designed to be used with two of the outputs from a capacitance meter, although they can be used with any type of instrumentation that has BNC outputs. A schematic of a single ASU is shown below:
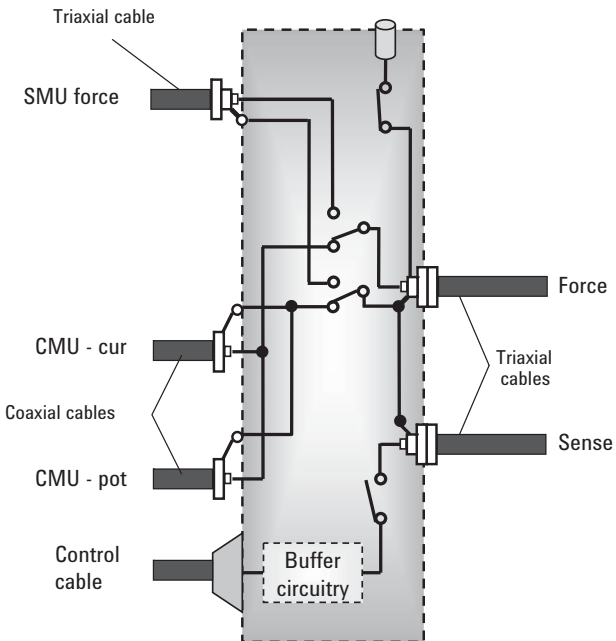
Figure 4.17. A simplified ASU schematic showing its switching capabilities.

A picture of the ASU's inputs is shown below.

Figure 4.18. The inputs to the ASU, showing the two coaxial inputs (top), the triaxial input (from the force line of the HRSMU, lower right) and the control cable input (also coming from the HRSMU, lower left).

The advantage of combining switching capability with 0.1 femtoamp measurement capability should be obvious: you can switch between IV and CV measurements on positioners without having to change any cables. In addition, although the issue of capacitance compensation will not be discussed until Chapter 8, it is sufficient to state here that the B1500A will automatically detect and compensate for the effects of the two supported cabling connections (1.5 m and 3.0 m) between the HRSMU and the ASU. This means that the only further cabling compensation required is for the additional items (triaxial cables, positioner and probe tip) that are attached to the triaxial outputs of the ASU. ***Note:*** Of course, the length of these "cable extensions" should be kept as short as possible.

SMU CMU unify unit
The ASU works well for CV-IV positioner-based wafer probing, but its 0.1 fA current measurement resolution is not always necessary. For this reason, Agilent supports two other positioner-based switching solutions for the HRSMU and MPSMU. Both of these solutions use the SMU CMU unify unit (SCUU). The SCUU is a switching module designed to work with the B1500A's MFCMU and two of either the high-resolution or medium power SMUs. The SCUU form-factor is designed to accept the four BNC outputs of the MFCMU and the two pairs of Kelvin triaxial outputs from the two HRSMUs and/or MPSMUs. The SCUU provides the capability to switch between the MFCMU and the HRSMU/MPSMU pair automatically, and it outputs are two pairs of triaxial outputs. A diagram of the SCUU showing its connections is shown below.
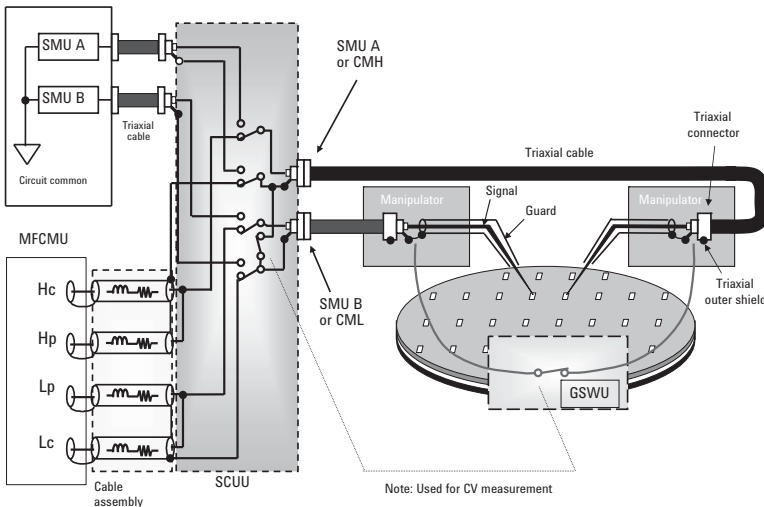
Figure 4.19. The SCUU supports switching between the MFCMU and two SMUs (HRSMU or MPSMU).

A picture of the SCUU (mounted on the rear of the B1500A) is shown below:

Figure 4.20. A picture of the SCUU mounted on the rear of the B1500A to illustrate how the MFCMU and SMU outputs are converted into two sets of triaxial outputs (force and sense).

Note: It is important to understand that the previous picture is for illustrative purposes only and that when making on-wafer measurements the SCUU should never be mounted directly on the rear of the B1500A.

The advantages of this solution for a positioner-based wafer probing environment are the same as those of the ASU; namely, you can switch between IV and CV measurements without having to change any cables. Also, as in the case of the ASUs, Agilent can supply a cable (3.0 m) that allows the SCUU to be mounted on an analytical wafer prober in close proximity to the DUT and the B1500A will automatically detect and compensate for the effects of the supported cabling connection. Just like for the ASU, the only further cabling compensation required is for the additional items (triaxial cables, positioner and probe tip) that are attached to the triaxial outputs of the SCUU. Agilent can also supply a magnetic mount for the SCUU, although most analytical wafer prober companies can also supply mounting hardware for the SCUU.

There is one other optional component of the SCUU solution that has not yet been mentioned even though it was shown in the diagram and the cable going to it was shown in the SCUU picture. It is the guard switch unit (GSWU). The GSWU can be connected to the SCUU, and its purpose is to short the guard lines of the two positioners used with the SCUU together (using an internal relay) whenever a capacitance measurement is being made. In normal current-voltage (IV) operation the GSWU does not short the guards together, which is of course what you want since the guards have to track the signal lines in order to permit low-current measurement. The reasons for shorting the guards together during CV measurement will be discussed in the next section.

Shorting the guards together during measurement
Although it is not possible to go into all of the details why shorting the guards together is important when making capacitance measurements until after we cover capacitance meter operation in Chapter 8, we can at least give a basic explanation here. In Chapter 8 the issues mentioned here will then be covered in greater depth.

In order to measure capacitance we have to measure impedance, which is a complex quantity consisting of a magnitude and a phase (or a resistance and reactance if you prefer). In addition to having resistance associated with it, all cabling also has capacitance and inductance associated with it. When we perform capacitance compensation we essentially are measuring the impedance characteristics of the cabling so that we can subtract them out from our measurement in order to arrive at the true (actual) value of the capacitor that we are trying to measure. Implicit in this scheme is the assumption that the impedance characteristics of the cables being used are stable. However, in the case of flexible cabling it turns out that this is not necessarily a valid assumption. If the guards are not shorted then any movement of the cables can cause their series inductance to vary, thereby invalidating any compensation performed and resulting in inaccurate capacitance measurements. The series inductance of coaxial cables can vary by several hundred nano-Henries (nH) per meter unless the cable guards are shorted together. The graph shown below illustrates how much measurement error can occur when the guards are not shorted during a capacitance measurement.
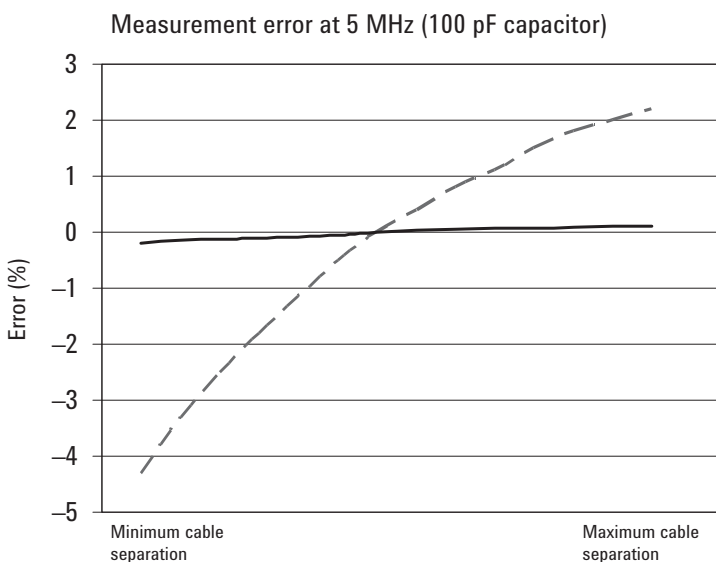
Figure 4.21. A comparison of the measurement error (%) for a 5 pF capacitor measured at 5 MHz when the guards are shorted (solid line) versus not shorted (dashed line).

In addition to the concerns with variable series inductance, there is another reason for shorting the guards together during capacitance measurement. A wafer resting on a chuck creates an LCR circuit, which has a resonant frequency. In order to short-out this circuit and prevent oscillations, you need to connect the probes together using a short cable. This is illustrated in the figure shown below.
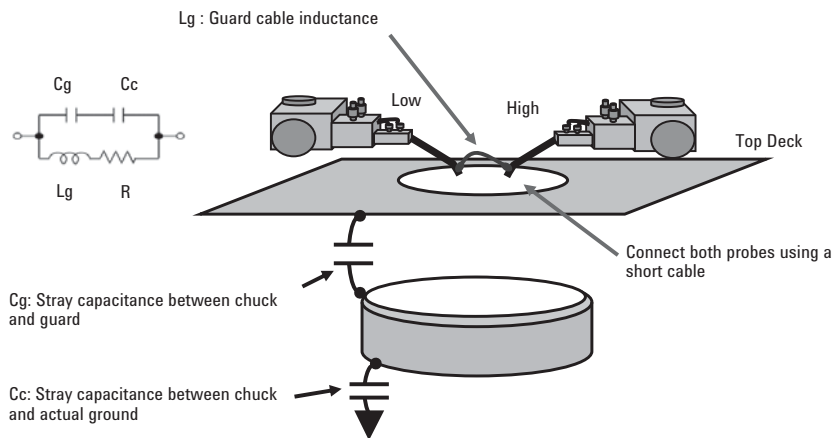
Figure 4.22. When parallel resonance occurs during a measurement, shorting the guards together will eliminate the problem.

Based on this discussion, it should be clear that we need to short the guards together during capacitance measurements for both the ASU and SCUU cases. In the case of the ASUs, Agilent supplies a shorting cable with the units for this purpose. If you are using the ASUs with the B1500A and EasyEXPERT software, then the software will automatically take care of shorting the guards together whenever you make a capacitance measurement. A simplified schematic showing the ASU connections for making CV-IV measurements is shown below:
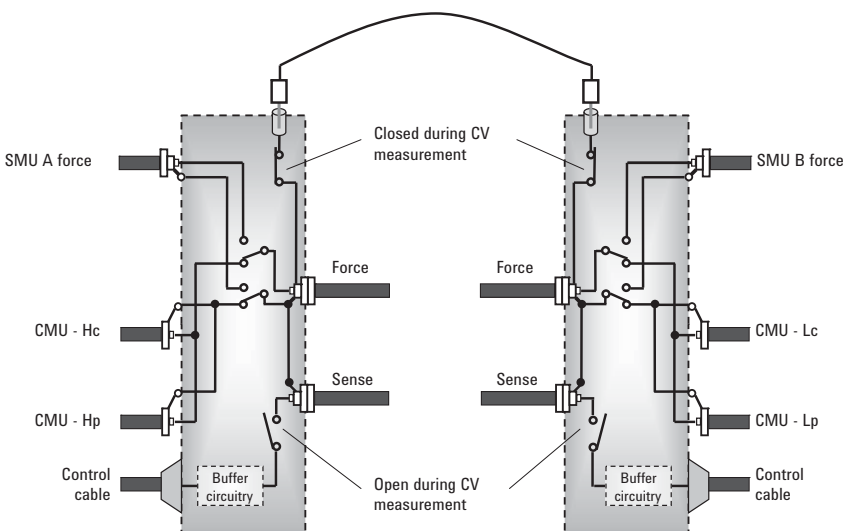
Figure 4.23. Simplified schematic showing the ASU connections when performing both CV and IV measurements.

A photo of two ASUs mounted on an analytical wafer prober with their guards tied together is shown below.
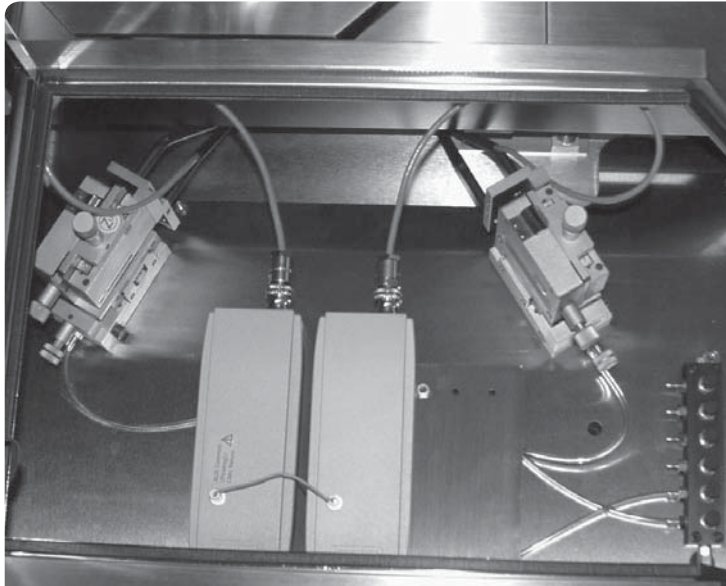
Figure 4.24. A picture showing the ASU guard connection as implemented on an analytical wafer prober. [Note: Photo courtesy of SUSS Microtech]

In the case of the SCUU, you need the optional GSWU in order to short the guards together during CV measurements. As in the case of the ASU, EasyEXPERT software automatically takes care of shorting the guards together during a CV measurement (and an LED on the GSWU will light up when this is occurring). A photo showing the GSWU mounted on a wafer prober and connected to the guards of two positioners is shown below.
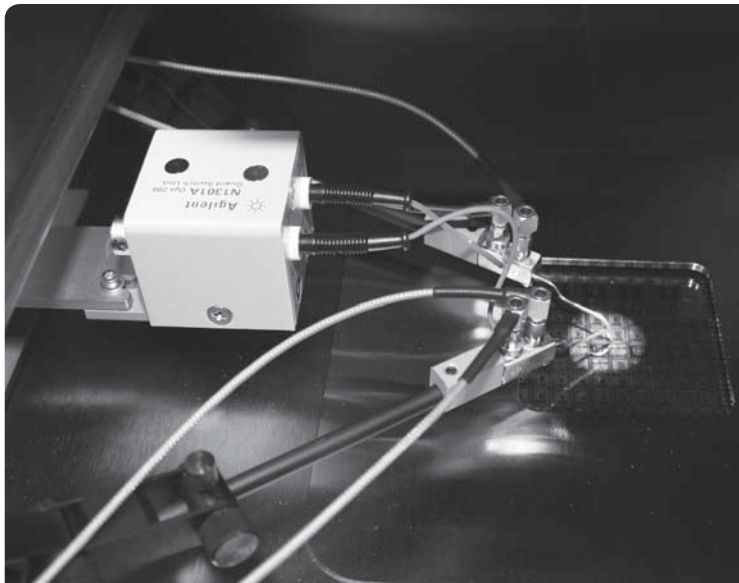
Figure 4.25. Picture showing the GSWU mounted on an analytical wafer prober and connected to the guards of two positioners. [Note: Photo courtesy of SUSS Microtech]

Summary of Agilent positioner-based switching solutions

The following table summarized the three Agilent positioner-based switching solutions. You can combine the B1500A's MFCMU and MPSMUs or HRSMUs with the SCUU or ASUs to create positioner-based CV-IV switching solutions to meet a variety of needs. All solutions are Kelvin, they support CV-IV switching, and they have integrated capacitance compensation.

10 fA/0.5 µV solution	1 fA/0.5 µV solution	0.1 fA/0.5 µV solution
1 × MFCMU 2 × MPSMU 1 × SCUU	1 × MFCMU 2 × HRSMU 1 × SCUU	1 × MFCMU 2 × HRSMU 2 × ASU
Kelvin measurement Multi-frequency capacitance measurement CV-IV switching Integrated capacitance compensation		

Figure 4.26. Table showing the three positioner-based switching solutions supported by the B1500A.

Note: The techniques involved in making capacitance measurements through a switching matrix will be deferred to Chapter 8 after we have completed a more thorough discussion of capacitance measurement.

High voltage and high current wafer probing

Overview

To perform high-voltage and high-current wafer probing you need to use a wafer prober designed for this purpose. While this statement might seem obvious, it is sometimes forgotten or ignored by users eager to begin making high power on-wafer measurements. High power analytical wafer probers differ from standard analytical wafer probers in several key ways:

1. Carefully designed safety schemes to prevent any chance of accidental exposure to high-voltage.
2. Support for the unique connectors used to make high-voltage and high-current measurements.
3. Special positioners for high-current wafer probing.

As we will see as we progress through this section, the cabling and connectors for high-voltage and high-current wafer probing are much more complicated than for standard wafer probing. In addition, different analytical wafer prober companies use different connector schemes. This makes it extremely important to specify exactly what wafer prober company you plan on using so that Agilent can correctly configure a solution for you.

Connection considerations

The first point you need to consider is exactly what type or types of high-voltage or high-current wafer probing you want to perform. If you only plan on performing one specific type of high power wafer probing (only high-voltage or only high-current), then you probably do not need the module selector unit. However, if want to perform both high-voltage and high-current measurements (or you think you might want to do both in the future) then it is highly advisable to purchase the module selector unit. As we discussed in Chapter 3, the module selector unit allows you to switch between the HVSMU, HCSMU or HPSMU without having to change any cables. Of course you can still use the module selector unit even if only one of these modules is connected to it, making it easy to expand your B1505A's measurement capabilities in the future if you add additional modules to it.

When you order a B1505A, it includes the correct cables for all of the installed modules. While these cables are sufficient to connect to the N1259A high power test fixture for packaged device testing, you will need additional cables and connectors to interface to a wafer prober. Some of these can be supplied by Agilent, and some may need to be supplied by your wafer prober vendor. Another factor that must be taken into consideration is that some of the lower power B1505A modules (the HPSMU and the GNDU) can be damaged if exposed to the high voltages generated by the HVSMU. To avert this, Agilent can supply a variety of protection adapters that prevent accidental damage to these modules. There are several different options of the protection adapters to support the different types of connectors used the various analytical wafer prober companies. Agilent can also supply resistor boxes (R-boxes) in different resistor values (1 MΩ, 100 kΩ and 1 kΩ) to place a resistor in series with the cables that you are using. The 1 MΩ and 100 kΩ R-boxes are used to prevent device damage due to voltage spiking, while the 1 kΩ R-box can be used to suppress device oscillation. An illustration showing the most common connectors and adapters is shown below.
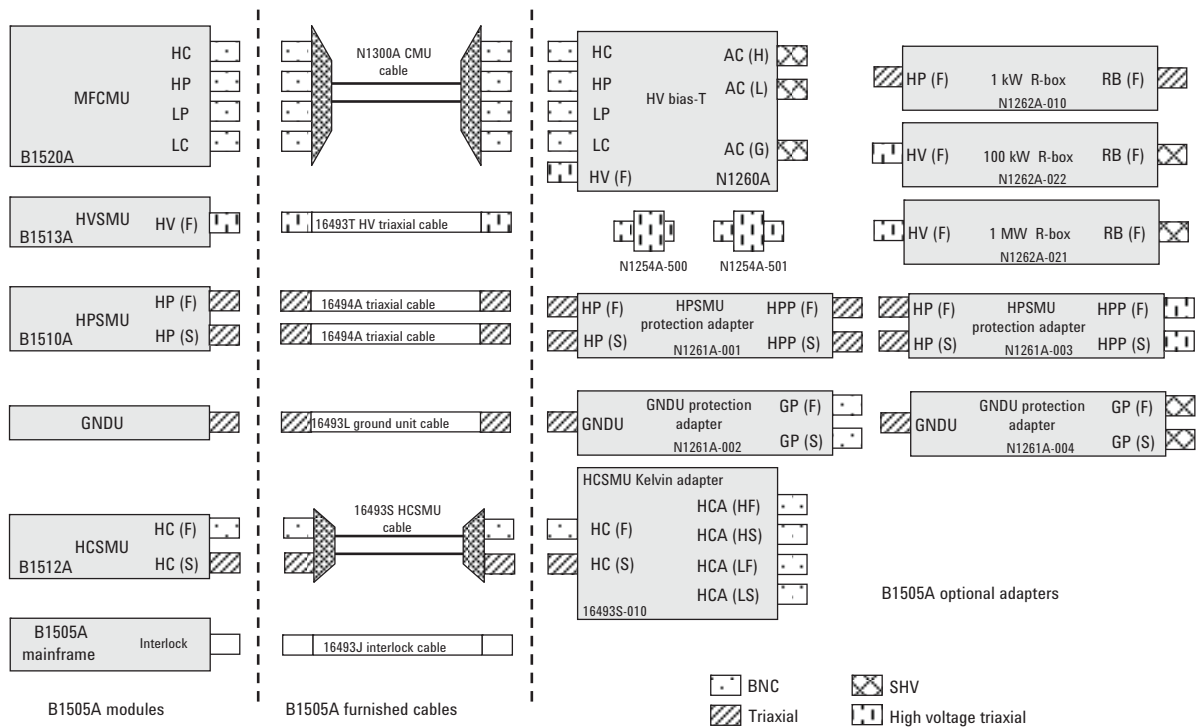
Figure 4.27. Agilent can supply a variety of optional adapters for use with the B1505A on an analytical wafer prober.

If you are planning on using the module selector unit with your wafer prober then the cables and accessories that you need will be different than the case when you are not using the module selector unit. The main difference is that the module selector unit contains device protection circuitry so that you do not need to use the protection adapters for the modules connected to the module selector. Of course, any modules not connected to the module selector would still need to use the protection adapters if there is any possibility of them being exposed to high voltages and currents. Again, Agilent can help you to configure the correct configuration for your use-model and the wafer prober that you plan on using. An illustration showing the most common connectors and adapters that would be used with the module selector unit is shown below.
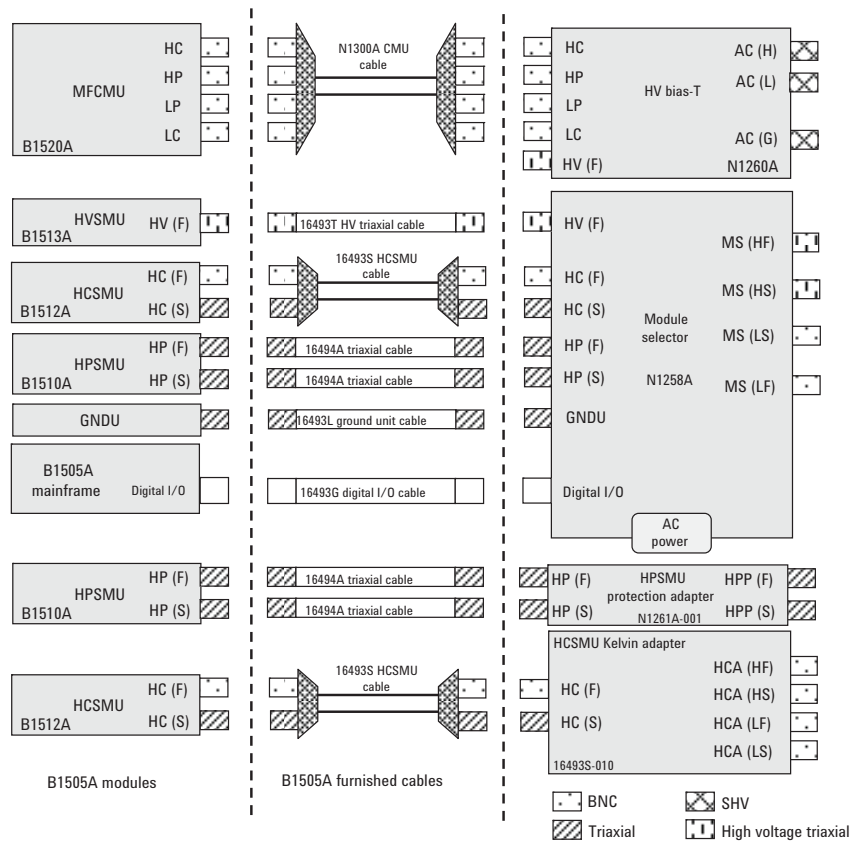
Figure 4.28. The module selector unit also performs protection functions, so the protection adapters are not needed for the modules connected to the module selector unit.

This collection of cabling and adapter accessories is rather daunting, and it is easiest to understand their use by showing some sample wafer probing configurations. ***Note:*** The explanation of how to use the high-voltage bias-T will be deferred until Chapter 8 (in which we will discuss capacitance measurement).

The first case we will examine is high-voltage wafer probing to measure lateral device breakdown and leakage currents (such as MOSFET drain to source breakdown). For high-voltage wafer probing it is essential to connect up the interlock from the B1505A to an approved interlock scheme supported by your wafer prober vendor. It is also important to use protection adapters for the HPSMU and GNDU modules to prevent high-voltages from accidentally damaging these modules. An illustration of this scheme is shown below.
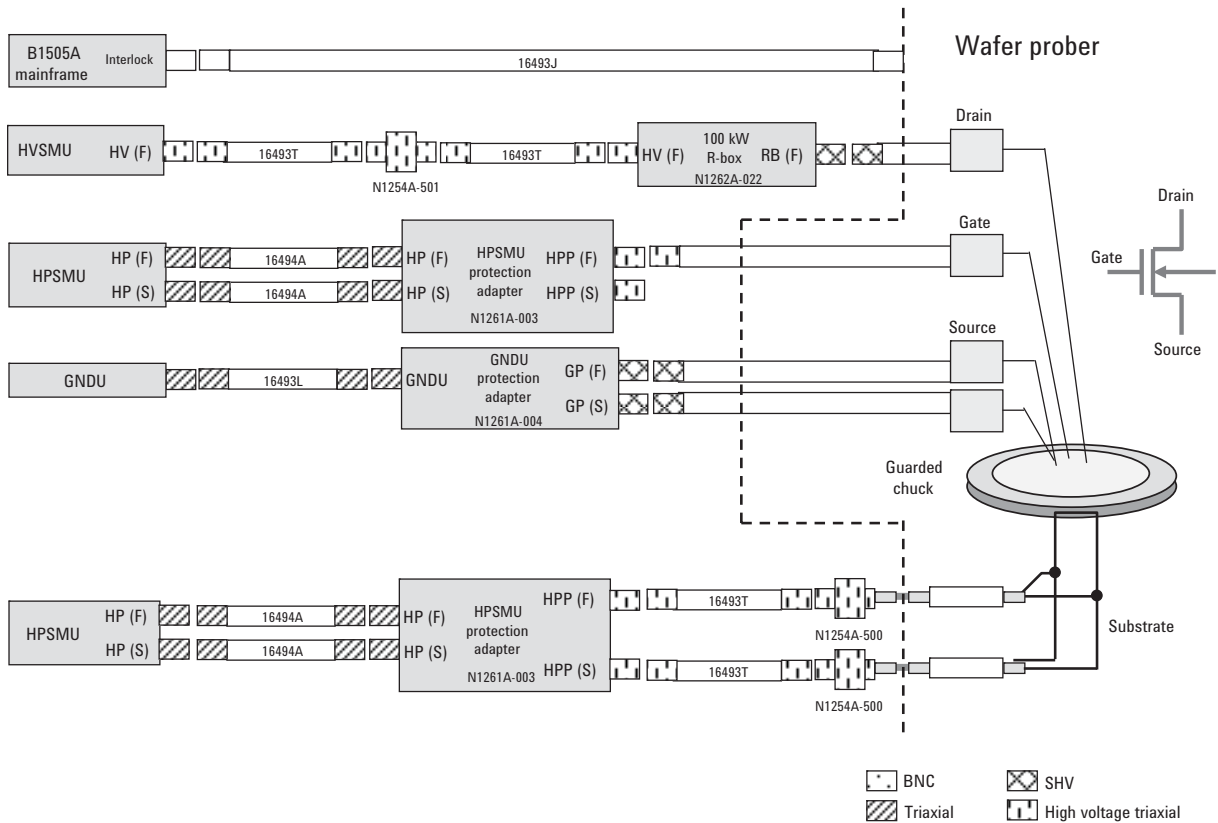
Figure 4.29. A B1505A wafer probing configuration to perform high lateral breakdown voltage measurements.

Note: If your wafer prober does not support SHV or high voltage triaxial connections then you can use the HPSMU protection adapter N1261A-001 which has triaxial outputs and the GNDU protection adapter N1261A-002 which has BNC outputs.

The next case we will examine is high-current wafer probing to measure high lateral currents using a Kelvin configuration (such as high-current MOSFET Ids measurement). Note that the BNC and triaxial outputs of the HCSMU must be separated out into the high force/high sense and low force/low sense BNC outputs using the HCSMU Kelvin adapter. In this situation Kelvin connections are virtually mandatory to prevent measurement error due to the large resistive voltage drops in the cabling that occur when many amps of current are pumped through them. Also remember that in Chapter 3 we explained that the HCSMU outputs are floating and that the low force (LF) and low sense (LS) outputs need to be tied to ground. Therefore, here they are tied to the ground unit using BNC T-connectors. An illustration of this scheme is shown below.
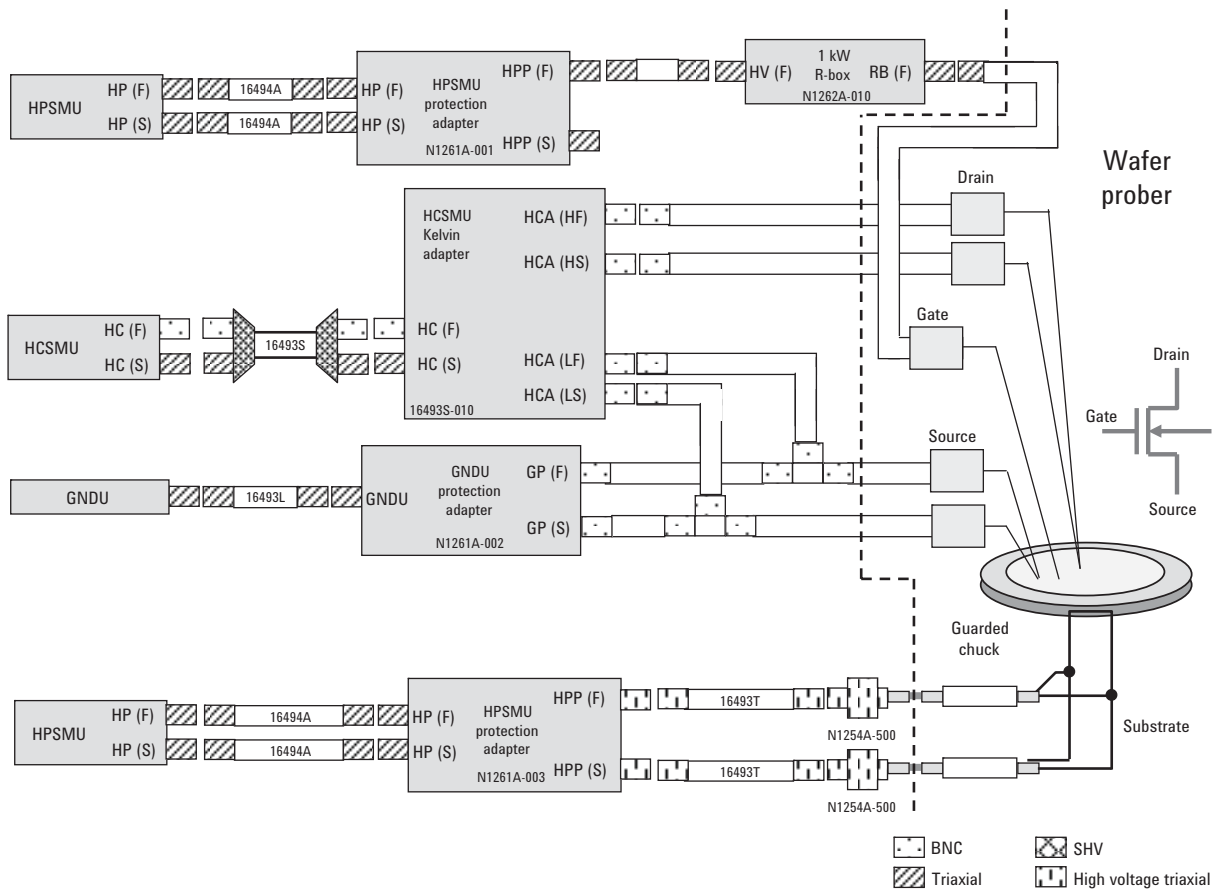
Figure 4.30. A B1505A wafer probing configuration to perform high lateral current measurements.

As can be seen from the previous two examples, it is not a trivial task to manually switch between high-voltage and high-current measurement on a wafer prober. Therefore, if you want to perform both types of measurements you should invest in a module selector unit. In addition to allowing you to easily switch between high-voltage and high-current measurements, the module selector unit also allows you to use the HPSMU to perform high-accuracy, low level measurements not possible with the HVSMU and HCSMU. As the illustration below shows, the wafer prober connections when using the module selector unit are greatly simplified.
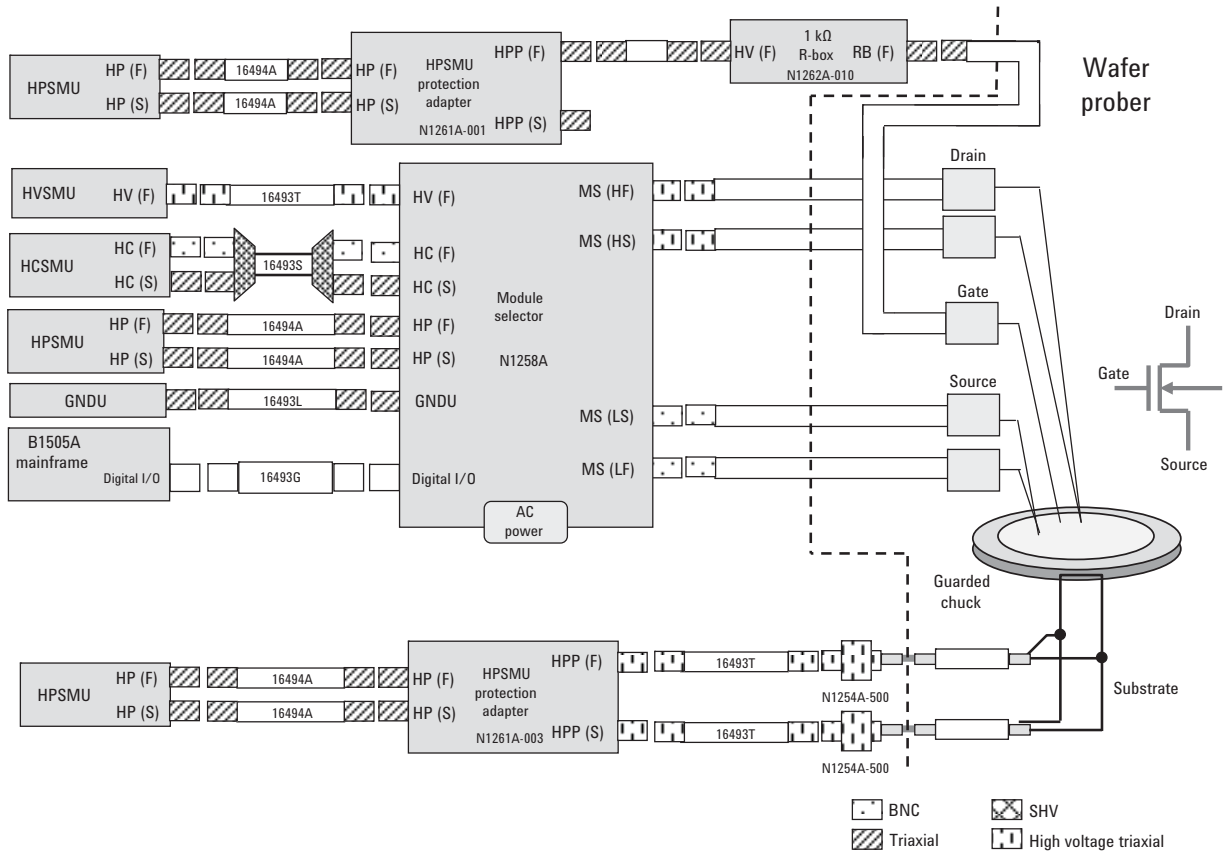
Figure 4.31. A B1505A wafer probing configuration with the module selector unit that can perform both high-voltage and high-current measurements without having to change any cabling.

The additional cost of the module selector unit is more than compensated for by the improvements in convenience and productivity.

Chapter 5: Time Dependent and High-Speed Measurements

"Experience is something you don't get until just after you need it." — Anonymous

Introduction

One of the most prominent trends in parametric test is the movement away from simple sweep and spot measurements (where measurement time is not an important part of the measurement) to measurements where the speed at which the measurement must be made is critical to the accuracy of the measurement. In some cases this requires new types of measurement modules other than SMUs, or even the integration of external instruments (pulse generators, oscilloscopes, etc.) with the parametric measurement equipment. However, some of the newer parametric instruments (such as the Agilent B1500A and B1505A) can make some impressive time sampling measurements just with their SMUs, so it is worthwhile to spend some time discussing all of the available measurement options.

Parallel measurement with SMUS

Parallel measurement is an important capability for all time dependent testing done with SMUs. True parallel measurement with SMUs is only possible when each SMU has its own analog-to-digital converter (ADC) available to it. The Agilent parametric instruments with this capability are the E5270B, the B1500A and the B1505A. For all of these instruments (except the E5260A Series), you can select either a shared high-resolution (HR) ADC or a per-SMU high-speed (HS) ADC for all of the various types of IV measurements as shown below.
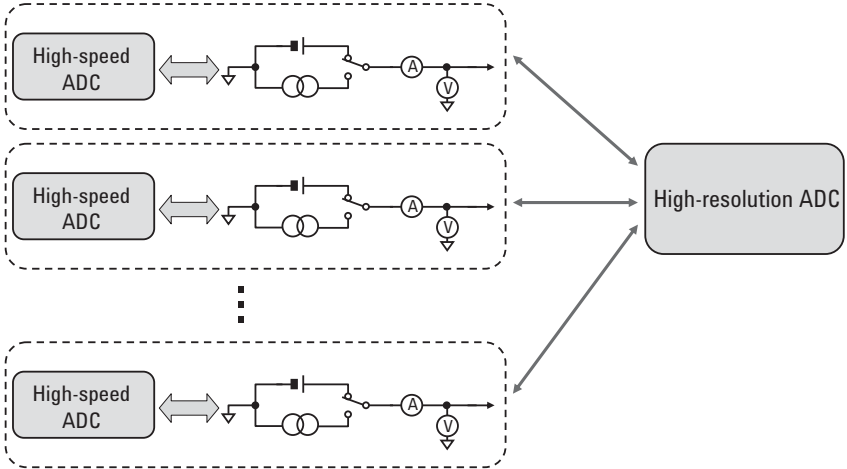
Figure 5.1. On instruments such as the E5270B, B1500A and B1505A you can select either the HS ADC or the HR ADC on a per SMU basis.

In Chapter 3 we discussed the multi-channel sweep mode of the B1500A and B1505A. This feature enables the B1500A/B1505A to make parallel measurements on multiple channels as long as the following conditions are met:

1. All parallel measurement SMUs must be in multi-channel sweep mode.
2. All parallel measurement SMUs must be using the HS ADC.
3. All parallel measurement SMUs must be in fixed measurement ranging.

If all three of these conditions are met, then parallel measurement occurs automatically. As was mentioned when we discussed multi-channel sweep mode, if the start and stop values for a VAR1 (sweeping) SMU are set to be the same then the SMU functions as a constant voltage or current source during the sweep (and it can also make parallel measurements).

One question often asked regarding this procedure is how to verify if the SMUs are indeed performing measurement in parallel (since there is no actual parallel measurement command). The suggested method to verify parallel test is to set both the HS ADC and HR ADC to very long integration times (4 or more PLCs). Repeat the measurement twice, the first time with the SMUs set to HS ADC and the second time with the SMUs set to HR ADC. It should be easy to detect a difference in the measurement times for these two cases (parallel measurement versus non-parallel measurement). Once you are convinced that the parallel measurement is working correctly, you can return the SMUs to their normal HS ADC setting and adjust the integration time for the HS ADC back to the value you want it at for the actual measurement.

Time sampling with SMUs

So far we have discussed making various types of sweep and spot measurements with SMUs, but we have barely mentioned the element of time. The 4155C, 4156C, B1500A and B1505A all possess a built-in time sampling capability that allows their SMUs to measure voltages or currents at specified intervals of time. This measurement capability is useful for certain types of reliability measurements such as time dependent dielectric breakdown (TDDB) where a device is placed under stress and continually monitored until an insulating layer ruptures. The B1500A/B1505A setup screen for a time sampling measurement is shown below.
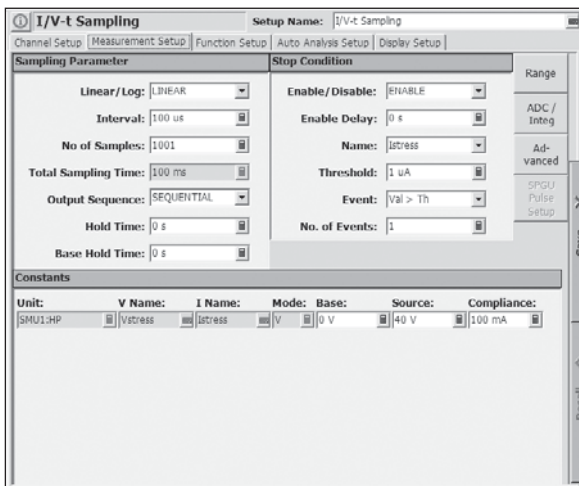
Figure 5.2. Screen capture showing how to set up a time sampling measurement on the Agilent B1500A/B1505A (Classic Test mode).

In addition to specifying how frequently samples should be taken in time (the sampling "Interval"), there are many other parameters that can be specified. The "No of Samples" parameter specifies the maximum number of sample points to be taken, and in the case of the B1500A/B1505A this can be as much as 100,001 points. Moreover, it is possible to take samples in either linearly or logarithmically spaced intervals and a variety of different logarithmic sampling intervals (points per decade) are supported. It is also possible to specify a "Stop Condition", which will stop the sampling measurement before the maximum number of sample points is reached if the specified stop conditions are met. This is a very important capability, since it can reduce test times by stopping a measurement once an event of interest (such as the rupture of a gate oxide) has occurred. An example of this is shown below.
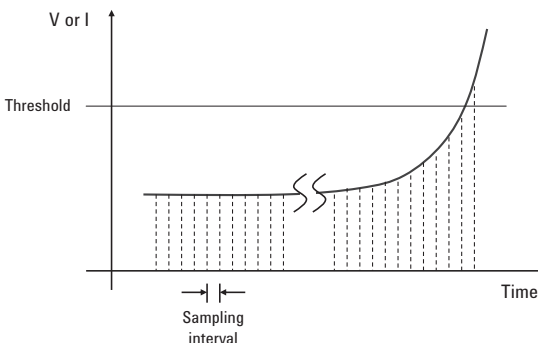
Figure 5.3. By setting up a stop condition you can terminate a time sampling measurement once an event of interest has occurred (such as the rupture of a gate oxide).

91

Sequential versus synchronous sequencing

For time sampling measurements it is sometimes important to consider how the SMU resources will be energized. For certain types of measurements you need to be very careful with respect to the order in which the SMUs are turned on. For other types of measurements, you may actually want ALL of the SMUs to be energized simultaneously. For this reason, Agilent parametric instruments that support time sampling (the 4155C, 4156C, B1500A and B1505A) all also support both sequential and synchronous SMU output sequencing. In the case of the B1500A/B1505A, if the "sequential" setting is selected then the SMUs will turn on in the order shown on the "Channel Setup" page (from top to bottom). However, if the "simultaneous" setting is selected then the SMUs will all turn on at the same time. Note: In both of these cases (sequential or simultaneous) the turn off order of the SMUs will follow the inverse order shown on the "Channel Setup" page (from bottom to top).

It is important to understand that the time sampling capabilities of the newer instruments (as well as their default settings) are different from those of the older 4145A/B. In many cases the failure to duplicate 4145A/B time sampling measurement results on the 4155C, 4156C, B1500A or B1505A can be traced to differences in the turn-on sequence of the SMUs. If you are having trouble correlating measurements results between these instruments then please check the settings on both instruments very carefully.

Setting up the time sampling interval

The most important point to understand when making time sampling measurements with SMUs is that accuracy trumps every other setting. This point is often missed by many users, resulting in much frustration and confusion. What is meant by this statement is that you cannot set up a time dependent measurement on an SMU that conflicts with the measurement accuracy you have specified. The SMU will always take as much time as it needs to meet the specified measurement accuracy, regardless of how fast you are telling it to make a measurement.

It is very easy to miss this point when setting up measurement conditions. For example, both limited ranging and auto ranging require the SMU to work its way down until the optimal measurement range (or range limit) is reached. Obviously, this takes time to complete. If you try to specify a time sampling interval that is smaller than the time required for the SMU to complete its measurement ranging, then the SMU will simply ignore the time settings and take however long is necessary to complete the measurement ranging. Therefore, it is always recommended to use fixed measurement ranging when performing time sampling measurements. Similarly, you cannot use power line cycle integration when making time sampling measurements. You must use an integration time of less than one PLC; in fact even when using integration times of less than one PLC you need to be careful not to specify an integration time that exceeds the measurement interval you are trying to achieve.

The exact procedure and limitations on setting up a time sampling measurement are idiosyncratic to the instrument that you are using. The following are important considerations to keep in mind if you set the sampling interval to less than 2 ms on the B1500A/B1505A:

1. All measurement channels must use the high-speed ADC; sampling intervals of less than 2 ms are not supported when using the high-resolution ADC.

2. If fixed measurement ranging is NOT selected, then the measurement channels automatically select the lowest fixed measurement range that covers the compliance value selected for that channel.

3. If the expected measurement time as determined by the number of averaging samples specified in the integration time is longer than the sampling interval, then the measurement channels will attempt to adjust the number of averaging samples so as to keep the specified sampling interval.

4. If multiple measurement channels are specified and all of the parameters are set up correctly, then all of the measurement channels will measure in parallel.

Advanced time sampling features (B1500A/B1505A)

We have spent a lot of time explaining how the SMUs make measurements in time sampling mode, but we have not discussed much about how they force voltage and current. In time sampling mode the B1500A and B1505A SMUs actually support some very useful sourcing features that have some important measurement uses. To begin with, let us review the basic output of an SMU during time sampling.
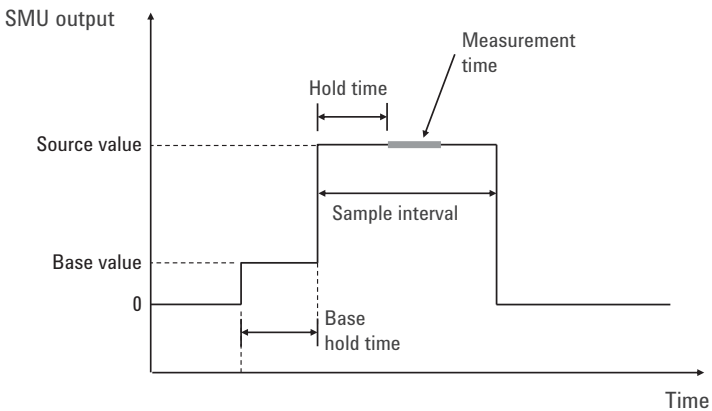
Figure 5.4. The basic SMU sourcing parameters for a time sampling measurement made on the B1500A/B1505A.

Note that both a "Base Value" and a "Base Hold Time" can be specified. It should also be pointed out that in both this diagram as well as the following diagrams the assumption made is that the measurement times are all less than the specified sampling interval (as is should be if everything is set up correctly).

The B1500A and B1505A also support a very useful feature (found under the "Advanced" settings) that permits SMU output to remain active AFTER the measurement completes. This feature is referred to as "bias hold after measurement", and it can be set to either the base or source value. The figure shown below illustrates the case where the SMU has been set to hold its base value.

93
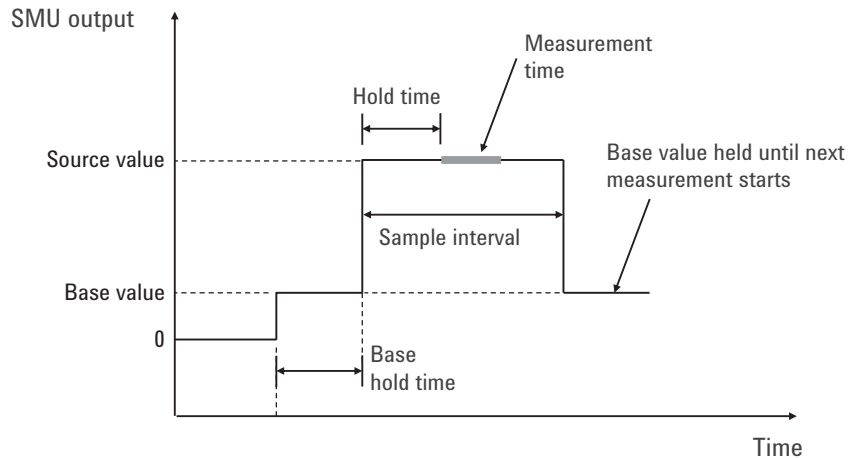
Figure 5.5. The B1500A/B1505A permit you to hold the bias to a non-zero value after measurement and this can be set to either the base or source value.

Note: It is important to understand that the bias is held only under two conditions:

1. When running multiple tests within another application test
2. When sequencing tests in "Quick Test" mode

The bias hold feature is NOT the same as the SMU Standby Mode, which energizes the SMUs to a specified value at all times (even when a measurement is not being made).

Finally, there is one other very interesting feature supported by the B1500A and B1505A in time sampling mode. This is the ability to specify a negative hold time. The figure shown below illustrates what is meant by a negative hold time.
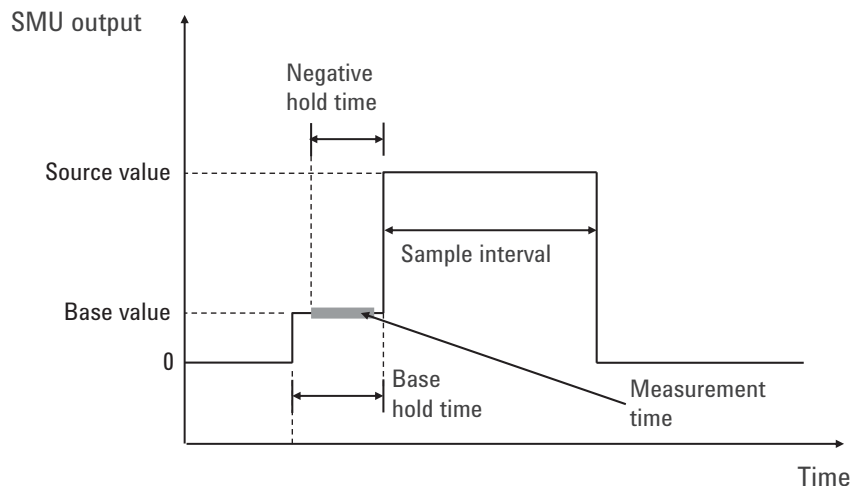
Figure 5.6. The B1500A/B1505A support a negative hold time feature that permits the SMU to start measuring BEFORE the SMU starts to supply stimulus.

It is important to note that the negative hold time feature is only supported when the sampling interval is less than 2 ms, and that the maximum negative hold time supported is –90 ms.

It turns out that these advanced time sampling features can be used very effectively for certain types of reliability measurements. In particular, negative bias temperature instability (NBTI) measurements can be performed very fast and effectively with SMUs by employing the aforementioned measurement features. Essentially, in an NBTI measurement we apply stress bias to a transistor gate and then periodically remove the stress to measure the transistor characteristics. However, if the stress drops to zero during the stress-to-measure transition then the transistor can "recover" and the measurement results will be invalid.
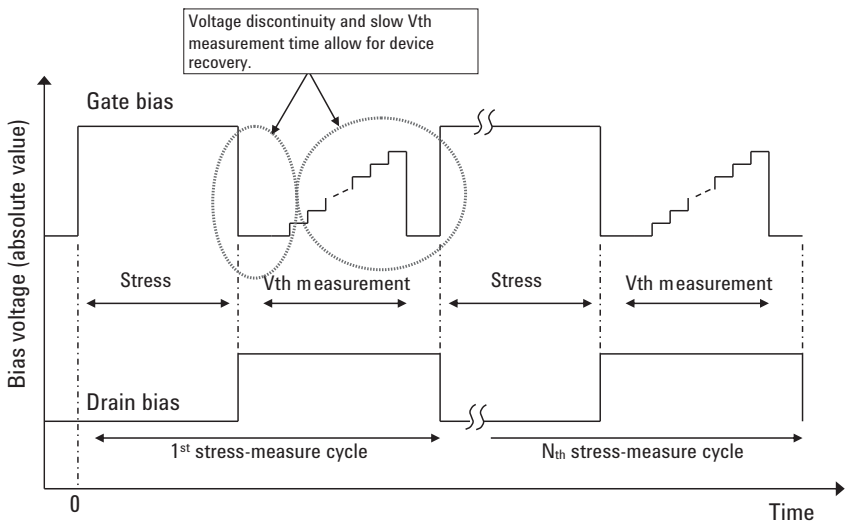
Figure 5.7. When making NBTI measurements it is essential that no discontinuities occur during the transition from stress to measurement.

By using the bias hold feature on the B1500A/B1505A we can make sure that this does not happen. In addition, by using the negative hold time feature we can actually start measuring while still in the stress phase. This ensures that no data is lost during the stress to measure transition. This technique is illustrated in the figure shown below.
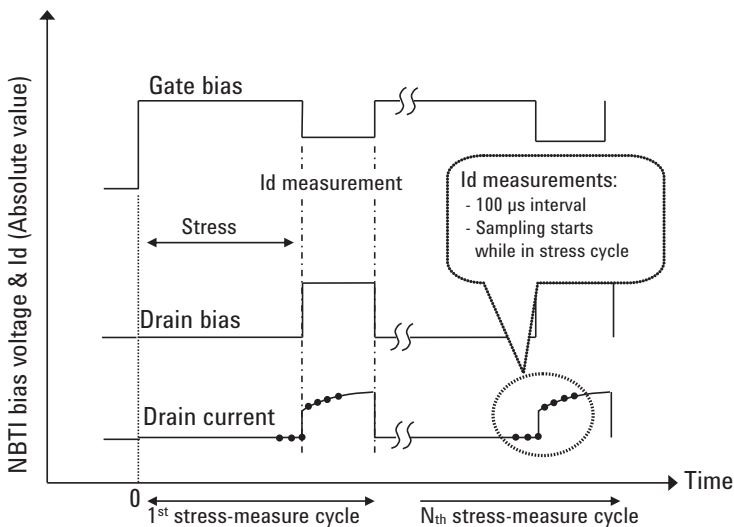
Figure 5.8. Making ultra-fast NBTI measurements using B1500A SMUs.

This technique is often referred to as "on-the-fly NBTI" measurement. The B1500A can make these measurements on a single SMU in 100 μs intervals. Using parallel measurement techniques, the B1500A can make parallel measurements using multiple SMUs in intervals given by the following equation:

$$t = 100\ \mu s + [20\ \mu s \times (n - 1)]$$

Here "n" is the total number of SMUs measuring in parallel. For more information please refer to the relevant B1500A application notes on this topic.

Maintaining a constant sweep step

It is sometimes desirable to perform a sweep measurement with constant sweep steps. However, it is important to understand that in a standard sweep measurement there is no guarantee of the measurement time to complete each point in the sweep. The following plot illustrates this situation.
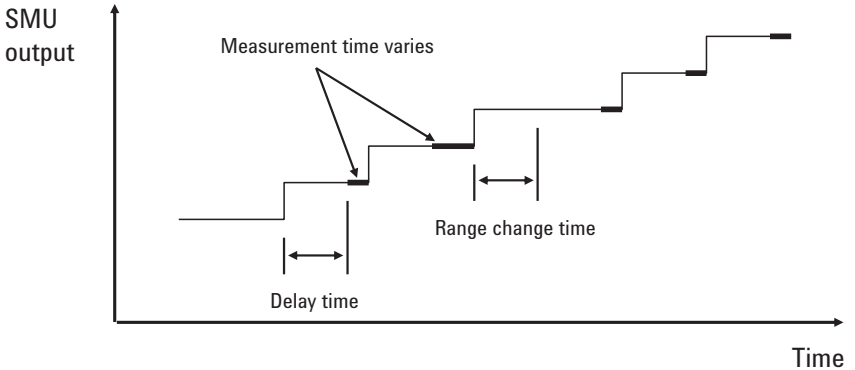
Figure 5.9. On a standard sweep measurement the amount of time spent at each point in the sweep can vary greatly.

There are two issues impacting the sweep measurement just shown. The first is that the measurement time varies from one measurement point to the next. The second is that if a range change occurs, then there will be a very long delay between measurement points.

There are ways to solve both of the aforementioned problems. Parametric measurement instruments support a feature known as "Step Delay". This feature allows you to specify the time that the instrument waits from the beginning of the actual measurement to when it increments to the next point in the sweep. Therefore, as long as the step delay is set to a value that is greater than that of the maximum expected measurement time then we can specify a constant total step time. Of course, this assumption is only valid if we do not encounter any range changes. To make sure that no range changes occur, fixed ranging should be used. The key thing to remember is that the fixed range chosen must include the maximum expected value to be measured during the sweep.
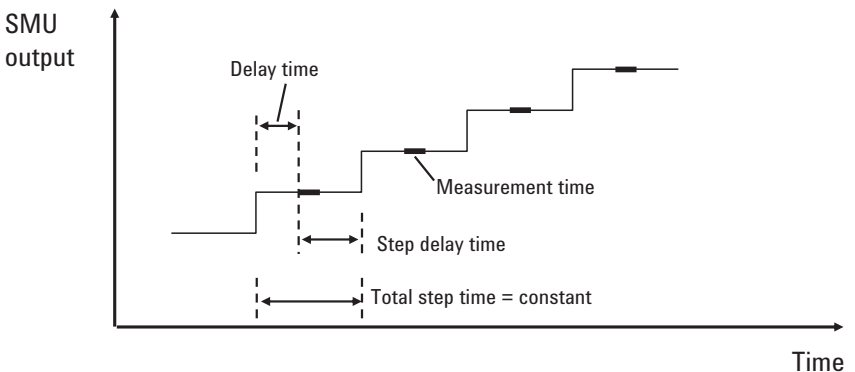
Figure 5.10. By using step delay and fixed measurement ranging, a constant sweep step time can be achieved.

There are some reliability measurements where maintaining a constant sweep step time is very important. The most common of these are the voltage ramp (VRAMP) and current ramp (JRAMP) tests. An example of a VRAMP test that uses the step delay feature to create uniform time steps for the voltage ramp is shown below.
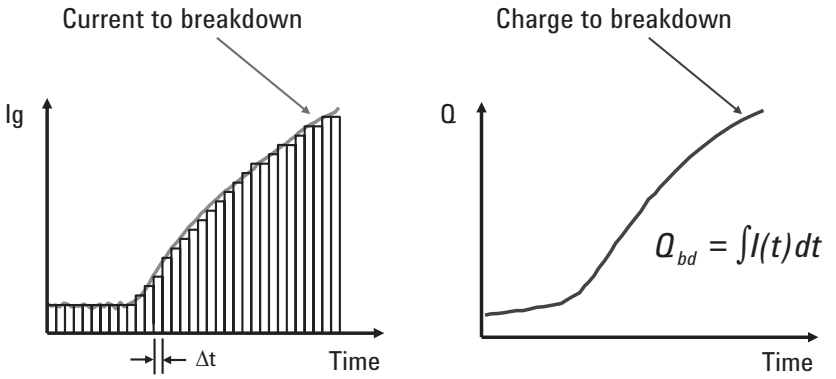
Figure 5.11. By using the step delay feature to create uniform time steps for a voltage ramp (VRAMP) sweep, it is very easy to calculate the total charge-to-breakdown (Qbd).

By maintaining a constant time step for the voltage sweep it is very easy to calculate out the charge-to-breakdown (Qbd) from the measured gate current at each point in the sweep using a simple rectangular approximation:

$$Q_{bd} = \int I(t)dt \approx \sum I \bullet \Delta t$$

High speed test structure design

To succeed in making high speed measurements you need more than just the correct measurement instrumentation; you also need to put sufficient forethought into the creation of the test structures that will be used to make the measurement. Attempts to make fast pulsed measurements with conventional DC test structures using DC positioners are unlikely to yield good measurement results. In general, fast pulsed measurements require test structures designed for a ground-signal (GS) or ground-signal-ground (GSG) measurement environment and RF positioners. The following figure illustrates this point.

Structure for conventional
DC measurement

Structure optimized for
RF measurement

Large overshoot and ringing

Clean pulse shape
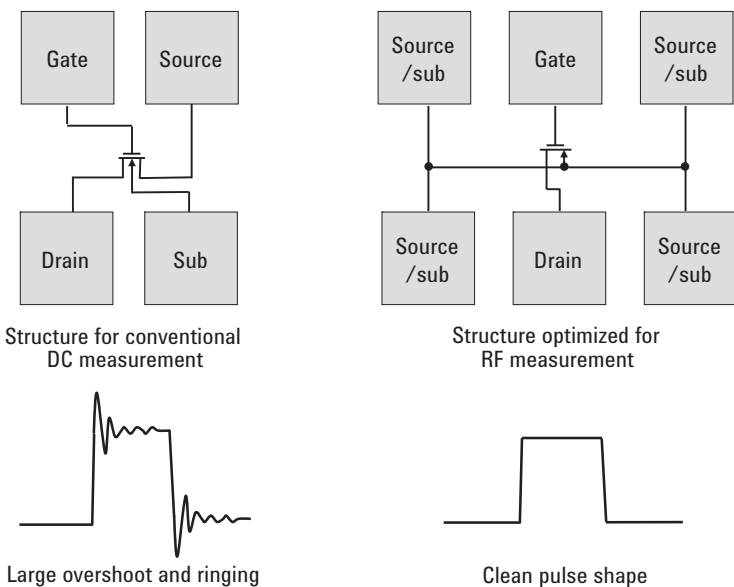
Figure 5.12. Illustration showing how conventional test structures using DC probes do not yield satisfactory results when attempting to make fast pulse measurements.

A photograph of a GSG RF probe tip is shown below.

Figure 5.13. A GSG probe tip. Note that this probe tip uses an SMA style coaxial connection. [Photo courtesy SUSS Microtech]

GSG probes do not necessarily produce superior measurement results to GS probes, since the results obtained also depend strongly on the pad layout and DUT structure. However, the following figure illustrates the reason why GSG probes can be superior to GS probes at high frequencies.
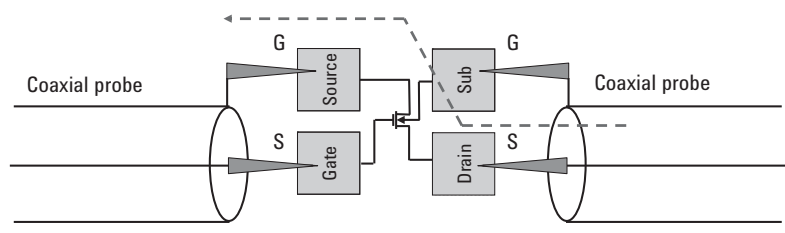
Some of the incoming signal energy must travel a long distance through the DUT to ground.

→ Large signal loss at high frequencies (not a concern for DC measurement)
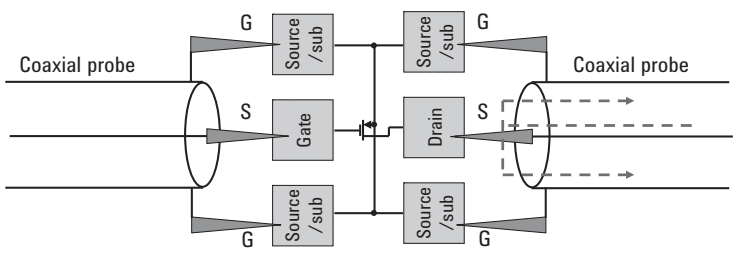
All of the incoming signal energy has a short return path through the DUT to ground.

→ Signal loss over frequency is minimized.

Figure 5.14. GSG probes can provide superior performance at higher frequencies relative to GS probes due to their shorter return path for the incoming signal energy.

The basic limitation of GS probes is that the energy entering the DUT can have a potentially long path length to return to ground, which results in significant signal loss. However, GSG probes have a much shorter return path to ground, which minimizes the signal loss over frequency.

Another consideration is the bandwidth of the RF probes that you are using. Since the shortest pulse widths that we use in parametric test are about 10 ns, at first thought it might seem that one or two hundred Megahertz of bandwidth would be sufficient. However, this is actually not the case. Remember that a square wave (pulse) is actually the summation of an infinite series of sine waves (odd harmonics). You need to have enough bandwidth in your RF probes (and indeed in your overall system) to support the higher harmonics, or you will get a distorted pulse shape. Therefore, to produce a clean 10 ns pulse you typically need at least 1 GHz of bandwidth.

In addition to the type of RF probes chosen and their bandwidth, the physical layout of the DUT also has a big impact on the integrity of the waveform pulse. In particular, placing the drain and source connections on MOSFET on opposite sides of the layout creates extremely long signal paths that will cause lots of ringing and greatly distort the shape of the applied pulse as shown below.

OK for pulse widths ≥ 200 ns
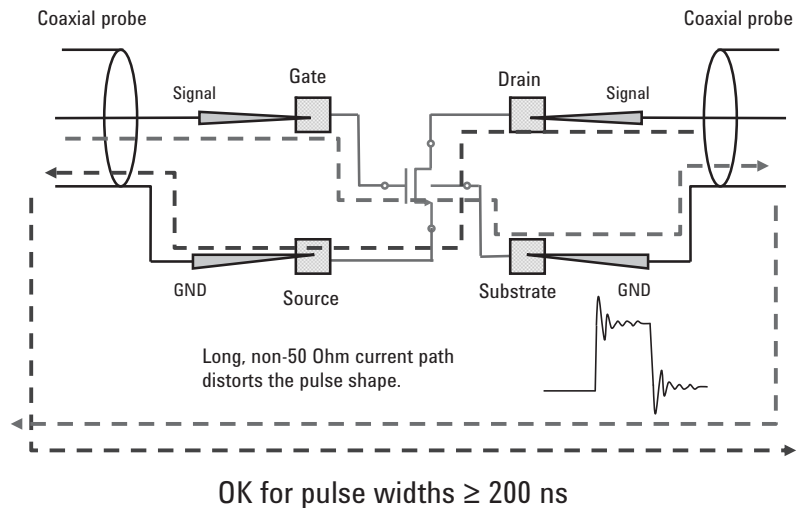
Figure 5.15. A GS probing arrangement with the drain and source on opposite sides as shown in this diagram can only support pulse widths of 200 ns or greater without causing pulse distortion.

Layouts with the drain and source terminals on opposite sides will generally only support pulse widths of 200 ns or greater.

A much better arrangement is to place the source and drain on the same side of the layout as shown below.

OK for pulse widths ≥ 100 ns
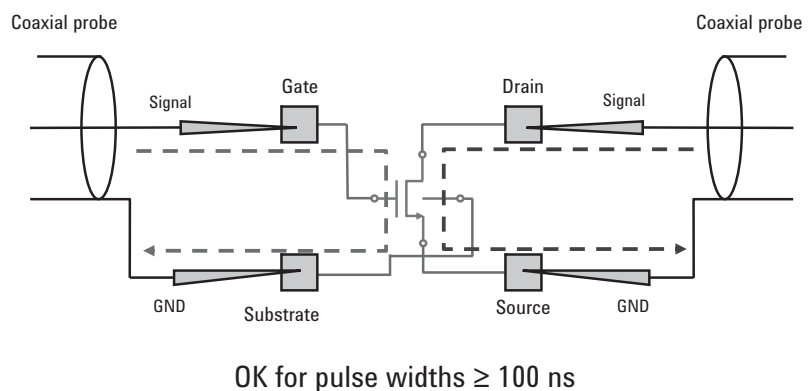
Figure 5.16. A ground-signal probing arrangement with the drain and source on the same side as shown in this diagram can support pulse widths down to around 100 ns.

Placing the source and drain on the same side creates a much shorter path for the signal energy, and a ground-signal pad layout can support pulse widths down to approximately 100 ns.

In order to achieve pulse widths down to 10 ns, both a ground-signal-ground pad layout and a structure design that shorts the source and substrate connections together are required. An example of this is shown below.
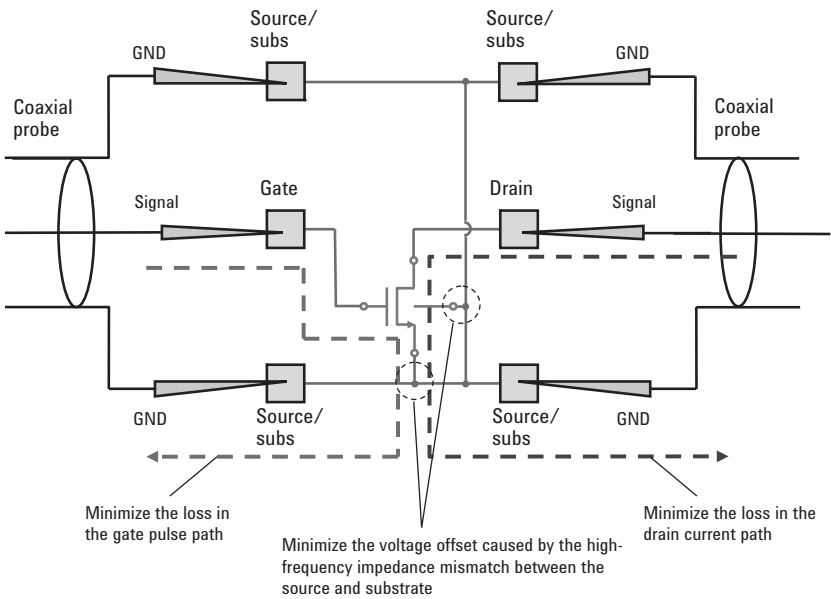
Good for pulse widths ≤ 10 ns

Figure 5.17. A ground-signal-ground probing arrangement with the source and substrate connections tied together as shown in this diagram can support pulse widths down to 10 ns or less.

Experimental results show that this layout scheme can reliably produce clean 10 ns pulses at the gate of the MOSFET.

Fast IV and fast pulsed IV measurements

Overview

Increasingly, many parametric measurements require some sort of fast IV or fast pulsed IV measurement capability. There are many factors driving this need, including new physical effects caused by decreasing lithographic and voltage scaling, the use of novel and exotic materials, and faster operating speeds that can generate large transient currents. Given this wide range of needs, it is difficult to create a universal solution that is both flexible and cost-effective. For example, some pulsed measurements may require large voltages but relatively slow pulse capability; others may require small voltages but relatively fast pulse capability. While it is certainly possible to design a single solution that can do both, the cost of creating an all-in-one solution can become prohibitive (especially if it is important to support multiple measurement channels for parallel testing). It is actually more efficient to support a range of solutions so that the user can select the one that best fits within their test needs and budget.

Until recently, solutions for fast pulsed IV measurements typically consisted of a precision pulse generator and oscilloscope as shown in the following illustration.
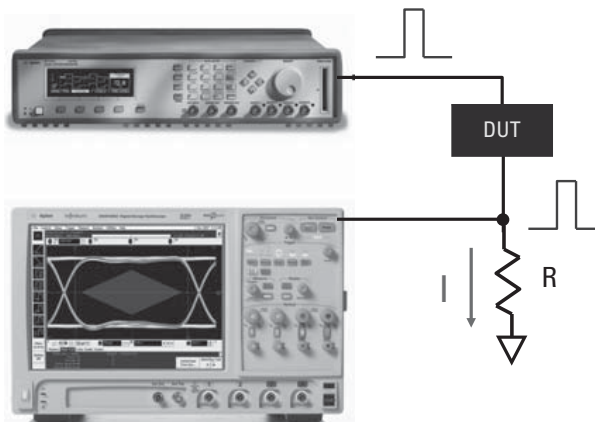
Figure 5.18. A conventional fast pulsed measurement solution using off-the-shelf instrument components.

In this solution, the current flowing through the DUT is calculated by measuring the voltage across a shunt resistor using an oscilloscope. However, this type of solution has several drawbacks:

1. It requires a precision resistor.
2. It requires some sort of software to integrate everything together.
3. It has limited accuracy.
4. It can only test one device at a time.
5. It requires a lot of external cabling.

While none of these factors are insurmountable, they can take many engineering hours of effort to overcome. It is typically more cost effective to use either a solution that consists of self-contained hardware or a solution that consists of external hardware and professionally developed software. In the following sections we will look at several different fast pulsed IV measurement solutions and discuss the types of measurement challenges for which they are best suited.

High-voltage semiconductor pulse generator unit (HV-SPGU)

The B1500A supports a single-slot, two-channel HV-SPGU module that can generate pulses with amplitudes of ±40 V (80 V peak-to-peak) into an open load and that can supply pulse widths from 5 μs to 10 s for pulsed IV measurement. Note: Although the HV-SPGU module is capable of generating voltage pulse widths as narrow as 50 ns, when used in pulsed IV measurements the HV-SPGU's pulse width is limited by its minimum current measurement interval (which is 5 μs). A photograph of an HV-SPGU module (2 output channels total) is shown below. Up to five HV-SPGU modules can be installed in a single B1500A, for a maximum of ten channels. Synchronization between the modules is maintained using external connections between the "Ref Out", "Ref In", "Sync Out", and "Sync In" outputs and inputs.

Figure 5.19. The B1500A´s HV-SPGU module has ±40 V output capability and supports two channels per module.

Although this pulse generator module has many features that make it ideal for flash memory testing, it also can be used for fast pulsed measurement. In order to improve accuracy, the HV-SPGU has a built-in voltage monitoring capability. The output voltage monitoring capability permits the actual voltage at the output to be measured. A simplified schematic of the HV-SPGU module is shown below.
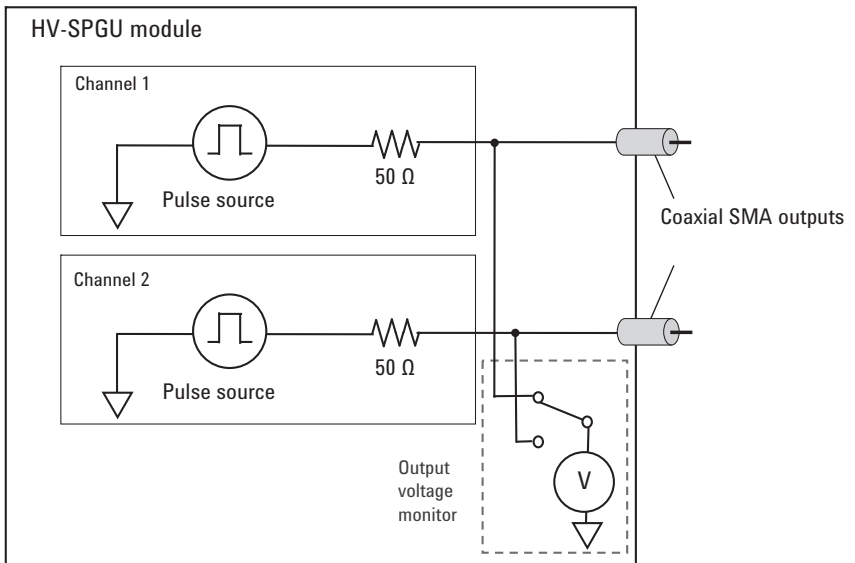
Figure 5.20. Simplified HV-SPGU schematic showing the shared output voltage monitor circuit.

Basic pulsed IV measurement on a MOSFET can be performed using only a single HV-SPGU module by connecting the drain to one channel and the source to another as shown below.
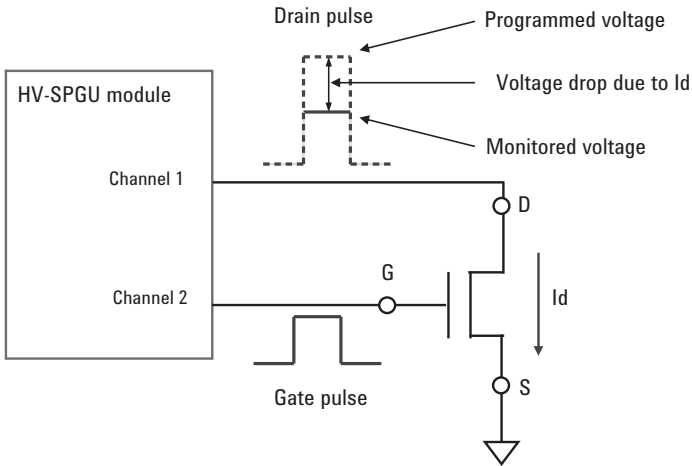
Figure 5.21. Performing pulsed IV measurement on a MOSFET using a single HV-SPGU module.

By determining the difference between the programmed voltage and the monitored (measured) voltage and by making use of the fact that the output impedance is a well-controlled and known value (50 Ohms), it is possible to determine the current flowing through the HV-SPGU output using a simple calculation.

Of course although this calculation is simple in principle, in practice obtaining an accurate value for Id can be challenging. To simplify this, Agilent supplies a library of application tests for performing pulsed IV measurements using the HV-SPGU. The most basic of these application tests permits you to specify single (spot) pulsed measurements on two HV-SPGU channels. The setup screen for this application test is shown below.
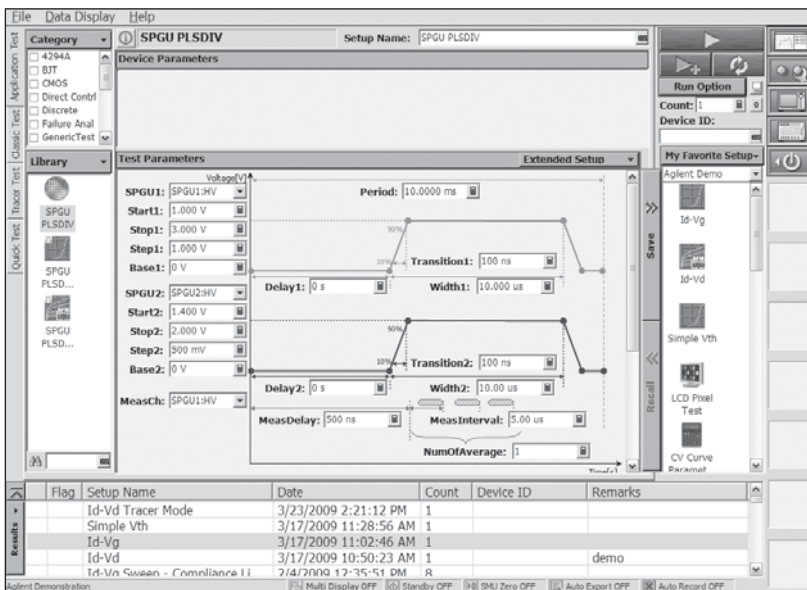
Figure 5.22. The basic B1500A EasyEXPERT application test for performing pulsed IV measurement using the HV-SPGU module.

By itself this application test is not particularly useful; however, it can be used within other application tests to create more complicated pulsed IV application tests. Using this technique, Agilent supplies the two most commonly needed application tests to perform pulsed Id-Vg and Id-Vd sweeps. An example of the input screen for the pulsed Id-Vg application test is shown below.
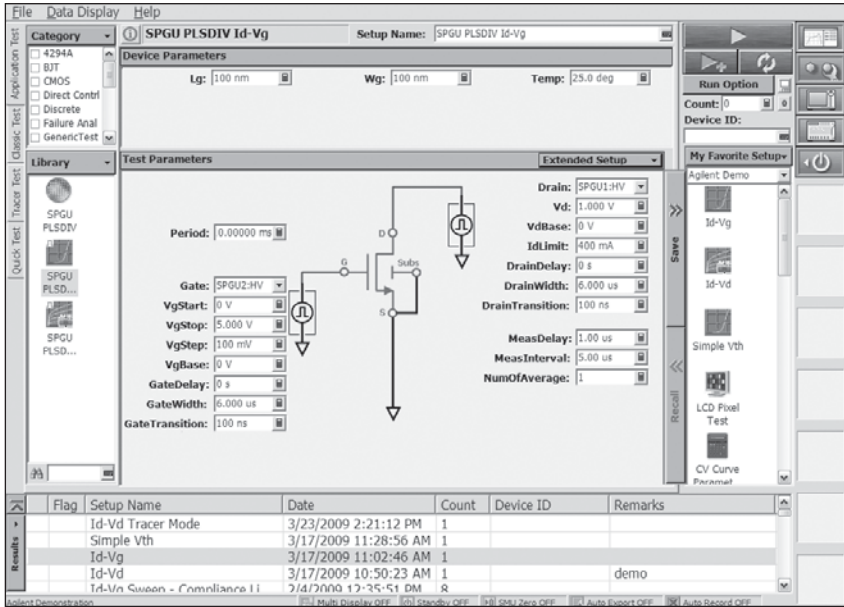
Figure 5.23. A sample B1500A EasyEXPERT application test to perform fast pulsed Id-Vg measurement using the HV-SPGU module.

The input screen for the Id-Vd application test is similar.

Due to its large voltage output capability, the HV-SPGU module is a good choice for testing devices in the "medium" power range (such as GaAs and HEMT devices used in RF applications). In addition, it is a good choice for measuring transient currents in circuits where more than 10 V of amplitude are required.

Waveform generator/fast measurement unit (WGFMU)
The B1500A supports a single-slot, two channel WGFMU module for pulsed IV measurement that can generate pulses with amplitudes of ±3 V, ±5 V, 0 V to 10 V or -10 V to 0 V and that can supply pulse widths from 100 ns to virtually any maximum length. Up to five WGFMU modules can be installed in a single B1500A, for a total of ten channels maximum. Similar to the case of the HV-SPGU module, synchronization between the WGFMU modules is maintained using external connections between the "Sync Out", and "Sync In" outputs and inputs. A photograph of the module outputs is shown below.

Figure 5.24. The B1500A's WGFMU module can output 10 V peak-to-peak and has a minimum pulse width of 100 ns.

The B1500A's WGFMU solution consists of the WGFMU module as well as two remote-sense and switch units (RSUs) connected to the WGFMU module via furnished cables. The WGFMU module contained in the mainframe generates the arbitrarily waveforms, and these waveforms are then transmitted through the cables to the RSUs. The RSU, which performs the actual current or voltage measurement, is separate from the WGFMU module so that it can be placed near the device under test (DUT) to minimize cable lengths and guarantee accurate high-speed measurement.
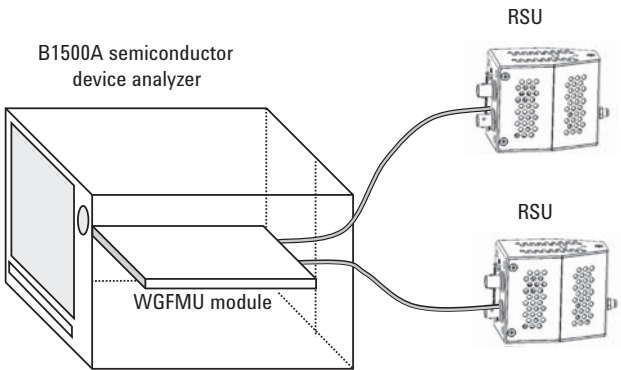
Figure 5.25. Each single-slot WGFMU module has two channels, and each channel has an RSU associated with it.

Each WGFMU channel has independent arbitrary linear waveform generator (ALWG) voltage generation capability and two user selectable operation modes: PG mode and Fast IV mode. The PG mode combines a very fast voltage measurement capability with 50 Ohm output impedance to minimize waveform reflections. The Fast IV mode has a slightly slower measurement speed and slower waveform rise/fall times than the PG mode, but it can measure both current and voltage. The figure shown below shows a simplified circuit diagram of one channel of the WGFMU and RSU combination.
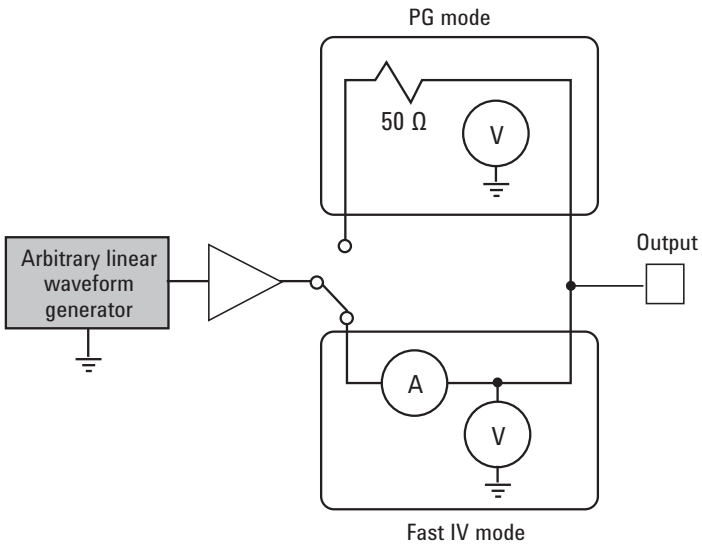
Figure 5.26. A simplified schematic of one channel of the waveform generator/fast measurement unit (WGFMU) and remote-sense and switch unit (RSU) combination.

In addition to its high-speed measurement capabilities, each RSU has the ability to multiplex its output between the WGFMU and an optional SMU. Even though the triaxial to coaxial cable conversion does not permit measurements to the full limit of the SMU specifications, this switching capability does permit the user to verify DC measurement results made using the WGFMU against those of an SMU. Moreover, each RSU also has a coaxial output that can be connected through an internal buffer circuit to the SMA coaxial output; this permits monitoring of the RSU output waveform with an external instrument such as an oscilloscope. A schematic illustrating these features is shown below.
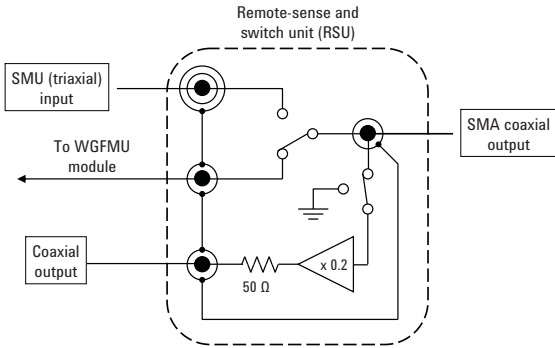
Figure 5.27. The RSU has features that allow it to work with an optional SMU and with an optional external monitoring device such as an oscilloscope.

A photograph of the RSU module inputs is shown below.

Figure 5.28. A photograph of the RSU module showing its voltage monitor (coaxial) output, SMU (triaxial) input and connector input for the cable that goes to the WGFMU module.

The WGFMU module has the ability to not only generate arbitrary waveforms, but also to measure either current or voltage very quickly at any point on those waveforms. In addition to this powerful measurement capability the WGFMU module also allows the user to dynamically change the measurement range to optimize measurement accuracy. Of course, the user can also specify the sampling rate and averaging time. The following diagram illustrates all of these features.
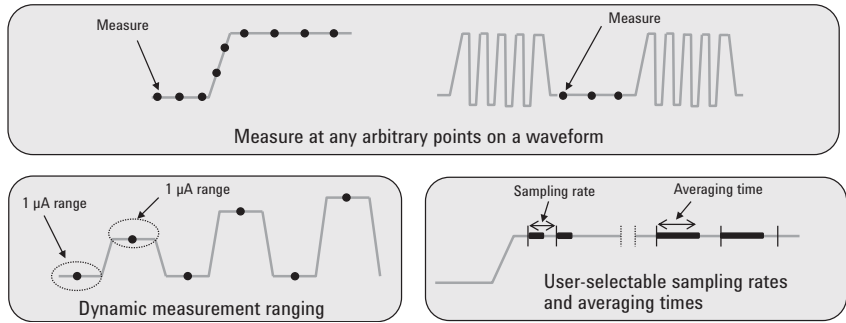
Figure 5.29. The key measurement feature of the WGFMU is its ability to measure at any point on an arbitrarily generated waveform.

An Id-Vg measurement made with the WGFMU module using three different measurement conditions is shown below.
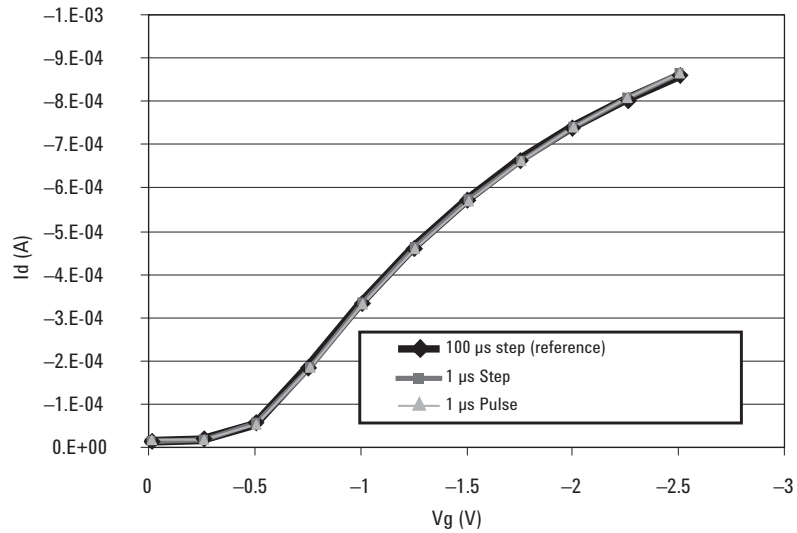
Figure 5.30. The fast pulsed Id-Vg sweep measurement (~ 1 µs per point) made using the WGFMU module correlates well with the non-pulsed, slower measurements.

To create a baseline, a staircase sweep measurement with 100 ms per step (50 ms delay) was performed first. Next, a staircase sweep measurement with 1 ms per step (500 ns delay) was performed. Finally, a pulsed sweep measurement with a 1 ms pulse (100 ns rise/fall times, 500 ns delay, 2 ms period) was performed. These measurement results confirm that the WGFMU module can indeed make reliable high-speed pulsed measurements.

The WGFMU module is an ideal solution for many types of reliability test that require fast measurement capability. Not only can the WGFMU module measure very quickly, but its arbitrary linear waveform generation capability makes it very easy to create complicated AC stress conditions. The following figure shows examples of reliability measurements made with both DC and AC stress conditions.

DC Stress

AC Stress (100 kHz, Duty cycle 50%)
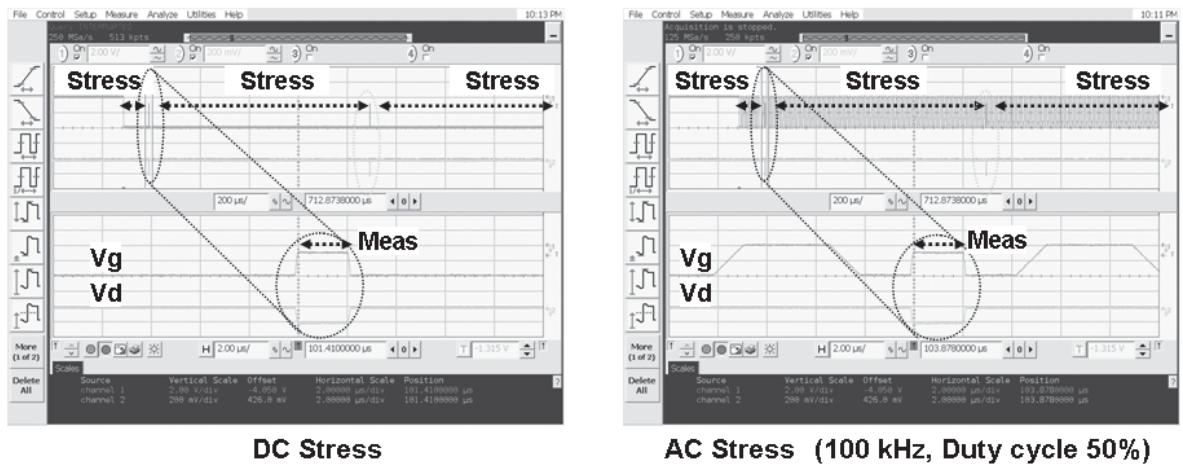
Figure 5.31. The WGFMU can easily apply both DC and AC stress, and there is no delay in transitioning from stress to measure and there is no glitching.

The advantage of the WGFMU module is that the transition from stress to measure and back to stress again is virtually seamless.

Similar to the HV-SPGU module, Agilent supplies a library of application tests to control the WGFMU module.
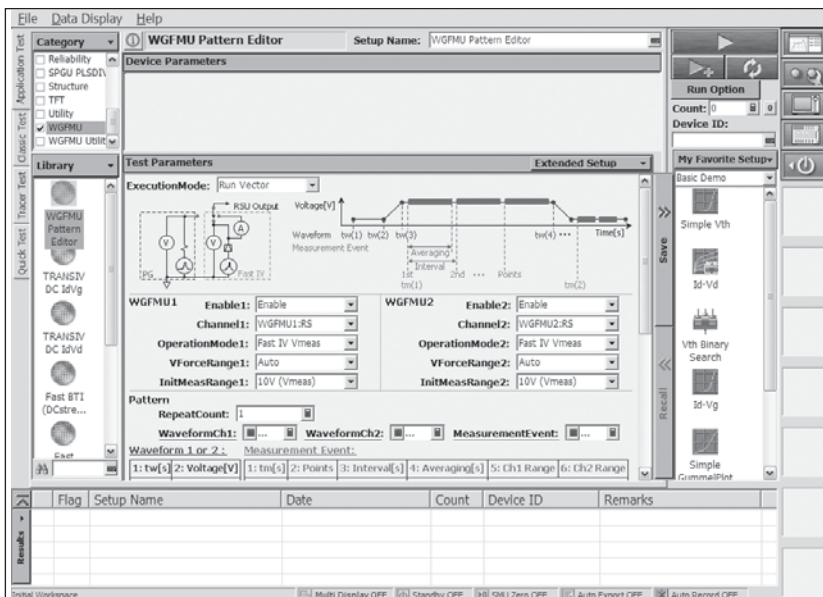
Figure 5.32. The WGFMU module pattern editor application test allows you to control the WGFMU from within EasyEXPERT.

An ideal application of the WGFMU module is ultra-fast NBTI testing.
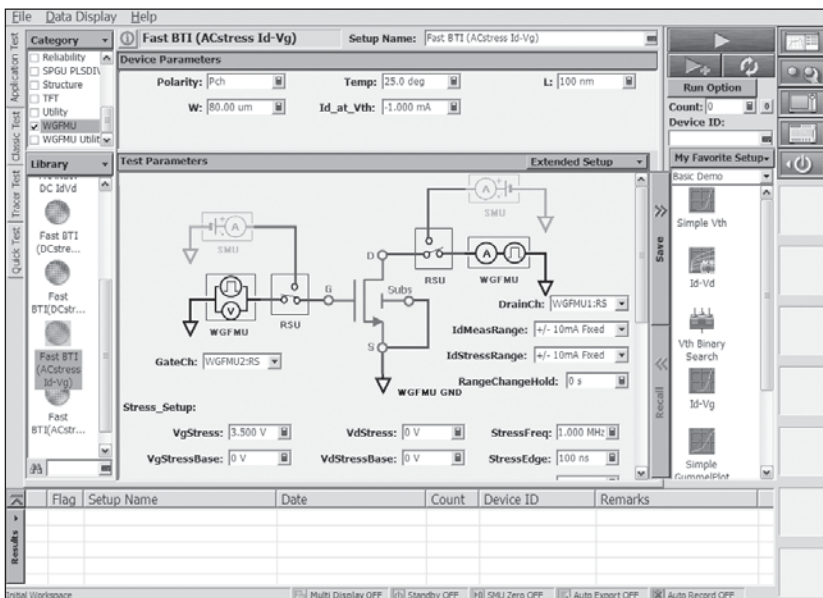
Figure 5.33. The fast BTI application test comes in versions that support both DC and AC stress (the AC stress version is shown here).

For complicated stress and measurement reliability testing, it is difficult to control the WGFMU using an application test due to all of the various parameters that need to be set. In these cases, it is better to control the B1500A remotely using an external controller and software. Agilent can supply the necessary API to control the WGFMU module using a variety of third party software.

Ten nanosecond pulsed IV solution

Agilent's 10 ns pulsed IV solution (B1542A) is designed to meet the most exacting needs of pulsed MOSFET transistor characterization. Before proceeding to explain the details of this solution, it is worthwhile to first understand why this solution is necessary. There are two unrelated situations that can require pulsed IV measurements with pulse widths as narrow as 10 ns:

1. MOSFET transistors fabricated on silicon-on-insulator (SOI) based technologies.
2. MOSFET transistors fabricated using high-k gate dielectrics.

We will briefly explain the physical phenomena underlying these two cases.

The silicon dioxide used on MOSFET SOI transistor structures is an extremely good thermal insulator. This makes these devices especially susceptible to the effects of device self-heating
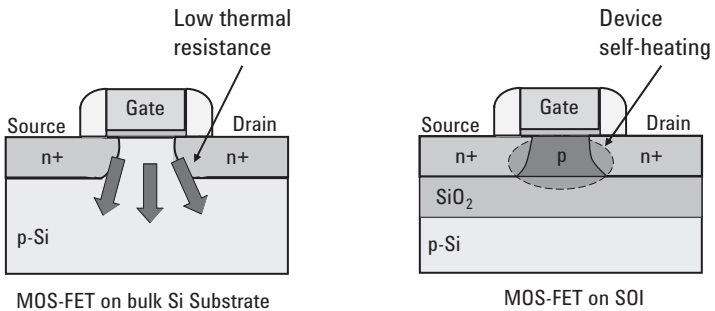
Figure 5.34. SOI based MOSFETs cannot easily dissipate the heat generated by continuous operation as can MOSFETs fabricated on bulk silicon.

The heat generated by measuring SOI transistors using DC measurement techniques lowers the electron mobility in the channel, which increases the turn-on voltage (Vth) for the MOSFET. Since the transistors are constantly cycling during actual circuit operation, measuring the Id-Vg characteristics of SOI MOSFETs with DC does not yield correct in-circuit estimate of device behavior. However, these types of transistors are so sensitive to device heating that the turn-on times must be very small (on the order of nanoseconds) with duty cycles that are << 0.1%.

MOSFET transistors that use high-k gate dielectrics as part of their gate insulation can experience electron trapping the gate dielectric.
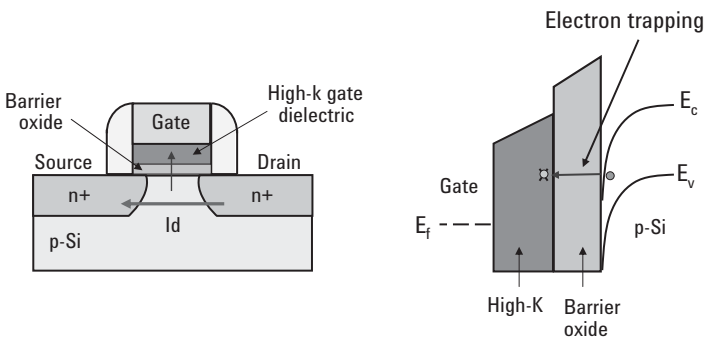
Figure 5.35. High-k gate dielectric MOSFETs are susceptible to electron trapping in the gate dielectric.

Electrons trapped in the boundary defects will work against the electric field applied by the transistor gate (at least in the case of an n-channel MOSFET). The result is that it takes more applied voltage to turn on the transistor, which increases the turn-on voltage (Vth) for the MOSFET. Fortunately, this electron trapping is not permanent and it can be reversed (or the electrons can be "de-trapped") by applying a negative voltage to the transistor gate. The time constant for electron trapping/de-trapping is on the order of 1 μs, so Id-Vg measurements need to be made on the order of nanoseconds to avoid seeing this effect.

Before continuing it is important to understand that Agilent's 10 ns pulsed IV measurement solution is specifically designed to characterize MOSFETs for the two cases just explained (SOI-based processes and high-k gate dielectrics). It cannot be used to make other types of pulsed IV measurements. However, most other pulsed IV measurement situations can be met using either the HV-SPGU or WGFMU solutions previously discussed. A basic schematic of Agilent's 10 ns pulsed IV solution is shown below.

$$I_d = \frac{V_{Drop}}{R_{Shunt}}$$
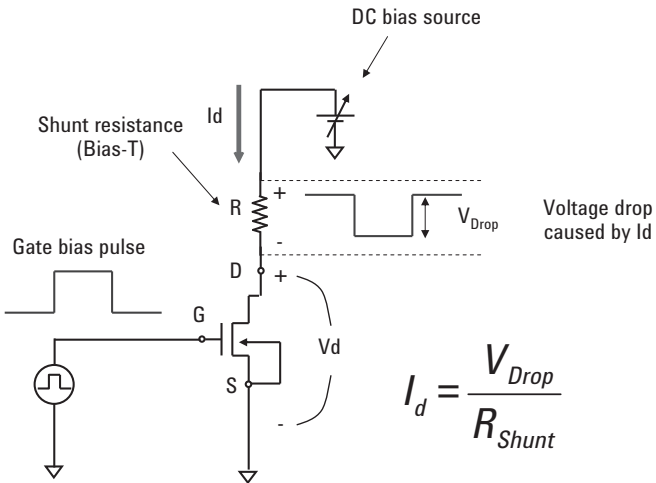
Figure 5.36. The basic circuit schematic outlining the operation of Agilent's 10 ns pulsed IV measurement solution.

While this solution is simple in concept, as previously mentioned it can be extremely difficult to implement in practice when using external instruments. Therefore, Agilent supplies the B1542A as a ready-made solution for these types of measurements. The B1542A is available in different options, and it supports a variety of semiconductor analyzers, oscilloscopes and pulse generators. You can run this solution from either EasyEXPERT running on the B1500A or from Desktop EasyEXPERT running on an external PC. The following two figures illustrate the key options of each of these two solutions.
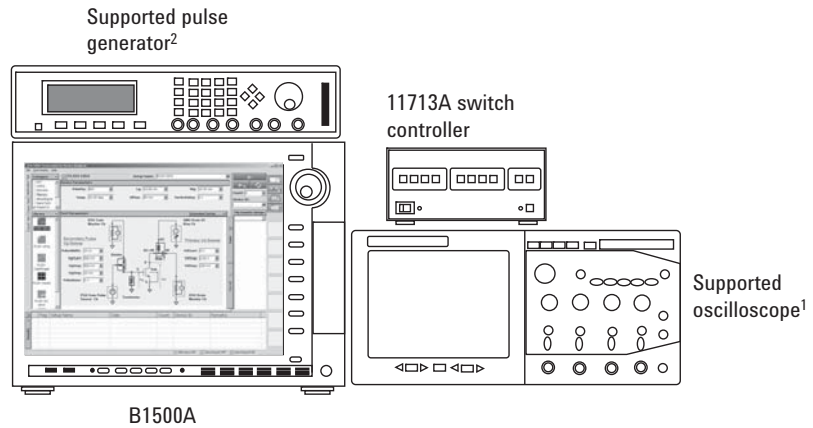
1. Supported oscilloscopes: 54850 series, MSO8000A series, DSO8000A series, DSO80000B series, DSO90000A series.
2. Supported pulse generators: 8110A, 81101A, 81110A.

Figure 5.37. The B1542A 10 ns pulsed IV solution using the B1500A as the system controller.
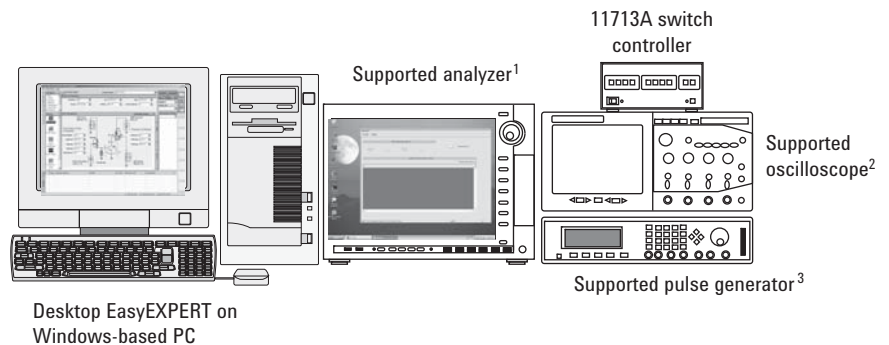
1. Supported analyzers: 4155B/C, 4156 B/C, E526xA, E5270B, B1500A.
2. Supported oscilloscopes: 54850 series, MSO8000A series, DSO8000A series, DSO80000B series, DSO90000A series.
3. Supported pulse generators: 8110A, 81101A, 81110A.

Figure 5.38. The B1542A 10 ns pulsed IV solution using an external PC running Desktop EasyEXPERT.

Of course, if you already own any or all of the supported equipment then all that you need to purchase are the B1542A software, cables and connectors. Note that a switch option is also available that allows you to switch between making standard IV measurements using SMUs and pulsed IV measurements using the pulse generator and oscilloscope without having to change any cabling.

The B1542A solution has a pulse verify feature that allows you to verify your gate voltage and drain current waveforms. By selecting and running this application test you can see the pulse shapes as measured by the oscilloscope. This makes it easy to verify the quality of the pulse shapes that you are generating and measuring. The following figure shows the results of an actual pulse capture performed using the B1542A solution.
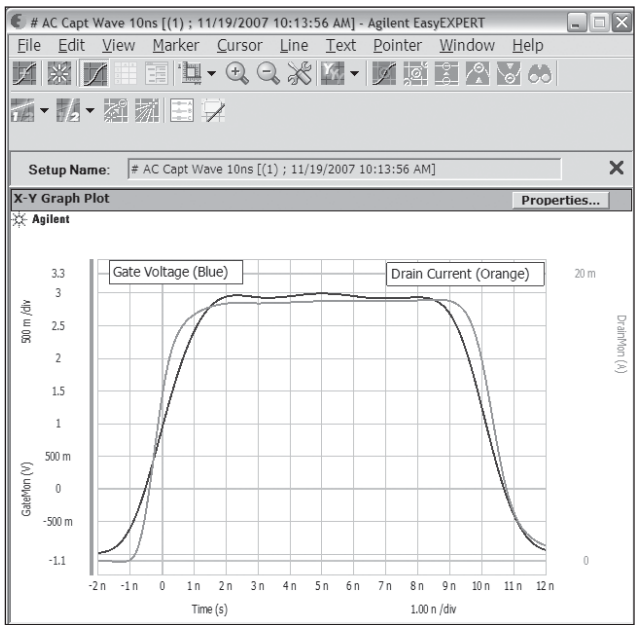
Figure 5.39. A plot of the gate and drain voltages showing the 10 ns pulse widths and fast pulse edge rates obtained using the B1542A pulsed IV solution.

As can be seen, the quality of the pulses is excellent. The rise and fall times are approximately 2 ns, and the pulse width is 10 ns.

The application test for performing pulsed Id-Vd measurement is shown below.
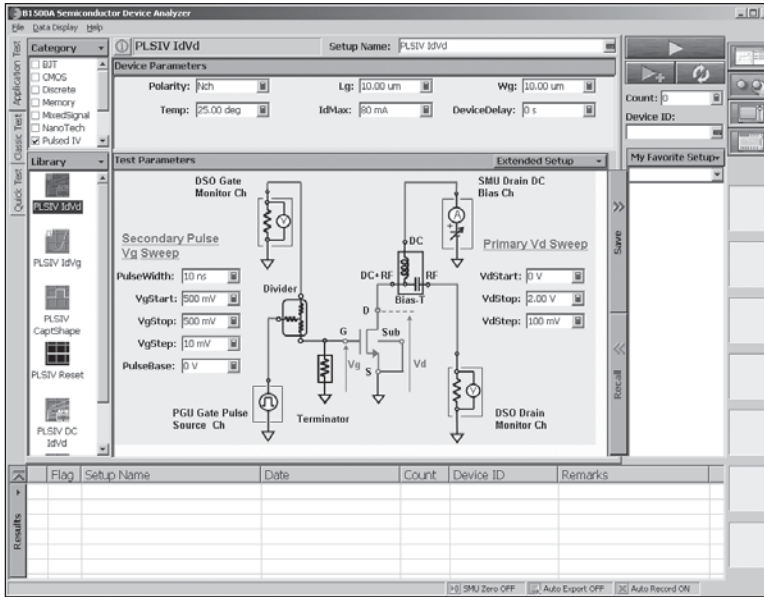
Figure 5.40. The B1542A 10 ns Id-Vd application test.

One simple method to determine the accuracy of the pulsed measurement scheme is to compare the measurement data taken using the pulsed IV solution with that of measurement data taken using SMUs on a standard transistor. Since a standard transistor does not have either the self-heating or electron trapping issues, the results of the two measurements should be the same. Comparisons of these two different types of measurements for an Id-Vd sweep on a standard transistor are shown below.
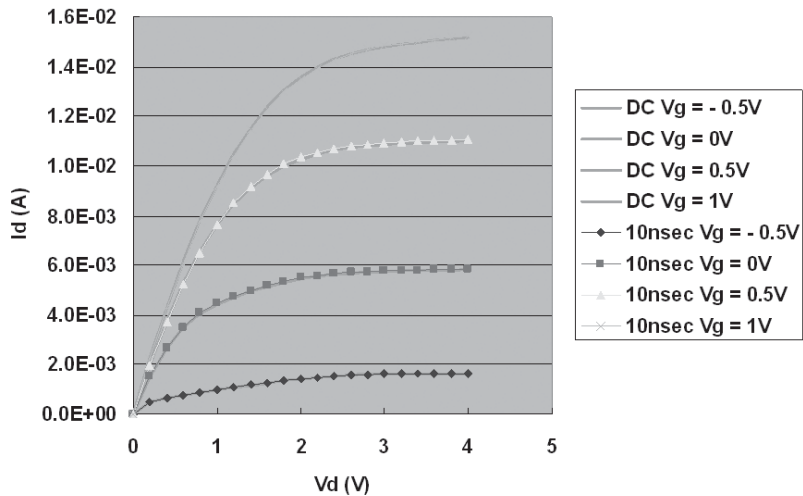
Figure 5.41. Comparisons of Id-Vd measurements taken on a standard transistor using both the pulsed IV solution and SMUs shows excellent correlation.

As these measurements show, the B1542A measurements give exactly the same results as do SMU-based measurements. This gives confirmation that pulsed IV measurements performed on devices with self-heating or electron trapping issues will yield accurate measurement data.

Summary of Agilent pulsed IV measurement solutions

Given the number of choices available, it is clear that a number of factors need to be weighed against one another when selecting a pulsed IV solution. The following steps are listed as a guide to help you with this process.

1. Make sure that you understand your measurement requirements.
    The key factors to consider are listed below:
    - Range of pulse widths
    - Required current measurement resolution
    - Maximum voltage and current output capability
    - Multiple pulsing capability
    - ***Note:*** Make sure that you understand that some solutions only work for specific device types and configurations.
2. Determine which solution or solutions meet your pulse width requirements, taking into account both your future as well as current needs.
3. If more than one solution meets your pulse width requirements, then select among the solutions using the other measurement parameters (current measurement resolution, current/voltage output capability, etc.). Please keep in mind that there may be some trade-offs among these various parameters (such as accuracy versus voltage/current output capability).
4. Once you have decided upon a solution, re-verify all of the specifications of that solution to make sure that it meets the measurement needs of your applications and devices.

As we discussed earlier in this chapter, SMUs can also make pulsed measurements. Although the SMU pulsing capabilities are slow relative to the three solutions discussed in this section, in certain cases (such as the on-the-fly NBTI measurement) SMU pulsed measurement can provide a very satisfactory and cost-effective measurement solution. Therefore, SMUs need to be included in any comparison of available pulsed measurement solutions. The following table summarizes the key specifications of the Agilent's various pulsed measurement solutions.

| Pulsed measurement solution	SMU	HV-SPGU	WGFMU	10 ns pulsed IV (B1542A)
Supported instruments	4155, 4156, B1500A	B1500A	B1500A	4155, 4156, E526xA, E5270B, B1500A
Maximum voltage	100 V[1,2] /200 V[3]	± 40 V	−10 V to 0 V or 0 V to 10 V	4.5 V (Gate) 10 V (Drain)
Maximum current	100 mA[1,2]/1 A[3]	400 mA	10 mA	80 mA
Current resolution	1 fA[2]/10 fA[1,3]	40 mA	0.014% of range[4]	1 mA
Pulse width range	500 μs − 2 s	5 μs − 10 s	100 ns (min.)[5]	10 ns − 1 ms
SMU measurement support	Yes	No	Yes	Yes (using 11713B)
Waveform monitor function	No	No	Yes	Yes
Multi-channel pulsing	No	Yes	Yes	No
External equipment required	None	None	None	Oscilloscope & PGU[6]
Connection	Direct	Direct	Direct	Bias-T/splitter

1. MPSMU
2. HRSMU
3. HPSMU
4. Requires averaging; see data sheet for more details.
5. The maximum pulse width is difficult to specify for this module, but it can easily be 20 million seconds (effectively infinite).
6. Supported oscilloscopes: 54850 series, MSO8000A series, DSO8000A series, DSO80000B series, DSO90000A series; Supported PGUs: 8110A, 81101A, 81110A.

Chapter 6: Making Accurate Resistance Measurements

"It requires a very unusual mind to undertake the analysis of the obvious."
— Alfred North Whitehead

Resistance measurement basics

Upon first thought it might seem strange to devote a chapter to resistance measurement. After all, resistance is governed by the simplest of all rules (Ohm's law) that every electrical engineering student learns in their first week of class:

$$V = I \times R \qquad \text{(Equation 6.1)}$$

However, despite the simplicity of this equation it turns out that measuring resistance accurately is actually one of the more challenging areas of parametric test. The reason for this is that the above equation is overly simplistic, and ignores the fact that resistance generates heat, which in-turn affects the value of the resistance itself. Therefore, it would be more accurate to re-write the above equation as:

$$V = I \times R(T) \qquad \text{(Equation 6.2)}$$

In this equation the resistance (R) is a function of temperature (T). This phenomenon, whereby the actual value of the resistance being measured changes due to heat generated by the current flow, is commonly referred to as the Joule self-heating effect.

Another factor requiring consideration is the resistance of the cables used to make a resistance measurement. When attempting to measure very small resistance values, Kelvin measurement techniques must be employed. The basics of Kelvin measurements have been explained in earlier chapters, and it is straightforward to apply these techniques to resistor measurements. However, it is worthwhile to point out that the combination of Joule self-heating effects and cable resistance make resistor measurements especially challenging. To reduce Joule self-heating you need to reduce the current (power) flowing into the resistor being measured. However, small currents flowing through the cable resistance require measurement equipment with very accurate voltage measurement capabilities in order to distinguish the cable voltage drop from the voltage drop across the actual resistor. Therefore, for these reasons resistor measurements can easily require a voltage measurement resolution capability of less than 1 millivolt.

The final factor to consider when making resistance measurements is electro-motive force (EMF). EMF is the technical name for the burst of electrical noise that is generated when a mechanical relay opens or closes. All SMUs have mechanical (reed) relays in their output paths that generate EMF when the SMU is activated. EMF can have significant impact upon measurement accuracy for all types of measurements, but it can especially impact low-level resistance measurements because they require extremely sensitive voltage measurement resolution.

Resistivity

Resistivity is a basic material property. If we measure the current flowing through a bar of homogeneous material with a uniform cross section when a voltage is applied to it, we can determine its resistance from the equation $R = V/I$. The related material property of resistivity can be calculated from the resistance measurement if the cross-sectional area (width × depth) and length of the sample are known as shown below.

$$R = \frac{V}{I} \qquad \rho = R\,\frac{A}{L} = R\,\frac{Wd}{L}$$
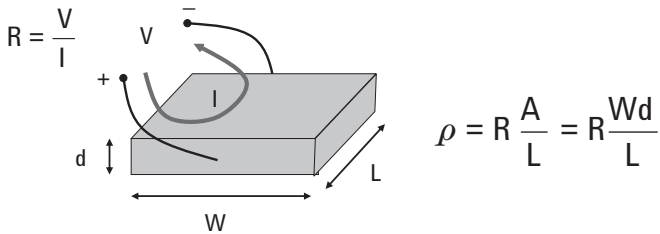
Figure 6.1. Calculating resistivity from a resistance measurement.

Sheet resistivity is typically measured for all of the implantation layers as well as the metal interconnect layers. It is common to see resistivity expressed in "Ohms per square", in recognition of the fact that the resistance from one side of a square of homogeneous material of uniform depth to the other side is always the same regardless of the size of the square. A little reflection on the effect of combining squares of material of uniform thickness together into larger squares should convince the reader of this fact.

$$R_S = (R_S \parallel R_S) + (R_S \parallel R_S)$$
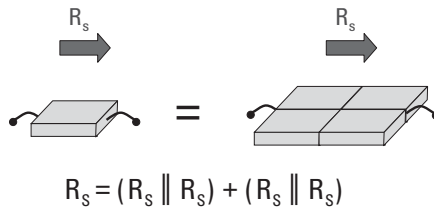
Figure 6.2. The Ohms per square of a homogenous material of uniform thickness is always the same regardless of the size of the square.

Van der Pauw test structures

The resistances measured when making resistivity measurements can be extremely small (on the order of milliohms), which mandates the use of Kelvin measurement techniques. Therefore, it is extremely convenient to be able to combine the Kelvin measurement technique with a simple geometry in order to determine the sheet resistance of a given material. The configuration universally used to achieve this is known as a Van der Pauw structure, named after L.J. Van der Pauw who first proposed this measurement technique in 1958. The Van der Pauw structure consists of a square sample with an electrical contact at each of the corners as shown below.
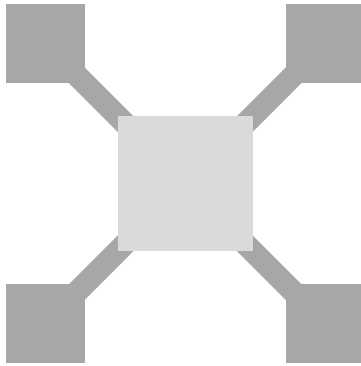
Figure 6.3. The Van der Pauw test structure for determining resistivity.

Van der Pauw showed in his original paper that the following relationship holds:

$$\exp \frac{-\pi d R_{AB,CD}}{\rho} + \exp \frac{-\pi d R_{BC,DA}}{\rho} = 1$$

Here d is the thickness of the sample, ρ is the sheet resistance of the sample, and the resistance measurements are defined as shown below.
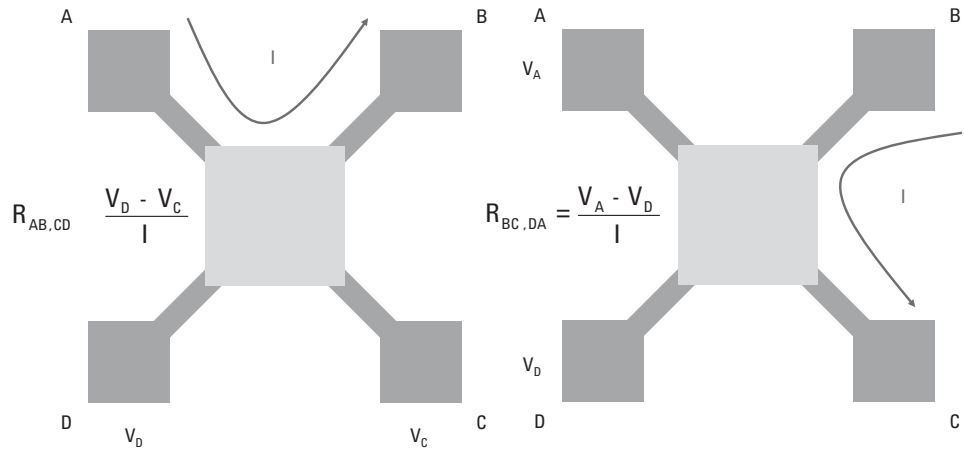
Figure 6.4. Calculating sheet resistance from a Van der Pauw test structure.

The calculation of the resistivity from this equation and these measurements is straightforward.

120

Accounting for Joule self-heating effects

Joule self-heating limits resistance measurement accuracy because it changes the value of the resistor being measured. However, once you select the maximum level of resistance change that you can tolerate you can translate this into maximum allowable power dissipation. We know that the equation for power dissipated in a resistor is given by:

$$P = V \times I = V \times \frac{V}{R} = \frac{V^2}{R}$$

By rearranging the above equation and solving for voltage we can convert the maximum power dissipation into a minimum acceptable voltage resolution necessary to meet the resistance measurement requirements. The equation describing this relationship is given by:

$$V_{max} = \sqrt{P_{max} \times R(T)}$$

A simple example illustrates how this process works. The maximum allowed power dissipation is determined by the thermal resistance of the underlying silicon dioxide layer. Silicon dioxide (SiO_2) has a thermal conductivity of $0.014 W/°C\text{-}cm^2$. We can use this to calculate out the thermal resistance as shown below:

Thermal conductivity   Thermal resistance
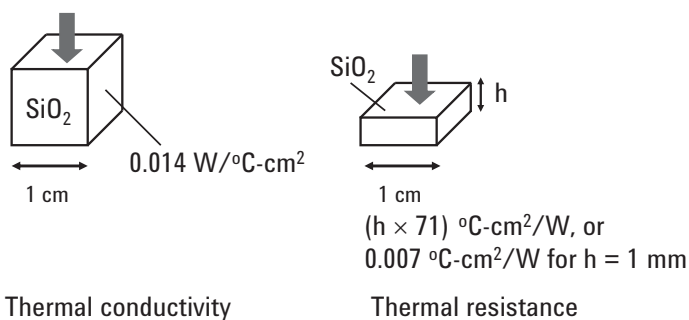
Figure 6.5. The thermal properties of silicon dioxide.

This means that if you apply 1 W to a 1 cm square metal structure of this thickness the temperature will rise by 0.007 °C; if you apply 1 W to a 10 μm square metal structure of this thickness then the temperature will rise by 7000 °C. Of course, in this example the material on top of the SiO_2 would vaporize long before a temperature of 7,000 °C was reached. This example is really here to emphasize the fact that one cannot arbitrarily increase the power (= current) into a resistor in an attempt to compensate for insufficient accuracy in the measurement hardware. At or near room temperature the resistance of a Cu or Al metal line changes by about 0.35%/°C. For the case shown above, we can compute the maximum amount of power that can be dissipated to produce a 0.1% change in resistance for a Cu or Al metal square 10 mm by 10 mm.

$$0.1\% = P_{max} \times 0.007\ ^\circ C - cm^2/W \times 0.35\% /^\circ C \times 1/(10\,mm)^2 \times (10\,mm/1\,cm)^2$$
$$\therefore P_{max} = 0.04\ mW$$

To achieve 0.1% accuracy in a copper Van der Pauw structure with an equivalent resistance of 10 mΩ per square (1 mm thick film) we have the following:

$$V_{max} = \sqrt{(0.04\ mW) \times (10\ m\Omega)} \approx 0.000632\ V$$

In other words, the measurement instrumentation must be able to resolve down to about 1 mV! *Note:* The Van der Pauw structure is a relatively extreme example. A metal line 1 mm wide and 100 mm long typically requires that the measurement instrumentation be able to resolve to the tens of micro-volts in order to obtain 0.1% accuracy. Since we know that the maximum allowable voltage is proportional to the square root of (Pmax × R), the required voltage resolution increases as the square root of the resistance value. Since this is a relatively weak dependence on R, 10 mV is a good estimate of the measurement resolution requirement for most resistance measurements.

Eliminating the effects of electro-motive force (EMF)

What is EMF?

Electro-motive force (EMF) is the name given to the transient voltage pulse created when reed relay switches open and close. Conventional reed relay switches, which can be obtained from a variety of sources, typically generate a thermo-EMF ranging from a few tens of micro-volts to a few hundreds of micro-volts after the relay activation current is turned on or off. This voltage drift, which can continue for several minutes before dying out, is not acceptable when making precision measurements such as those required for parametric test. The figure shown below is an example of the thermal-EMF generated by a commercially available reed relay.

Thermo-EMF of typical commercial reed relay
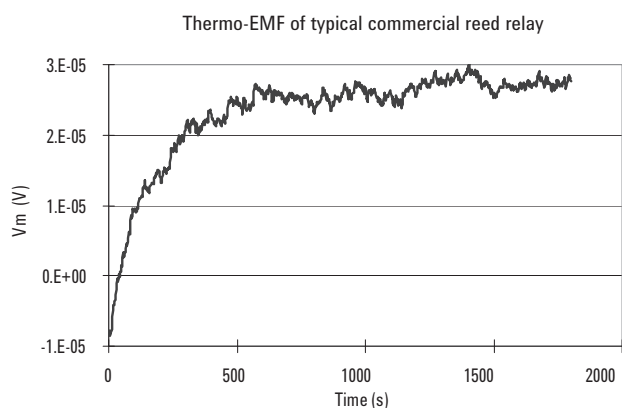
Figure 6.6. This graph shows the thermo-EMF generated by a typical reed relay.

It is important to note that the above graph is not characteristic of the reed relays used in parametric measurement instruments (for obvious reasons). Agilent Technologies has developed a proprietary reed relay technology known as the "Cool Guard" relay. These relays are used in parametric measurement equipment to minimize the effects of EMF, thereby improving both throughput and measurement accuracy. The characteristics of a Cool Guard relay are shown below.
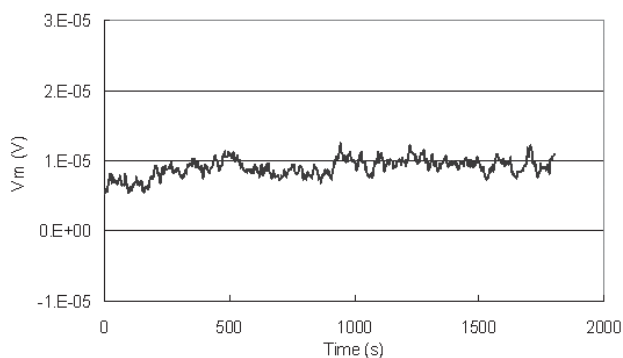
Figure 6.7. This graph shows the thermo EMF generated by an Agilent Cool Guard relay.

The SMU schematic discussed previously in Chapter 3 left out the output relay switch for purposes of simplicity. However, it is important to understand that all SMUs have a switch in their output path, and that this switch generates EMF when opened or closed. The revised SMU schematic is shown below.
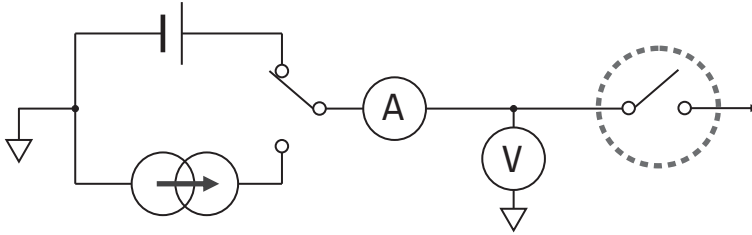
Figure 6.8. SMU schematic showing the output connection relay.

When using the SMUs from the instrument front panel, the software takes care of opening and closing this relay so its presence is invisible to the user. However, when using instrumentation under program control the user must issue a connect ("CN") command in order to close this relay, and the relay remains close until a command is issued (such as "*RST") to open the relay(s).

Mitigating the effects of EMF

Reducing the impact of EMF on parametric measurements is relatively straight-forward. It should be pointed out that the data sheet specifications take thermo EMF into account, so under normal measurement conditions you do not need to be concerned about thermo EMF. However, when performing extremely sensitive measurements or when trying to extend the measurement accuracy of your parametric measurement equipment, the following guidelines can minimize or eliminate thermo EMF effects:

1. Make sure that the output relay is closed and that the relay has stabilized to its final value. This means that the instrument should be warmed up and that the SMU output relay should remain closed.

2. Complete any measurements as quickly as possible (within less than 10 seconds). Doing this will keep the measurement drift to within a few microvolts.

One valuable resistance measurement technique that eliminates both the effects of EMF as well as any offset voltages in the measurement equipment is to force current and measure the resistance twice using the Kelvin measurement technique, reversing the polarity of the force current for the second measurement. If you work out the Kirchhoff's current law (KCL) and Kirchhoff's voltage law (KVL) equations for this circuit, you can see that by averaging these two measurements the offset voltages and EMF voltages cancel out and you get the true value of the resistance.
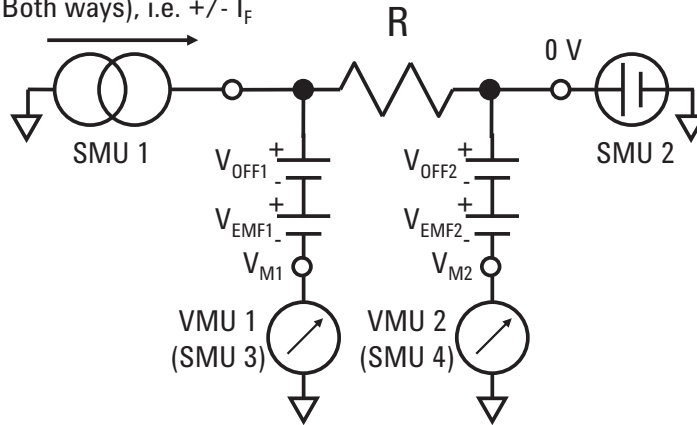
Figure 6.9. By measuring resistance twice and averaging the two resistor values the effects of EMF and voltage offsets are eliminated.

$$R = \frac{(V_{M1} - V_{M2})}{I_F}$$

In summary, the key points of the above technique are:

1. Make sure that you keep the applied current small enough to eliminate any Joule self-heating effects.
2. Measure the resistance twice by reversing the polarity of the applied current and average the two results.
3. If available, use a voltmeter rather than an SMU to measure the voltage since voltmeters generally have superior voltage measurement accuracy.

Chapter 7: Diode and Transistor Measurement

"Choose a job you love, and you will never have to work a day in your life"
— Confucius

Introduction

It is not the intent of this handbook to teach a course on semiconductor device physics as there are already an abundance of excellent textbooks available on this subject. However, it is difficult to discuss making parametric diode and transistor measurements without first spending a little time understanding their operation. Therefore, we will give a brief review of pn junctions, diodes, and MOS and bipolar transistor operation with an emphasis on how we characterize them in parametric test as opposed to detailed theoretical derivations.

PN junctions and diodes

Review of PN diode operation

Intrinsic semiconductor materials (such as silicon) do not have an abundance of either electrons or electron holes. However, silicon can be doped with other materials such that it becomes either n-type (possessing excess electrons) or p-type (possessing excess electron holes). When considered individually these materials are not particularly interesting. However, consider the case shown below when these two materials are brought into close contact.
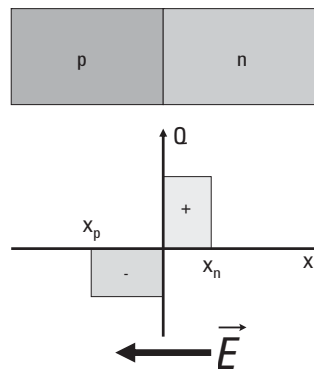
Figure 7.1. The cross section of a pn junction assuming an abrupt change from p-doped to n-doped material. The graph shows the fixed charge remaining after the mobile carrier diffusion has stabilized.

Assuming the extremely idealized case of an abrupt junction (i.e. one that instantaneously transitions from p to n material) as shown in Figure 7.1, we can see that something very interesting happens. The force of diffusion causes holes from the p-type material to flow into the n-type material (leaving behind fixed negative charge), and similarly the force of diffusion causes electrons from the n-type material to flow into the p-type material (leaving behind fixed positive charge). This diffusion process will continue until the electric field created by the fixed charge in what is normally called the space-charge region becomes strong enough to exactly balance the diffusion tendencies of the mobile carriers.

The one-dimensional (x-axis) equations defining current flow in semiconductor are shown below.

$$J_n = q\mu_n E_x + qD_n \frac{dn}{dx}$$ *(Equation 7.1)*

$$J_p = q\mu_p E_x - qD_p \frac{dp}{dx}$$ *(Equation 7.2)*

Where   J is the current density of electrons (n) and holes (p)
q is the electron charge
E_x is the electric field in the x-dimension
μ is the mobility of electrons (n) and holes (p)
D is the diffusion constant for electrons (n) and holes (p)
n is the electron density
p is the hole density

These equations basically state what was alluded to in the previous discussion of an abrupt pn junction. Namely, current flow in a semiconductor consists of two parts: a drift current proportional to the applied electric field and a diffusion current proportional to the spatial first derivative of the mobile carrier density. In addition to the above current flow equations we also have the Einstein relationship which relates the ratios of the mobility and diffusions constants as shown below.

$$\frac{D_n}{\mu_n} = \frac{D_p}{\mu_p} = \frac{kT}{q}$$ *(Equation 7.3)*

Where   q is the magnitude of the electron charge (1.602×10^{-19} Coulomb)
k is Boltzmann's constant (1.38×10^{-23} J/K)
T is the absolute temperature [deg K]

The general form of Poisson's equation relates the second derivative of the electric potential to the total space charge density (ρ). Since we know that in a semiconductor this has to be related to the densities of mobile and fixed charge, we can write this as follows.

$$\frac{d^2\varphi}{dx^2} = -\frac{q}{\varepsilon_{Si}}(p - n + N_d - N_a)$$ *(Equation 7.4)*

Where   Nd is the donor density concentration
Na is the acceptor density concentration
ε_{Si} is the permittivity of silicon

In the case of the abrupt junction shown in Figure 7.1 we make what is known as the depletion approximation, which assumes that the semiconductor is divided into distinct regions which are either completely neutral or completely depleted of mobile carriers. Therefore, in the depletion region we can write the above equation as follows.

$$\frac{d^2\varphi}{dx^2} = -\frac{q}{\varepsilon_{Si}}(N_d - N_a)$$ *(Equation 7.5)*

Using the depletion approximation we can integrate the above equation to get the electric field in both the p and n regions as shown below.

$$E_x(x) = -\frac{qN_a}{\varepsilon_{Si}}(x + x_p) \qquad -x_p \leq x \leq 0 \qquad \text{(Equation 7.6)}$$

$$E_x(x) = -\frac{qN_d}{\varepsilon_{Si}}(x_n - x) \qquad 0 \leq x \leq x_n \qquad \text{(Equation 7.7)}$$

Where x_p is the width of the space charge in the p region (see Figure 7.1)
         x_n is the width of the space charge in the n region (see Figure 7.1)

Graphically, these equations have the appearance shown below.
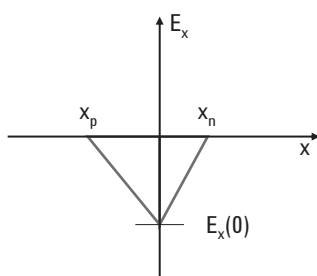
Figure 7.2. The electric field in an abrupt pn junction under the depletion approximation assumption.

We know that the electric field has to be continuous at x = 0.

$$E_x(0) = -\frac{qN_d x_n}{\varepsilon_{Si}} = -\frac{qN_a x_p}{\varepsilon_{Si}} \qquad \text{(Equation 7.8)}$$

This gives us the result shown below.

$$N_d x_n = N_a x_p \qquad \text{(Equation 7.9)}$$

Equation 7.9 shows an important characteristic of pn junctions: the width of the depletion region varies inversely with the magnitude of the dopant concentration. In other words, higher dopant concentrations result in narrower space charge regions.

When no voltage is applied to the pn junction a barrier exists to current flow and the diode acts as an open circuit. The derivation of the current flow equations are involved and beyond the scope of this text. However, it should be somewhat intuitive that as we apply a positive voltage (i.e. electric field) to the p-region we are acting to reduce the built-in electric field of the pn junction. At some point the electric field is reduced enough to allow current to flow through the pn junction.
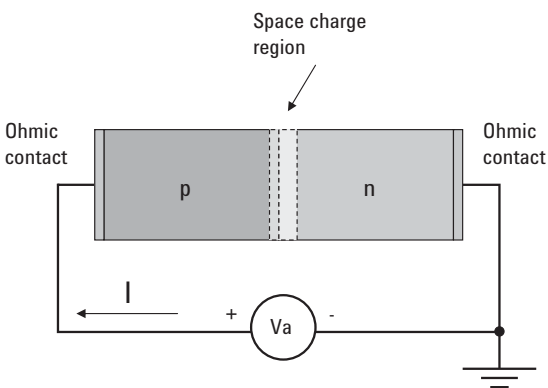
Figure 7.3. The behavior of a pn junction under positive applied bias.

Therefore, without any detailed derivations we will ask the reader to take it on faith that the current flow through a pn diode exhibits exponential dependence upon applied voltage (Va) and is given by the equation shown below.

$$I = I_o \left(e^{\left(qV_a / kT \right)} - 1 \right)$$

(Equation 7.10)

This is sometimes called the ideal diode equation. It predicts a saturation current of $-I_o$ for negative values of Va and an exponentially rising current for positive values of Va. To emphasize that a diode only conducts current in one direction, it has the circuit symbol shown below.
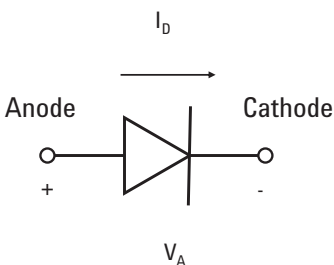
Figure 7.4. The circuit symbol for a diode.

The p-doped region is denoted as the anode, and the n-doped region is denoted as the cathode. Note: In actuality diodes can conduct current in both directions. However, typically much larger voltages need to be applied to the cathode (relative to the anode) in order for current flow to occur in the reverse direction. In this condition the diode is said to "breakdown", which is a logical term for this phenomenon since it is an aberration from normal diode behavior. The physics of semiconductor junction breakdown will not be discussed in this handbook, but later some practical measurement examples will be explored.

The "Ohmic contacts" shown at the ends of the diode in Figure 7.3 simply mean that the semiconductor material is heavily doped enough such that the metal to semiconductor contact does not present any sort of barrier to the flow of current. If the semiconductor material is lightly doped the metal to semiconductor contact can actually behave as another form of diode known as a Schottky barrier diode. Current flow in a Schottky barrier diode has a dependence on applied voltage as shown below.

$$ I = I'_o \left[e^{\left(qV_a / nkT \right)} - 1 \right] $$

(Equation 7.11)

In Equation 7.11 "n" is a constant usually ranging from between 1.02 and 1.15. The prime symbol is present on I_o to emphasize that this constant is different in value from that for the case of a pn junction. Schottky diodes typically have an effective "turn-on" voltage that is several hundred millivolts less than that of a pn junction diode, which makes them essential in the design of bipolar logic circuits since they can keep the base to collector junction from forward biasing and therefore keep the transistor out of saturation.

One important point to note about pn junctions is that they behave as voltage-dependent parallel plate capacitors, since as we apply external voltages we modify the charges in and around the space-charge region. Therefore, junction capacitance is one important parameter that must be characterized for all semiconductor devices, since this impacts the speed at which the devices will switch when used in an integrated circuit. However, since capacitance measurement is much more challenging to perform correctly than simple current and voltage (IV) measurements, we will defer a detailed discussion of semiconductor capacitance measurement to Chapter 8.

Basic diode characterization

Diodes are relatively simple devices to characterize. From Equation 7.10 we can see that a plot of the log of the diode current (Id) should be linear with respect to applied voltage. A plot of diode current and the log of the diode current for a "typical" diode are shown below.
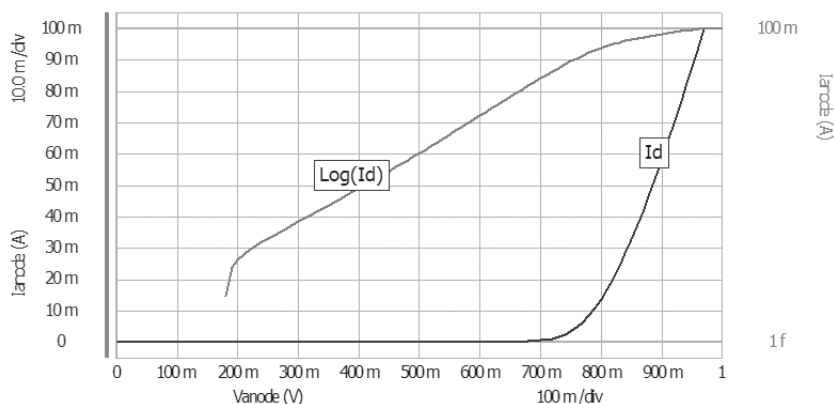
Figure 7.5. A pn diode sweep in the forward direction plotting both Id and Log(Id).

Of course, another important parameter is the reverse breakdown characteristics of the pn diode. A plot of this is shown below.
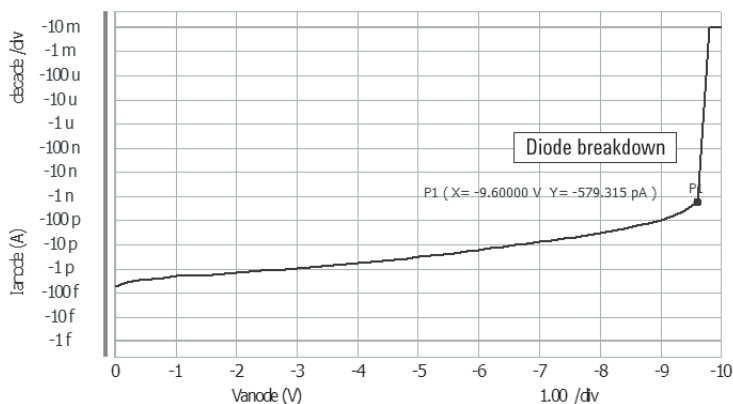
Figure 7.6. The reverse breakdown characteristics of a diode.

Using various processing techniques it is possible to control the reverse breakdown characteristics of certain classes of diodes very precisely. Diodes with these sorts of precisely controlled breakdown characteristics are known as zener diode. This has some obvious beneficial uses in circuit design, since it permits the zener diodes to be used as voltage clamps within the circuit.

MOS transistor measurement

Silicon-based MOS transistors are the dominant technology for digital electronics for many reasons, foremost of which is their ability to scale relatively easily (at least until recently) with improvements in lithography. The operation of a MOS transistor is more intuitive and easier to explain than that of a bipolar transistor. Essentially in a MOS transistor we have two highly doped contact regions (the source and the drain) separated by a channel region of opposite doping type. By applying a voltage to the gate, which is located above the channel region, we can induce charge in the channel region such that we create a conductive path between the source and the drain.

Review of MOSFET operation
A simplified cross-section of an n-channel type MOS (NMOS) transistor with positive bias on the gate and the source grounded is shown below. **Note:** For now we are also assuming that the voltage applied to the bulk is zero volts.

Figure 7.7. An n-channel MOS transistor with gate bias greater than the threshold voltage but with negligible drain to source voltage.

In this example, the voltage applied to the drain (Vd) is small relative to the gate voltage (Vg). As long as Vg exceeds the threshold voltage necessary to invert the channel (Vt) a conductive channel exists between the source and drain. In this situation the drain current (Id) is given by the charge in the channel (Q_{ch}) divided by the time it takes the carriers to transit across the channel (T_{tr}).

$$Id = -\frac{Q_{ch}}{T_{tr}}$$

(Equation 7.12)

The channel charge is dependent on the channel capacitance, area and applied gate voltage. The transit time is dependent on the drift velocity of the carriers and the channel length. In this regime the value of the drain current is approximated by the equation shown below.

$$Id = \mu_n \frac{W}{L} C_{ox}(V_g - V_t) \bullet V_d$$

(Equation 7.13)

Where μ_n is the electron mobility
       W is the channel width
       L is the channel length
       Cox is the capacitance per unit area in accumulation

If we increase the drain voltage (Vd) such that it is no longer negligible relative to the gate voltage then the space charge region and the channel are no longer spatially uniform as shown below.

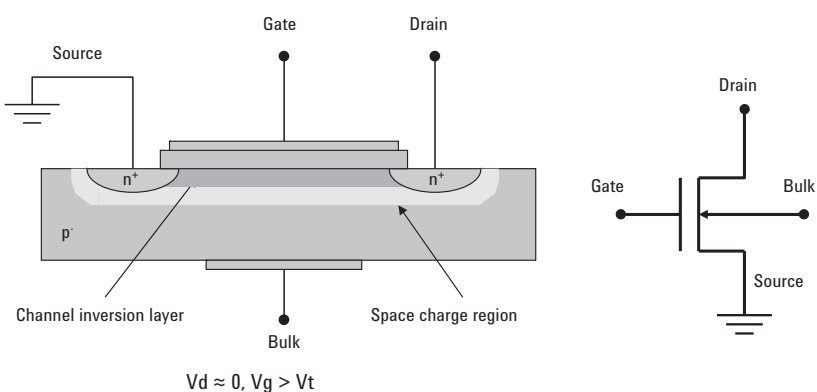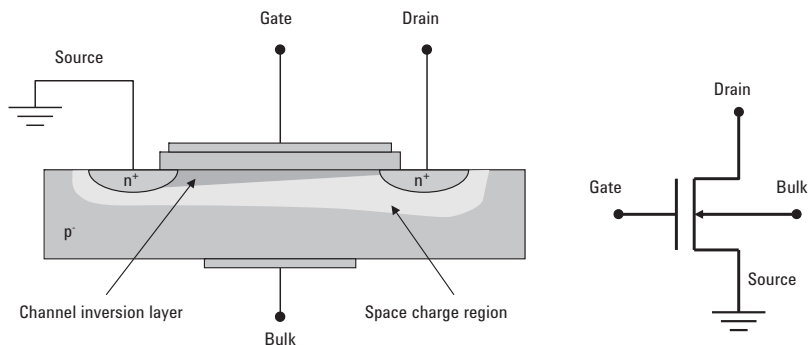0 < Vd < (Vg – Vt), Vg > Vt

Figure 7.8. An n-channel MOS transistor with gate bias greater than the threshold voltage but with non-negligible drain to source voltage.

The simple assumptions used to derive the previous equation for Id are no longer valid since the drain voltage acts to reduce the inversion layer near the drain. In essence this means that the inversion layer becomes more trapezoidal than rectangular. We can approximate this by assuming the average voltage in the channel to be (Vg – Vd/2). This leads to the revised equation for Id shown below.

$$Id = \mu_n \frac{W}{L} C_{ox} \left(V_g - V_t - \frac{V_d}{2} \right) \bullet V_d \qquad \text{(Equation 7.14)}$$

This equation defines the behavior of Id in the non-saturation region of operation.

Plotting out the previous equation for various values of Vg results in a series of downward-facing parabolas with maxima at the point where Vd = (Vg - Vt) as shown below.
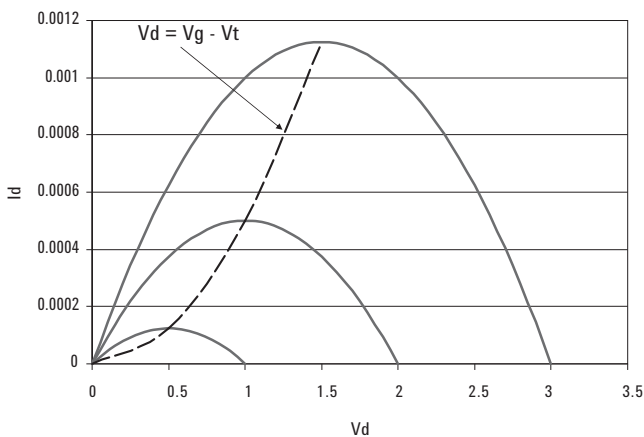
Figure 7.9. Plotting out the value of Id using the previous equation for different values of Vg yields a series of downward facing parabolas.

133

It should be obvious that this does not make any physical sense for values of Vd greater than Vg - Vt, since this would imply a negative transconductance (transconductance being the first partial derivative of Vd with respect to Vg). To understand what is going on we need to consider what happens when Vd becomes greater than (Vg - Vt).

Vd > (Vg – Vt), Vg > Vt
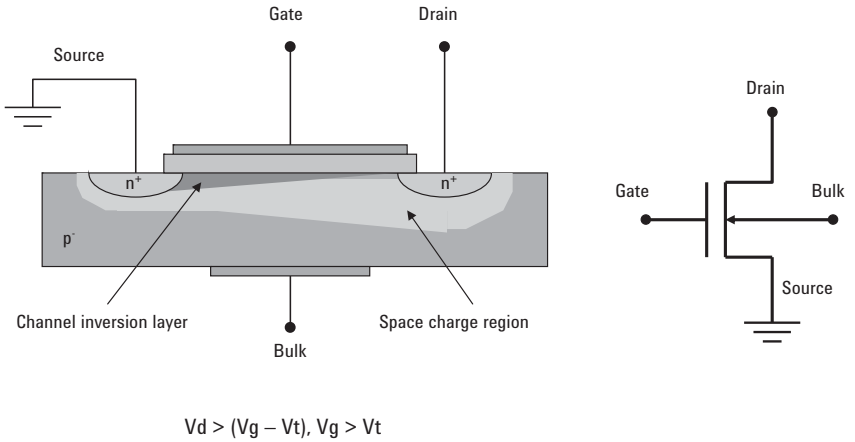
Figure 7.10. An n-channel MOS transistor with gate bias greater than the threshold voltage and with Vd greater than (Vg – Vt).

When Vd becomes greater than (Vg - Vt) the inversion layer becomes pinched off at the drain. However, as the electrons approach the drain there is no barrier to stop them and they are rapidly accelerated by the high electric field into the drain region. In this situation the current is determined by the rate at which electrons reach the edge of the depletion region, which is in-turn determined by the maximum electron drift velocity. Since the maximum drift velocity is (at least to first order) insensitive to Vd, once Vd reaches the value of (Vg - Vt) the drain current becomes saturated (constant). The equation for the drain current when in saturation is given below.

$$Id = k\frac{W}{L}\left(V_g - V_t\right)^2$$

(Equation 7.15)

A plot of the Id versus Vd characteristics for an n-channel MOSFET is shown below.
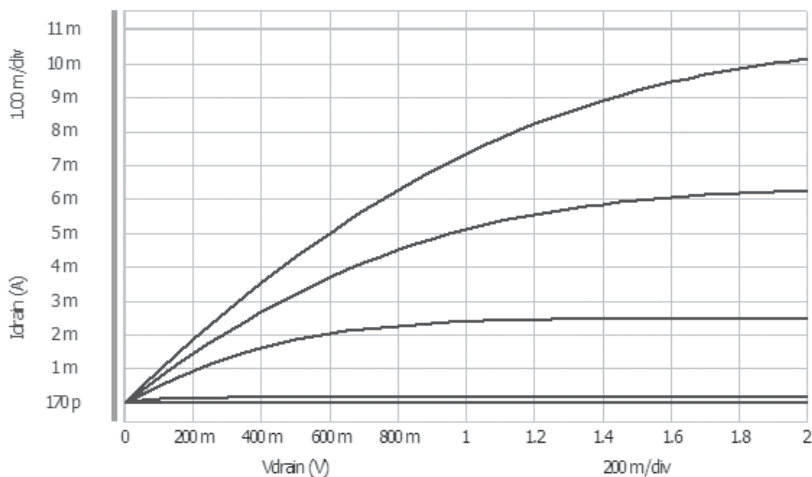
Figure 7.11. A "typical" Id versus Vd characteristic for an n-channel MOSFET.

The value of the constant k is proportional to the product of μ_n and Cox but is generally not exactly the same. In actuality the drain current does exhibit a slight positive dependence on the drain voltage, since the effective channel length (L) does decrease slightly for larger values of Vd. Many times the k, W and L factors are simply absorbed into another constant and written as shown below.

$$Id = \beta \bullet \left(V_g - V_t \right)^2$$ *(Equation 7.16)*

 Of course, the previous discussion of MOS transistor behavior was extremely simplified. However, the modification to the previous equations for other conditions is relatively straightforward. In the case where the transistor source terminal is not at ground it is a correct assumption that (at least to first order) an increase in Vs will have the effect of reducing the influence of the gate voltage. Therefore, it is reasonable to modify the equation for Id in saturation for the case where Vs ≠ 0 as follows.

$$Id = \beta \bullet \left(V_g - V_s - V_t \right)^2$$ *(Equation 7.17)*

135

Basic MOSFET characterization

From the equations we developed for MOSFET behavior we can see that a plot of the square root of Id versus Vg should be linear.

$$\sqrt{Id} = \sqrt{\beta} \cdot \left(V_g - V_s - V_t\right)$$

(Equation 7.18)

In addition, the intercept of this line with the x-axis (Vg) should equal the threshold voltage (when the source is grounded). Therefore, one simple technique for determining the threshold voltage (Vt) of a MOS transistor is to plot the square root of Id and extrapolate this to its intersection with the x-axis.

MARKER(1.2500000000E+000 V 3.1151725474E-002 ------)
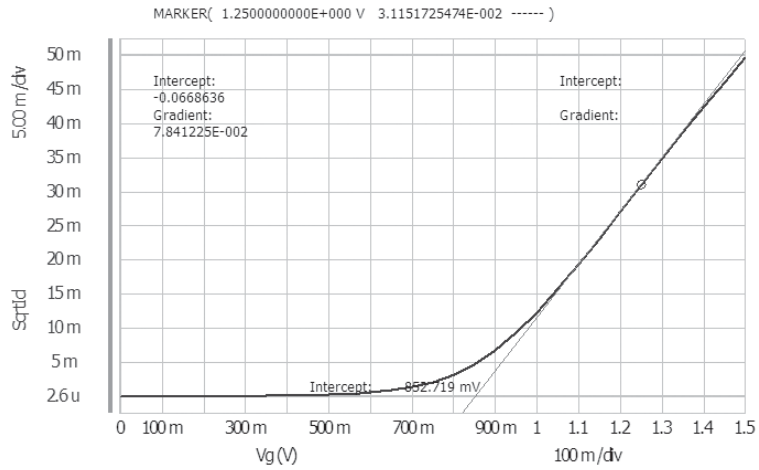
Figure 7.12. Determining the Vth of a MOSFET by plotting the square root of Id and extrapolating the slope back to the x-axis.

Note: In reality there is some subthreshold leakage current flowing in the transistor for Vg values less than Vt, so a tangent line needs to be extrapolated to the x-axis from a point on the square root of Id plot where Vg >> Vt.

The effect of a non-zero bulk voltage can be understood by examining its impact upon the charge within the channel. Consider the case shown below where a negative voltage is applied to the bulk.
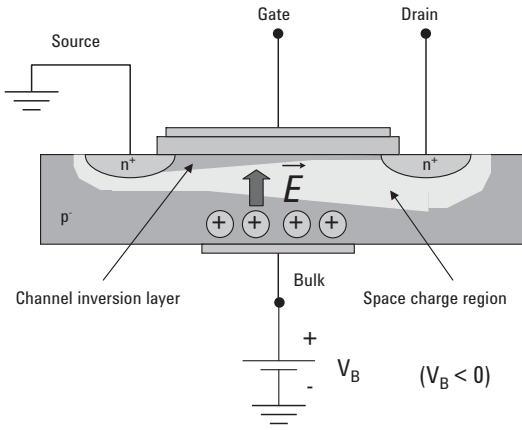
Figure 7.13. The effect of applying a negative bias to the n-channel MOS transistor bulk terminal.

Applying a negative charge on the bulk will induce positive charge to accumulate near the bulk contact. The effect of this positive charge is to act in opposition to the voltage applied to the gate, effectively requiring more gate voltage to achieve channel inversion than for the case where the bulk is at zero volts. Therefore, we would expect to see Vt increase with decreasing values of Vb (which is commonly known as the "body effect").

Similarly, applying a positive charge on the bulk will cause Vt to decrease. The equations that describe this behavior are beyond the scope of this simple handbook, but a plot of Id versus Vg for several values of Vb shows that physical reality matches our intuition.
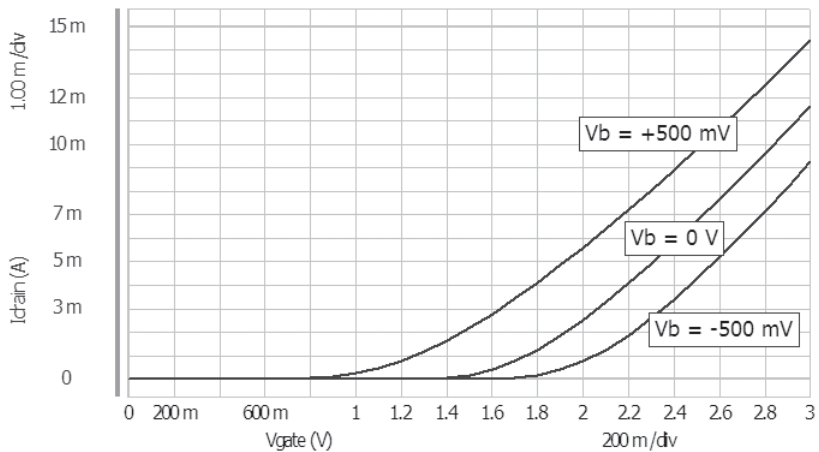
Figure 7.14. The effect of the bulk voltage on threshold voltage (Vt).

As can be seen from this plot, negative values of Vb shift the curve to right which implies an increase in the threshold voltage (Vt). Similarly, positive values of Vb shift the curve to the left which implies a decrease in the threshold voltage.

For power MOSFETs the on-resistance, Rds(on), is an important parameter. Rds(on) is defined as the forced drain voltage (Vds) divided by the drain current (Id) in the region where the MOSFET is in the on state. The B1505A can easily calculate the Rds(on) using its HCSMU and the built-in analysis functions in EasyEXPERT. In the following example an Rds-Vgs plot is shown. A pulsed drain current of 6 A is forced while the gate voltage (Vgs) is swept from 0 V to 20 V and the drain voltage is measured. From this measurement it can be seen that when Vgs is lower than 4 V the instantaneous value of Rds is large because the MOSFET is not fully turned on.
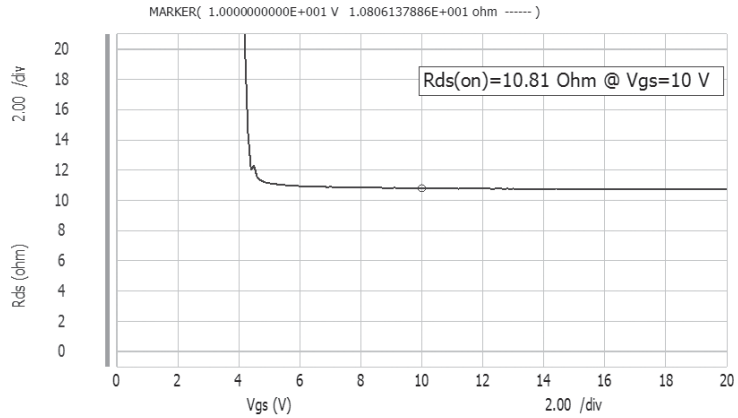
Figure 7.15. The B1505A's HCSMU module makes it easy to measure the on-resistance characteristics of a power MOSFET.

The drain-source breakdown voltage (BVdss) and drain cutoff current (Idss) are also very important parameters for high-power MOSFETs. To determine BVdss the gate and source are both shorted to ground and the drain voltage is swept from zero volts to a very large value. Since the voltage required to reach break-down is typically over 1000 V for power MOSFETs, the HVSMU of the B1505A is the only choice to make this measurement. EasyEXPERT's analysis functions make it easy to automatically display the BVdss once a current level (1 μA in this case) to define the onset of breakdown has been selected. Similarly, once a voltage level (1500 V in this case) has been selected to define Idss it too can be automatically displayed. The following plot illustrates this measurement.
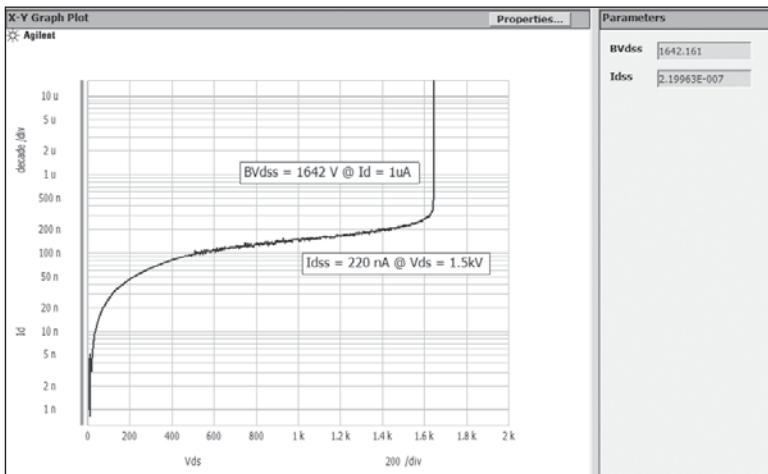
Figure 7.16. The B1505A's HVSMU module and EasyEXPERT software make it easy to measure the BVdss and Idss of a power MOSFET.

Bipolar transistor measurement

With the predominance of CMOS technologies in integrated circuits, bipolar transistors have largely been relegated to specialized (albeit not necessarily unimportant) niche applications. The most common uses for bipolar transistors today are in microwave circuitry, high-power applications and analog circuitry.

Review of bipolar transistor operation
The operation of a bipolar transistor is actually quite a bit more complicated to explain (and less intuitive) than that of a MOS transistor. The configuration of a bipolar transistor consists of three dopant regions of alternating type (either NPN or PNP). The following discussion will be centered on the operation of an NPN transistor which has the circuit symbol shown below.
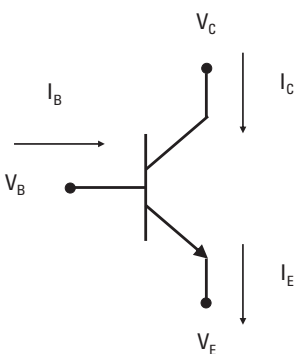
Figure 7.17. The circuit symbol for an NPN bipolar transistor.

Note that the base to emitter junction contains an arrow similar to that for the representation of a pn diode, and in fact the base and emitter will function as a pn diode. Similarly, the base to collector junction will also function as a pn diode. However, while it is true that any bipolar transistor will function in both the forward and reverse directions, due to intentional asymmetrical doping of the junctions optimal performance is obtained when the base to emitter junction is forward biased.

We will now consider the case where we have the emitter of the NPN transistor at ground and we apply a positive voltage to the base. We are also assuming the transistor collector is at a voltage much greater than that of the base ($V_C \gg V_B$). If the base to emitter voltage (V_{BE}) is sufficiently large enough to forward bias the base to emitter junction, then current will flow in the transistor as shown below.
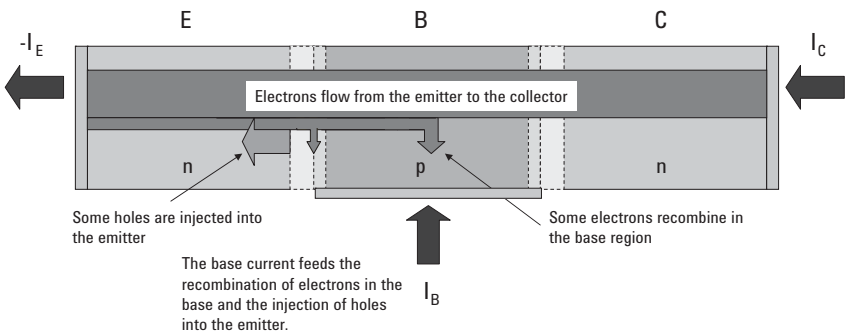
Figure 7.18. The basic operation of a bipolar transistor (NPN type).

Of course, electron flow from the emitter to the collector is equivalent to a positive current flow from the collector to the emitter.

Similar to the case of a pn diode, there is an exponential relationship between current flow and the base to collector (V_{BC}) and base to emitter (V_{BE}) voltages. Assuming that V_{BC} is zero or negative, the collector current is approximately described by the equation shown below.

$$I_C \approx I_o \left(e^{\left(qV_{BE}/kT \right)} \right)$$

(Equation 7.19)

As can be seen from this equation, the base to emitter voltage does not need to change much in order to have large changes in the collector current. Therefore, once the NPN transistor has a sufficiently large V_{BE} to become actively biased the magnitude of the collector current is limited by the amount of current entering into the base of the transistor. As the previous diagram showed, this base current is necessary to feed both the recombination of electrons in the base region as well as the injection of holes into the emitter. There is a very simple linear relationship between the base and collector currents given by the equation shown below.

$$I_c = \beta \cdot I_B$$

(Equation 7.20)

Basic bipolar transistor characterization

Probably the single most important bipolar transistor parameter is its current gain or beta. This parameter is also often written in terms of h-parameters as "h_{FE}", which stands for forward gain through the emitter. A Gummel plot is usually used to determine the value of beta for various values of collector current. One method to perform a Gummel plot on an NPN transistor is to set the collector and base to zero volts and to sweep the emitter negatively. The values of base current and collector current are then measured and plotted on a log scale (since by Equations 7.19 and 7.20 the natural logarithm of the collector and base currents should be straight lines).
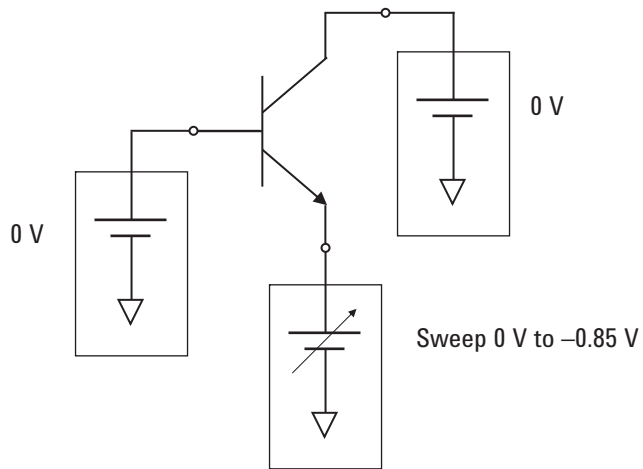
Figure 7.19. Typical setup for a bipolar transistor Gummel plot.

The results of an actual Gummel plot are shown below.
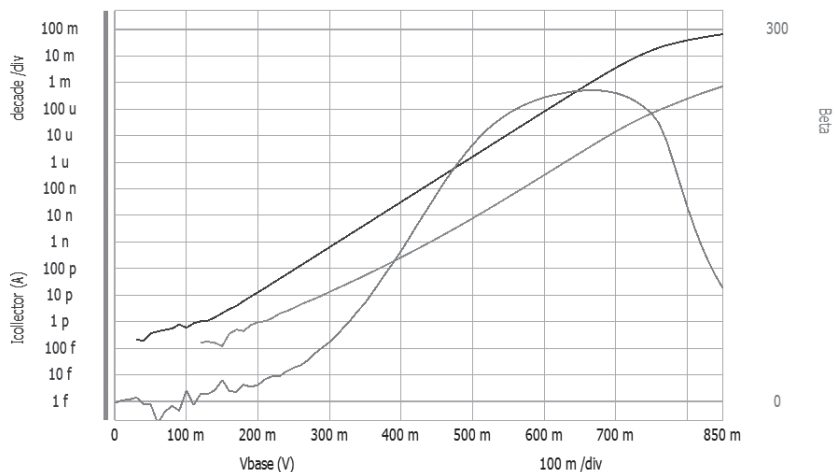
*Figure 7.20. A Gummel plot made on an NPN bipolar transistor using the B1500A.
Ic, Ib and the calculated beta are all shown on the same graph.*

As this plot shows, the value of the transistor beta is actually dependent upon
the large-signal value of the collector current, although it is approximately
constant over a fairly wide range of collector current values.

Another important bipolar transistor parameter is the emitter resistance (Re).
The flyback method is the most common means used to determine this param-
eter. In this measurement the SMU connected to the collector terminal is placed
in current force mode with a very small force value (almost zero amps), making
it essentially a high-impedance voltmeter. The SMU connected to the base
terminal is also set to current force mode and the collector-emitter voltage (Vce)
is measured and plotted versus the base current (Ib).
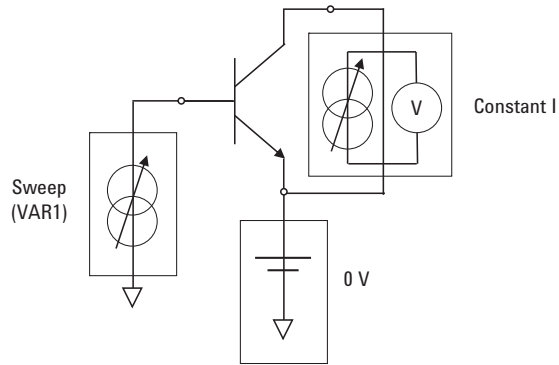
Figure 7.21. The configuration for measuring Re on a bipolar transistor using the flyback method.

The slope of the line in the flyback region can be determined by creating a regression line at predefined points using the built-in EasyEXPERT analysis functions. The slope of this line is in units of amps/volt; this is the inverse of resistance (which has units of volts/amp). Therefore, the inverse of the regression line slope gives us the value of emitter resistance (Re). A plot of Ib versus Vce showing the emitter resistance is shown below.
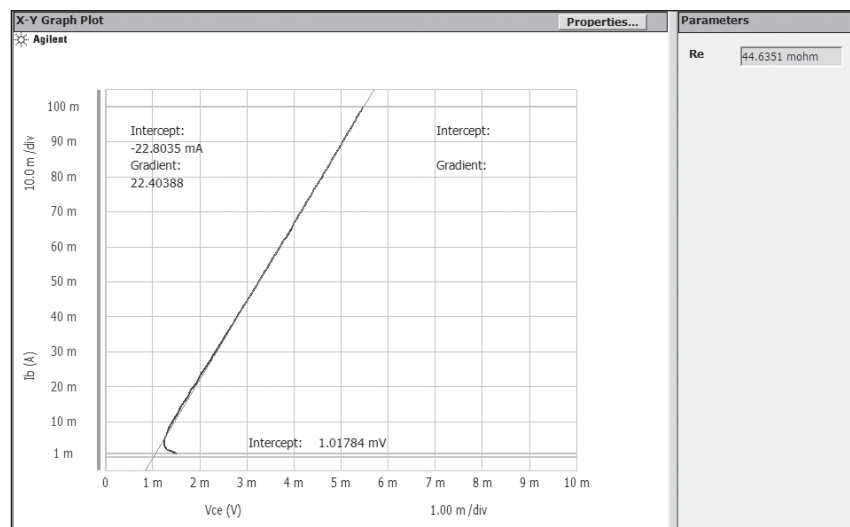
Figure 7.22. Extracting the emitter resistance (Re) using the flyback method.

As in the case of power MOSFETs, the breakdown voltages of power bipolar devices are also very important parameters. The junction breakdown voltages of power BJTs are defined as follows.

BVcbo – The collector-base breakdown voltage with the emitter open

BVceo – The collector-emitter breakdown voltage with the base open

BVebo – The emitter-base breakdown voltage with the collector open

The B1505A's HVSMU can measure breakdown voltages for most power BJTs, and it can also measure the leakage current with picoamp resolution. EasyEXPERT's graphical capabilities make it easy to superimpose these different breakdown measurements onto a single plot and to add comments to the graph for use in reports and presentation. An example of this is shown below.
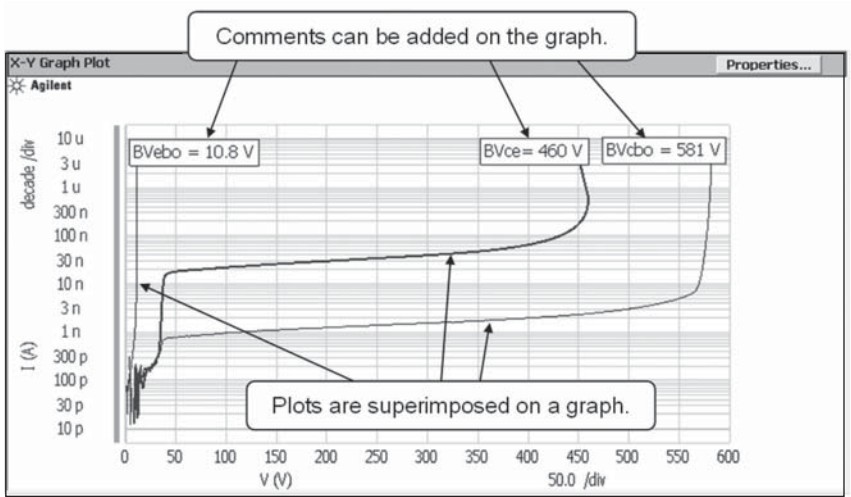
Figure 7.23. High power bipolar transistor breakdown voltages displayed on a single graph.

If you would like more detailed information on power device characterization, please refer to Agilent Technologies' Application Note B1505-1, "Accurate and Efficient Characterization of Power Devices at 3000 V/20 A" (5990-3749EN).

Chapter 8: Capacitance Measurement Fundamentals

"Furious activity is no substitute for understanding." — H. H. Williams

Introduction

Capacitance measurement is one area of parametric test where many easily preventable measurement mistakes are often made. The reason for this is not lack of intelligence on the part of the user, but rather a lack of fundamental training on capacitance measurement theory and how to make good capacitance measurements. Unfortunately, this type of information is typically not taught in universities and can usually only be learned either through (rather painful) experience or by reading about it in a publication such as this.

Why do engineers make so many mistakes when measuring capacitance (especially on-wafer)? The most common reasons are:

1. Capacitance measurement requires compensation to remove parasitic inductance and capacitance from the measurement cables and fixturing, and many times this is done improperly (or not at all).
2. An induced current flows through the outer shield of the BNC connectors (if used) on a capacitance meter, and this current is necessary to balance the measurement current of the capacitance meter. If the outer shield is grounded, then the induced current flow is shorted to ground and the bridge may not be able to balance. Many users are unaware of this issue.
3. Measuring capacitance on a semiconductor wafer on a wafer chuck is very different from measuring a discrete device. The effects of the wafer prober chuck on the measurement cannot be ignored.
4. For higher measurement frequencies (> 5 MHz), structure (layout) design has a major impact on the success or failure of the measurements.

As we progress through this (rather lengthy) chapter we will cover all of these issues in detail

MOSFET capacitance measurement

Review of MOSFET capacitance behavior

Before delving into capacitance measurement theory, it is useful to first review MOSFET device operation to remind ourselves as to why we are making these measurements in the first place. While the capacitance measurement techniques discussed in this chapter are general and can be used on a variety of different device types, the dominance of MOSFETs in modern electronics elevates them enough in importance to justify devoting a section to reviewing their operating characteristics.

MOSFETs are voltage-dependent capacitors. The MOSFET gate-to-substrate capacitance depends upon the applied dc voltage (which we measure using an ac voltage of much smaller magnitude that rests on top of the dc voltage). The following set of diagrams shows the behavior of an NMOS transistor as the voltage applied to the gate is varied from negative to positive.
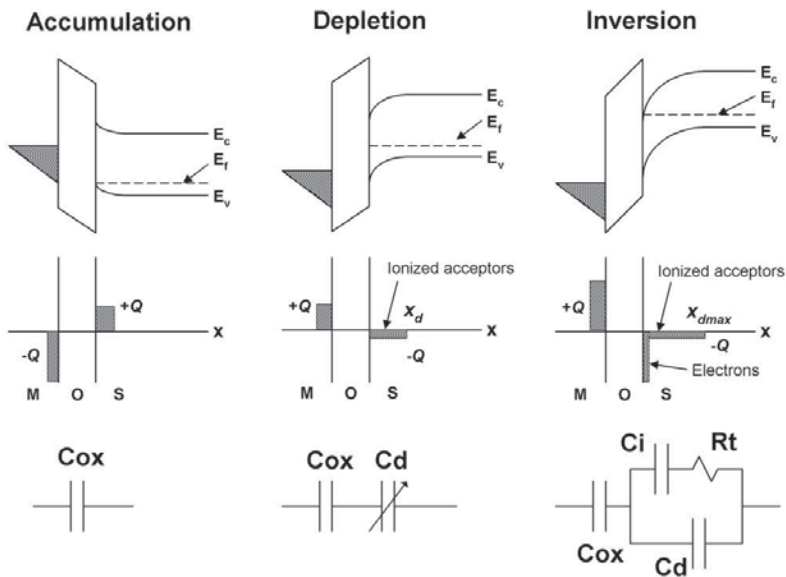
Figure 8.1. Capacitive behavior of an NMOS transistor to changes in the voltage applied to the gate.

If the silicon is held at ground and a negative voltage is applied to the gate, the MOS capacitor will begin to store positive charge at the silicon surface. The surface has a greater density of holes than Na (the acceptor density), and this condition is known as surface accumulation. In this condition the mobile charge on both sides of the oxide can respond rapidly to changes in applied voltage, and the device looks just like a parallel plate capacitor of thickness *tox*. Since it is a pure gate oxide capacitance, we denote its value as *Cox*.

If a positive gate voltage is applied to the gate relative to the silicon, the built-in positive voltage between the gate and silicon is increased. The silicon surface becomes further depleted of carriers as more acceptors become exposed at the surface, resulting in the condition known as surface depletion. In this condition electrostatic analysis shows that the total MOS capacitance consists of the series combination of Cox and the capacitance across the surface depletion region, Cd. Note that Cd depends upon the applied voltage.

If the positive gate voltage is further sufficiently increased, then the energy bands bend away considerably from their levels in the bulk of the silicon. The depletion region reaches a maximum width, $xdmax$, and all of the electron acceptors within this region are fully ionized. In the surface region generation of carriers exceeds recombination, and the generated electrons are swept by the electric field into the oxide-silicon interface where they remain due to the energy barrier between the conduction bands of the silicon and the oxide. Thus, the total charge in the silicon consists of the sum of these two charges. Electrostatic analysis again shows that the total MOS capacitance can be modeled as the oxide capacitance in series with the parallel combination of the depletion capacitance and the series combination of surface charge capacitance, Ci, and the depletion resistance, Rt.

Sample MOSFET parameter calculation

The reason that measurement of the gate to substrate capacitance of a MOSFET device is so important is that it is the only way to calculate many important device parameters such as substrate impurity concentration (Nsub) and flat band voltage (Vfb). It is illustrative to go through a sample parameter calculation to show how to extract these parameters from a capacitance versus voltage (CV) curve. All of the following calculations are based upon the (CV) curve shown below:
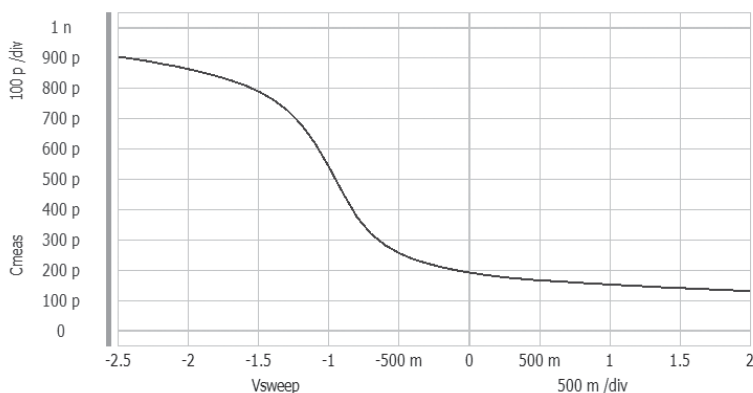

Figure 8.2. Capacitance versus voltage (CV) plot.

The gate oxide thickness of a MOSFET capacitor (t_{ox}) can be calculated from the standard equation for a parallel plate capacitor:

$$t_{ox} = \frac{A \cdot 10^8 \cdot \varepsilon_o \cdot \varepsilon_d}{Cox} \qquad \text{[Angstroms]} \qquad \textit{(Equation 8.1)}$$

Where: A is the capacitor gate area [cm^2]
　　　　ε_o is the free space permittivity (8.854×10^{-14} F/cm)
　　　　ε_d is the dielectric constant of SiO_2 (3.9)
　　　　Cox is the measured capacitance in heavy accumulation (Vg bias = Vdd) [F]

For Cox = 9.040×10^{-10} F and A = 0.001 cm^2, we have that:

　　　　t_{ox} = 37 Angstroms

 The next two parameters that we must calculate are the substrate impurity concentration and the Fermi potential. These are given by the following two equations:

$$Nsub = \frac{4 \cdot |\varphi_f|}{q \cdot \varepsilon_o \cdot \varepsilon_{Si}} \left(\frac{Cs_{min}}{A} \right)^2 \qquad \text{[1/cm}^3\text{]} \qquad \textit{(Equation 8.2)}$$

$$\varphi_f = \pm \frac{kT}{q} \cdot \ln\left(\frac{Nsub}{n_i} \right) \qquad \text{[Volts]} \qquad \textit{(Equation 8.3)}$$

Where: Nsub is the impurity concentration of the substrate
　　　　n_i is the intrinsic carrier concentration [$1/cm^3$]
　　　　Φ_f is the Fermi potential [Volts]
　　　　Cs_{min} is the minimum depletion layer capacitance [Farads]
　　　　ε_{Si} is the dielectric constant of Si (11.7)
　　　　q is the magnitude of the electron charge (1.602×10^{-19} Coulomb)
　　　　k is Boltzmann's constant (1.38×10^{-23} J/K)
　　　　T is the absolute temperature [deg K]

Note: The sign of the Fermi potential is determined by the doping in the channel. It is plus (+) for p-doped channels (NMOS transistors) and minus (-) for n-doped channels (PMOS) transistors.

These equations do not have a closed form solution; they must be solved iteratively. Using the value of Cs_{min} = 2.01×10^{-10} F in the above example, through repeated iterations on a computer we arrive at the values for Nsub and Φ_f shown below:

　　　　Nsub = 1.812×10^{17} cm^{-3}
　　　　Φ_f = 0.4315 V

Where   A = 0.001 cm^2 and T = 300 ^{o}K

The Debye length can be calculated from the following equation:

$$\lambda = \sqrt{\frac{2k \cdot T \cdot \varepsilon_o \cdot \varepsilon_{Si}}{q^2 \cdot Nsub}} \qquad \textit{(Equation 8.4)}$$

The value of the depletion layer capacitance under flat band conditions (Csfb) is given by:

$$Csfb = \frac{\sqrt{2} \cdot A \cdot \varepsilon_o \cdot \varepsilon_{Si}}{\lambda}$$

(Equation 8.5)

Plugging the value of Nsub obtained above into these two equations yields:

$Csfb = 1.0789 \times 10^{-9}$ F

Since the flat band capacitance (Cfb) is the series combination of Cox and Csfb, we know that:

$$Cfb = \frac{Cox \cdot Csfb}{Cox + Csfb}$$

(Equation 8.6)

We earlier calculated the value of Cox, so combining this with the value of Csfb just calculated we arrive at a value for Cfb of:

$Cfb = 4.9185 \times 10^{-10}$ F

To get the value of the flat band voltage (Vfb) we need to take this value of Cfb and perform a linear interpolation on our capacitance plot. The two points on either side of the value of Cfb that we have just calculated are:

V = −0.9 V, C = 454 pF
V = −1.0 V, C = 540 pF

Thus, we can determine the value of the flat band voltage to be:

Vfb = −0.9440 V

The next useful parameter to calculate is the surface charge density (Qss). Typically, this is divided by the electron charge (q) and expressed as Qss/q:

$$\frac{Qss}{q} = \frac{Cox}{q \cdot A} \cdot \left| \Phi_{ms} - Vfb \right| \qquad [1/cm^3] \qquad \text{(Equation 8.7)}$$

In this example,

$\Phi_f = 0.4315$ V
$\Phi_{ms} = -0.6 - \Phi_f = -1.032$ V

where Φ_{ms} is the difference in work functions of the semiconductor (Si) and the gate poly (poly-Si). Therefore, plugging in numbers we get that:

$Qss/q = 4.9375 \times 10^{11}$ cm^{-3}

We can also calculate the fixed charge in the depletion layer (Q_b) and the Vth using the following equations:

$$Q_b = \pm q \cdot Nsub \cdot \frac{A \cdot \varepsilon_o \cdot \varepsilon_{Si}}{Cs_{min}} \qquad \text{[Coulomb/cm}^2\text{]} \qquad \textit{(Equation 8.8)}$$

$$Vth = Vfb + \left(2 \cdot \Phi_f - \frac{A \cdot Q_b}{Cox} \right) \qquad \text{[Volts]} \qquad \textit{(Equation 8.9)}$$

Note: The sign of Q_b is determined by the doping in the channel. It is plus (+) for n-doped channels (PMOS transistors) and minus (-) for p-doped channels (NMOS) transistors. Plugging in values previously obtained, we have that:

$Q_b = -2.2785 \times 10^{-7}$ C/cm^2

Vth = 0.1711 V

It should be obvious that it is best to create a program to automate this parameter extraction process.
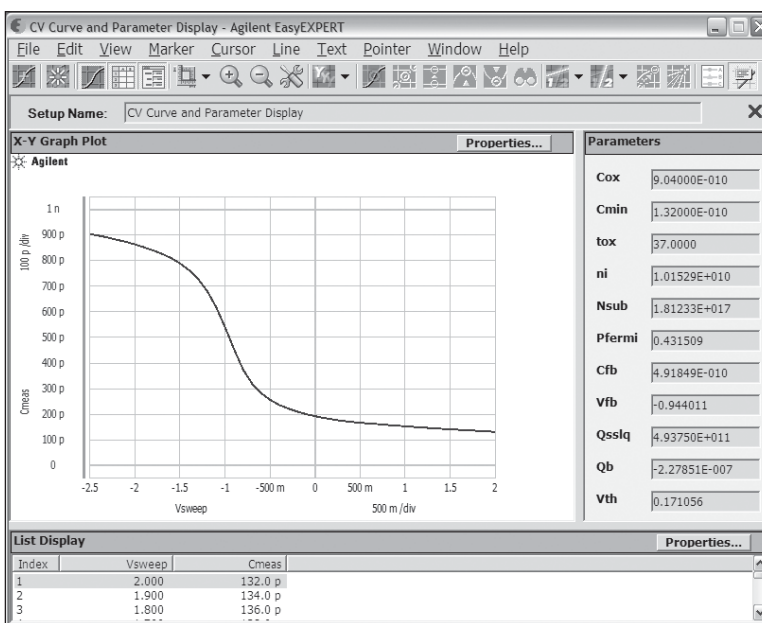

Figure 8.3. Automatic calculation of MOSFET capacitor parameters using Agilent EasyEXPERT software.

Note: The application test to perform this calculation can be downloaded from the B1500A product page on the Agilent Technologies web site.

Quasi-static capacitance measurement

It is important to understand that a quasi-static capacitance versus voltage (QSCV) MOSFET device response to a stimulus is not a measurement technique. While it is true that certain measurement techniques are more likely to result in a QSCV response from a MOSFET than are others, whether or not a stimulus results in a QSCV response depends upon a variety of process, device, and layout related factors in addition to the type of stimulus applied to the MOSFET. Therefore, as we discuss the step voltage QSCV measurement technique in the following section it is worthwhile to keep in mind that in certain cases QSCV device responses can also be observed using more traditional oscillator-based capacitance measurement techniques. Also, simply using the step voltage QSCV measurement technique does not guarantee a QSCV device response.

Obtaining a QSCV response
Before proceeding with an explanation of the voltage step QSCV measurement technique, it is useful to review the difference between a QSCV and a high-frequency capacitance versus voltage (HFCV) response.
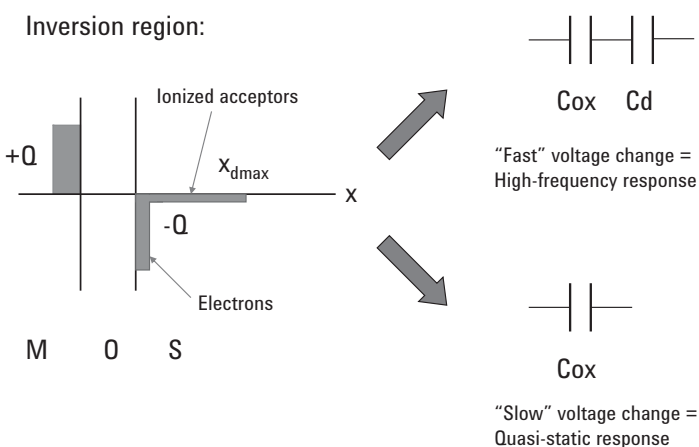
Figure 8.4. Quasi-static versus high-frequency capacitance versus voltage response for an NMOS transistor.

When the semiconductor channel is fully inverted (as shown above), the only sources for the free electrons are generation and recombination. If the applied voltage is changed slowly enough, then these mechanisms can supply mobile electrons and the capacitor looks just like a parallel capacitor with a capacitance value approximately equal to that of Cox. This results in a classic QSCV response. However, if the voltage changes more rapidly than generation and recombination can respond, then the ionized acceptor region has to modulate to maintain charge balance. In this case the total capacitance has to be the series combination of the oxide capacitance and the depletion region capacitance (Cox + Cd). This acts to reduce the overall value of the capacitance below that of Cox alone, resulting in what is usually known as an HFCV response.

While many textbooks have attempted to supply equations to define what exactly is a sufficiently "slow" enough voltage change to produce a QSCV response, measurement experience has shown that obtaining a QSCV response from a MOSFET is about equal parts art and science. In practice, it is usually best to start a sweep in the inversion region, and then proceed to sweep through the depletion and accumulation regions. In addition, it may be necessary to supply a DC bias to the capacitor for several seconds while it is inverted prior to the start of the sweep. Sometimes it is also useful to shine a light onto the device to stimulate the generation of minority carriers before the start of the sweep. However, in this case it is essential to turn the light off before the start of the sweep, since the light disturbs the thermal equilibrium of the capacitor and distorts the measurement results.

The voltage ramp QSCV measurement technique
The classical method to perform a QSCV measurement involved using a voltage ramp created with an instrument such as the Agilent 4140B. The 4140B could generate a voltage ramp with a constant slope, which allowed the capacitance to be determined using the basic relationship between current, voltage and capacitance.

$$i = C\frac{dV}{dt} \quad \Rightarrow \quad C = \frac{i}{dV/dt} \qquad \qquad \text{(Equation 8.10)}$$

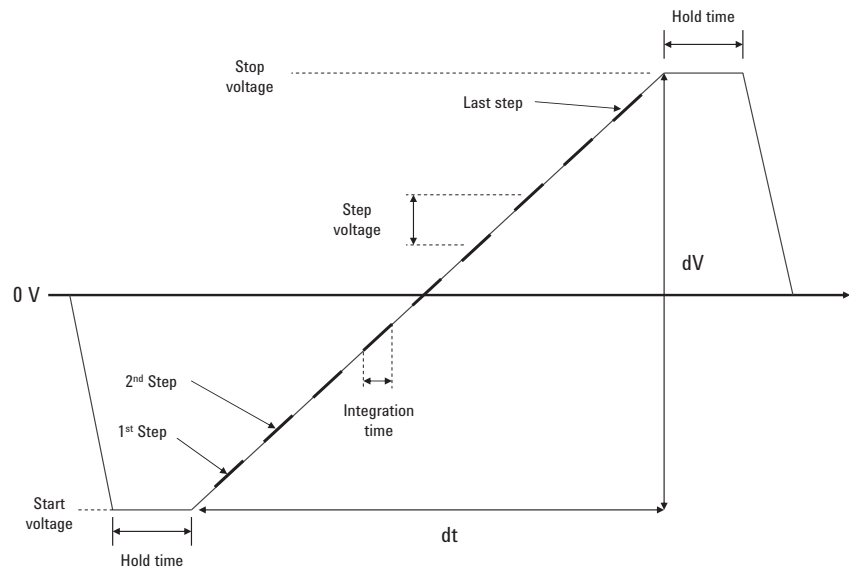A typical QSCV measurement using this technique is shown below.

Figure 8.5. The classical voltage ramp QSCV measurement technique using an Agilent 4140B.

Since the 4140B is an obsolete instrument that is no longer sold or supported, this QSCV measurement technique has been replaced by the voltage step QSCV measurement technique.

The voltage step QSCV measurement technique

The step voltage QSCV measurement technique does not use a voltage ramp or oscillator to measure CV. Instead, it uses SMUs and performs a measurement that it very similar to a standard IV sweep. The key difference is that, at each point on the sweep, a (user-specified) ΔV is applied. The corresponding charge ΔQ that results can be calculated by measuring the current induced by the ΔV and numerically integrating the area under the current versus time curve. This then allows the capacitance at each point along the sweep to be calculated by dividing ΔQ by ΔV. The figure shown below illustrates this technique:

$$C = \frac{\Delta Q}{\Delta V} \qquad \Delta Q = \int i\, dt \sim \sum i\, \Delta t$$
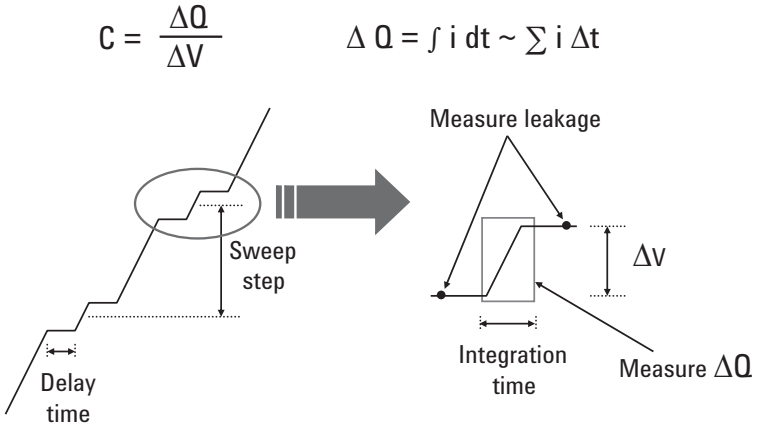
Figure 8.6. The step voltage QSCV measurement technique.

Obviously, the value for ΔV chosen has to be less than or equal to the size of the sweep step.

In addition to measuring capacitance at each point in the sweep, the step voltage QSCV algorithm also supports the ability to remove leakage currents caused by electron tunneling through the gate dielectric. This is done by measuring the leakage current before and after applying the ΔV, and then subtracting out the leakage current before calculating the capacitance. Unfortunately, this technique is ultimately limited by the current measurement range in which the capacitance sweep is being made. If the leakage current exceeds the current measurement range, then an error will occur. For extremely leaky gate dielectrics (with oxide thicknesses of approximately 25 angstroms or less), the step voltage QSCV measurement technique cannot be used. In this case, the only solution is to make extremely high frequency measurements using a conventional capacitance meter technique. This technique will be discussed later in this chapter.

A valid question to ask is: how well does the voltage ramp QSCV technique correlate with the voltage ramp QSCV technique? The answer is quite well as long as you pick the measurement parameters correctly.

- The same start, stop and step values can be used
- The voltage step can be determined by multiplying the 4140B ramp rate by the 4140B integration time
- The integration time should be set to the same value as the 4140B integration time
- The leakage compensation feature is disabled

The following graph shows actual measurement data taken on a device using the two different QSCV measurement techniques.
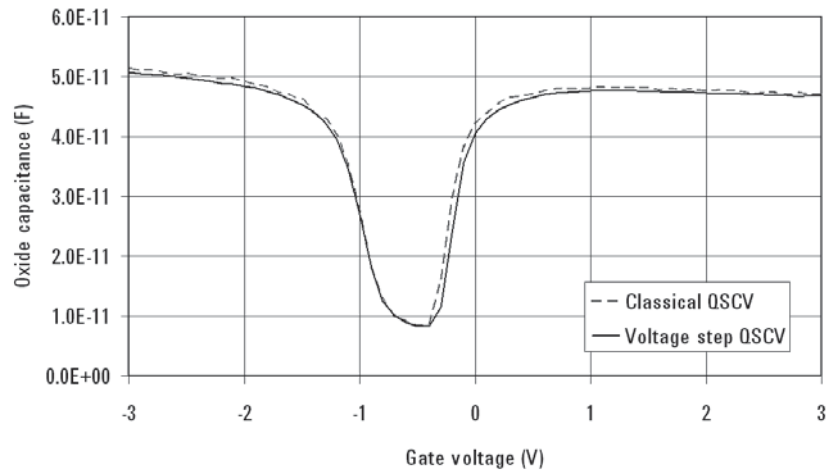
Figure 8.7. Correlation between the classical (voltage ramp) and step voltage QSCV measurement techniques.

As this graph shows, the correlation between the two techniques is quite good. If you have trouble correlating the two techniques then please check to see if one or more of the following conditions applies:

- The 4140B ramp rate (dV/dt) is too steep
- The *delay time* or *integration time* is too short
- The leakage current compensation feature is turned on

QSCV offset compensation

The main parasitic affecting the step voltage QSCV measurement technique is the offset capacitance between the triaxial cables. Therefore, it is sufficient for the QSCV measurement technique to support simple offset compensation. With the probe tips open, the instrument performs a capacitance measurement to determine the offset capacitance. This value is stored and automatically subtracted from subsequent QSCV measurements.
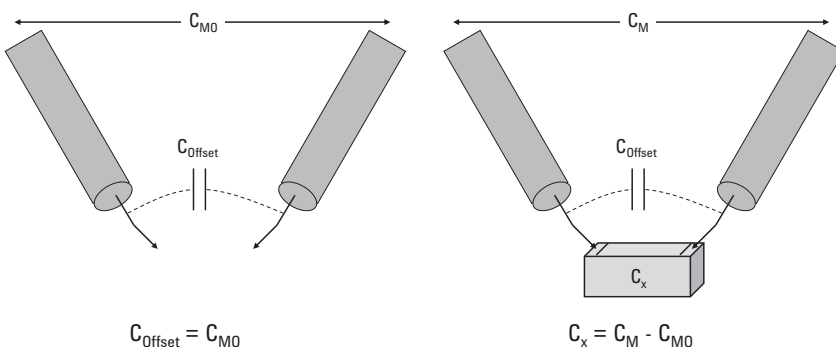
Figure 8.8. The QSCV measurement technique can automatically subtract the innate offset capacitance of the triaxial cables.

This compensation works well for the QSCV measurement technique and gives excellent correlation with oscillator-based capacitance measurement techniques.
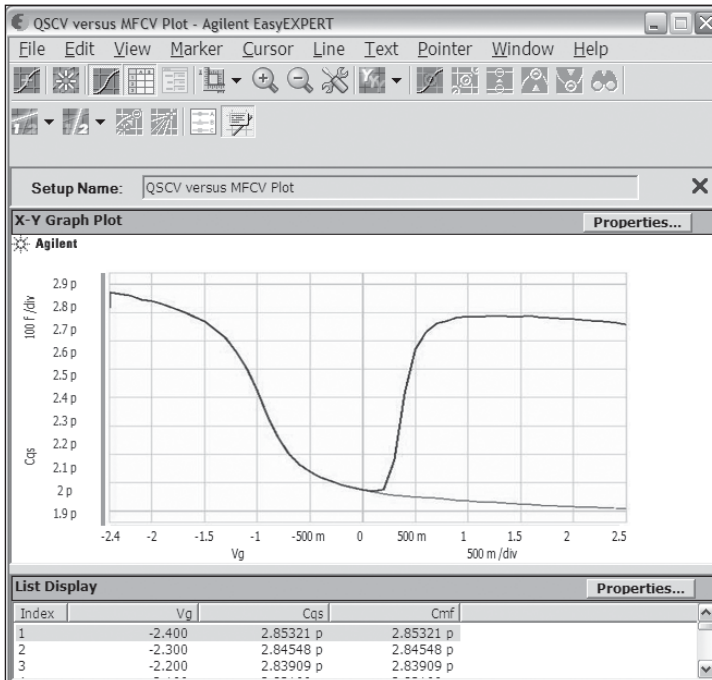
Figure 8.9. This superposition of quasi-static CV and high-frequency CV plots shows that you can get excellent measurement correlation between the two techniques.

Low frequency (< 5 MHz) capacitance measurement

For the purposes of this discussion we will consider low frequency measure-ments to be anything that uses an oscillator that has a frequency from a few Hertz up to 5 MHz.

Capacitance meter operation

For many people, the operation of a capacitance meter is a great mystery. However, the basic principal of the auto balancing bridge circuit are not difficult to understand. Please refer to the circuit schematic shown below.

$$V_2 = I_2 \times R_2$$

$$Z = \frac{V_1}{I_2} = \frac{V_1 R_2}{V_2}$$
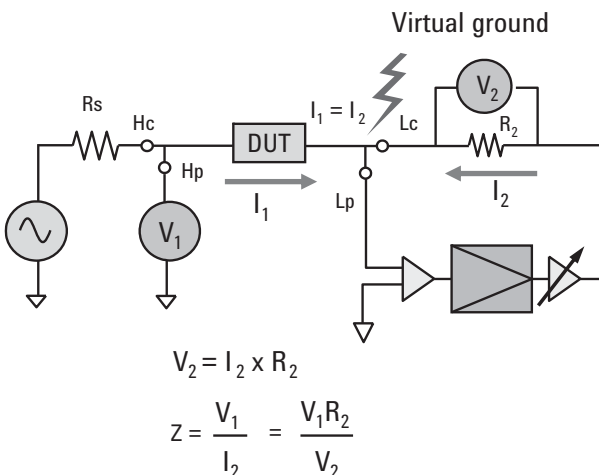
Figure 8.10. The auto balancing bridge capacitance measurement method.

The auto balancing bridge can be conceptualized as an operational amplifier (op amp) circuit. Ohm's law applies: $V = I \times R$. The device is stimulated by an AC signal, with the actual voltage applied to the device being monitored at the H (high) terminal. The L (low) terminal is driven to 0 volts by the virtual ground of the op amp. The current, I_2, through the range resistor is equal to the current through the DUT. Therefore, the output voltage is proportional to the current through the device. Voltages and current are automatically balanced, thus giving rise to its name. To cover a wide frequency range, a null-detector and a modulator are used instead of an amplifier in practical circuits. The four inputs to the capacitance meter are defined below:

    Hc: Signal source
    Hp: Potential meter
    Lc: Current meter
    Lp: Potential meter to Lock the phase of measurement signal

The 4TP measurement method

The most common measurement technique used with the auto balancing bridge capacitance meter is the 4 terminal pair (4TP) measurement method. In this procedure the Hc and Hp terminals and the Lc and Lp terminals are shorted together as shown below.
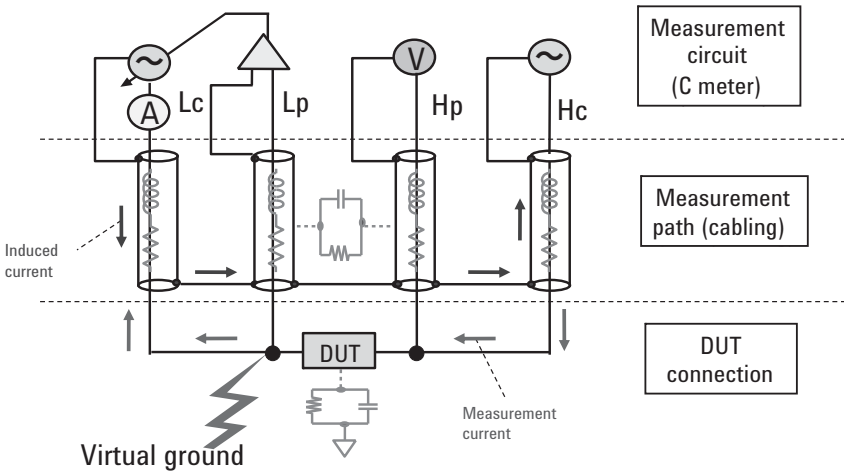
Figure 8.11. The 4 terminal pair (4TP) measurement method (including parasitic elements).

The Hp and Hc terminals are commonly referred to as the CMH (capacitance meter high) connection and the Lp and Lc terminals are commonly referred to as the CML (capacitance meter low) connection. There is residual inductance and resistance in the measurement path (cabling) and stray capacitance both between the cables and between the DUT and ground (which are also shown in the previous figure). The user must perform compensation through the measurement path (cables) in order to eliminate the effects of these parasitic elements or the accuracy of the measurement will be seriously degraded.

The CML terminal of the 4TP scheme is a virtual ground and it is very important NOT to tie this point to earth ground. If the CML terminal contacts earth ground then the auto balancing bridge circuit will not be able to balance and you will get measurement errors. The outer conductor (shield) of the capacitance cables has the same potential as the virtual ground, and it needs to be floated as well. In addition, the four outer shields should be connected together just before the probe tips with cables that are short in length. This creates a return path for the current that is induced in the shield, which stabilizes the series inductance of the cables. Unless the outer shields are shorted and the series inductance stabilized, the mere act of changing the cable separation will cause large variations in cable inductance that immediately invalidate any cable compensation that has been performed.

The importance of proper compensation

For capacitance meters, the "calibration plane" defines the point (normally the BNC outputs of the capacitance meter) at which the user can obtain the specified measurement accuracy of the capacitance meter. Of course, for parametric test we need to use cabling to test our DUT. In the case of the 4TP measurement method this introduces a variety of sources of error due to residual parasitic elements in the test fixturing (cables).
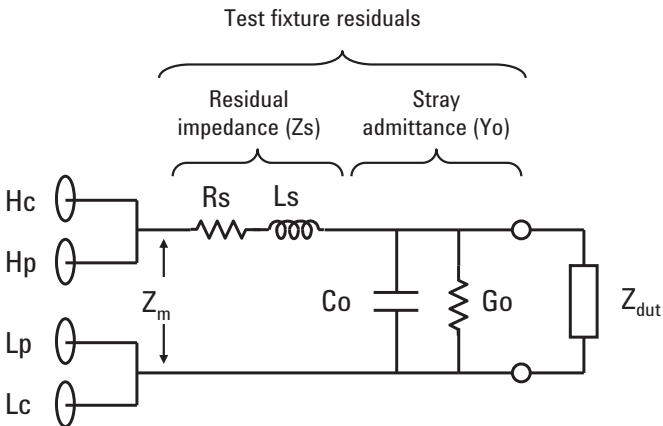
Figure 8.12. The test fixture residuals that you must account for in order to make accurate capacitance measurements through a test fixture (cable).

To remove the effects of these test fixture residuals, we perform two measurements: one with the test terminals open and one with the test terminals shorted as shown below.

$$Yo = Go + \omega jCo$$

$$Rs + j\omega Ls \ll \frac{1}{Go + \omega jCo}$$

$$Zs = Rs + j\omega Ls$$
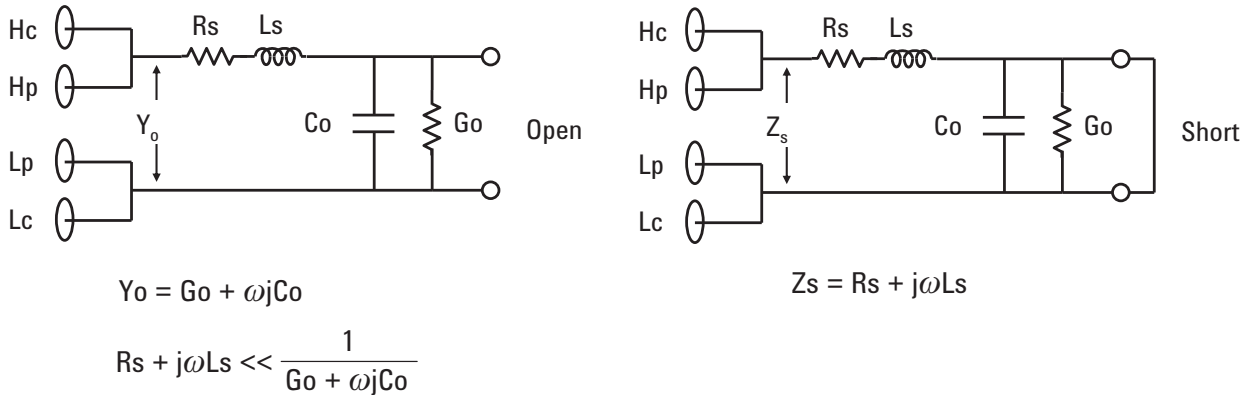
Figure 8.13. Using OPEN/SHORT compensation to determine test fixture residuals.

Using these two values, we can calculate the true value of Zdut from the measured value (Zm) using the equation shown below.

$$Z_{dut} = \frac{Z_m - Z_s}{1 - (Z_m - Z_s)Y_o}$$
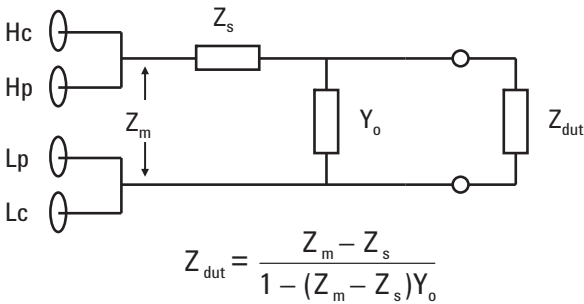
Figure 8.14. The true value of Zdut can be calculated using the OPEN and SHORT values.

In addition to OPEN/SHORT cable compensation, phase compensation should always be performed before making measurements for the first time (when this feature is supported). Phase compensation improves the bridge balance stability at high frequencies and minimizes phase shift effects due to variations in frequency and cable length. Phase compensation should be performed before performing the OPEN/SHORT cable compensation, and it generally does not need to be done again unless you change your measurement configuration. The Lc and Lp terminals need to be shorted together during the phase compensation. The procedure to follow to perform phase compensation varies by instrument, so you need to refer to your instrument's manual to understand exactly how to do this (if it is supported).
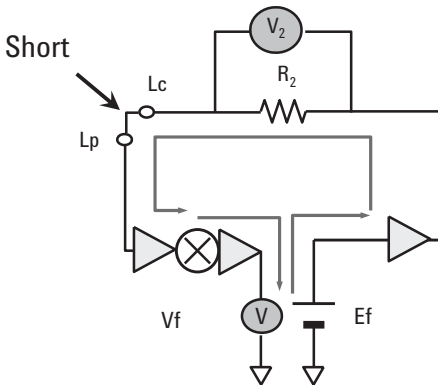
Figure 8.15. Phase compensation calculates the characteristic of the null loop circuit, which improves repeatability and phase accuracy.

For measurements at or below 5 MHz load compensation is not always necessary; its effect on measurement accuracy at these frequencies depends on the exact fixturing that you are using. However, for measurements above 5 MHz you should always perform OPEN/SHORT/LOAD compensation or significant measurement error will occur. This requires the on-wafer measurement of some sort of calibrated resistor, and it is normally done with RF wafer probes. We will return to this topic when we discuss high-frequency capacitance measurement.

On-wafer capacitance measurement

There are some important measurement issues to take into account when making on-wafer capacitance measurements. The most important of these is how the Hc and Hp and Lc and Lp terminals contact the wafer. Many people know from experience that they get much better measurement results by contacting the Lc and Lp terminals to the wafer (DUT) and contacting the Hc and Hp terminals to the wafer chuck when using the 4TP cabling method.
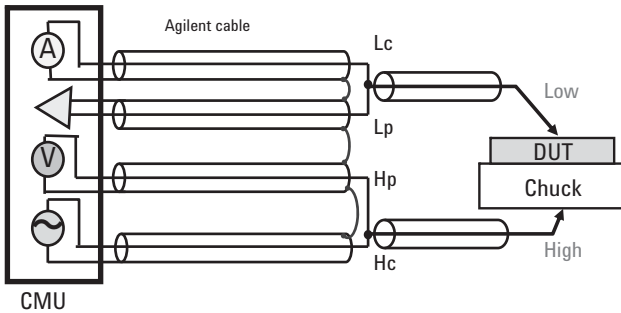
Figure 8.16. The proper way to make capacitance measurements on-wafer using the 4TP method.

However, most people do not know why this provides better measurement results. The reason that you obtain better measurement results using this technique can be understood by remembering that on a wafer prober you have to be concerned with chuck-to-ground capacitance. When making measurements at frequency, the chuck capacitance acts as a sneak path to ground. Consider the two different cases shown below.

Figure 8.17. The reason why measuring from the Lc and Lp terminals eliminates measurement error.
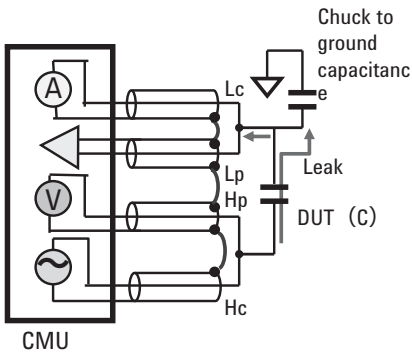
The actual capacitance measurement is performed by the Lc and Lp terminals, so if these terminals directly contact the DUT then all of the current flowing through the DUT is measured by the capacitance meter and the current flowing to ground through the wafer chuck is irrelevant. Of course when measuring in this fashion you need to reverse the applied DC bias, but this is a rather trivial consideration that can easily be automated.

160

Another reason to connect the Hp and Hc terminals to the wafer chuck is noise. The wafer chuck acts like a large antenna, collecting random electromagnetic noise. If the Lp and Lc terminals are connected to the wafer chuck, then this noise will flow directly into the capacitance meter. In extreme cases the capacitance meter may give an error message and return an unbalanced status error (UNBAL). However, if the Hp and Hc terminals are connected to the wafer chuck then noise from the wafer chuck is attenuated when it passes through the DUT.
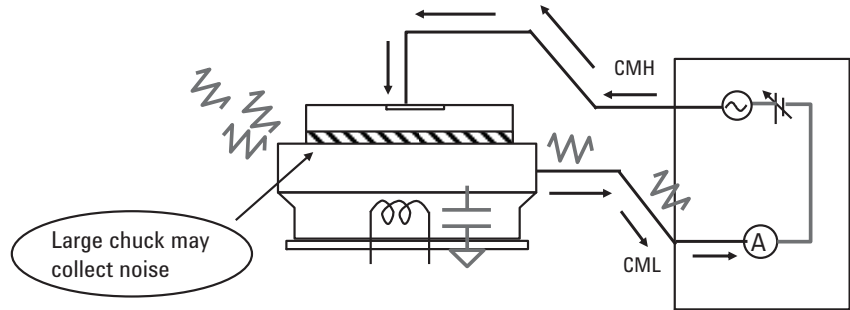
Figure 8.18. A large chuck can inject noise directly into the ammeter if the Hc and Hp terminals are connected to the DUT.

The following graphs compare the measurement results obtained for different frequencies and integration times when connecting the CMH terminals to both the gate and to the bulk.
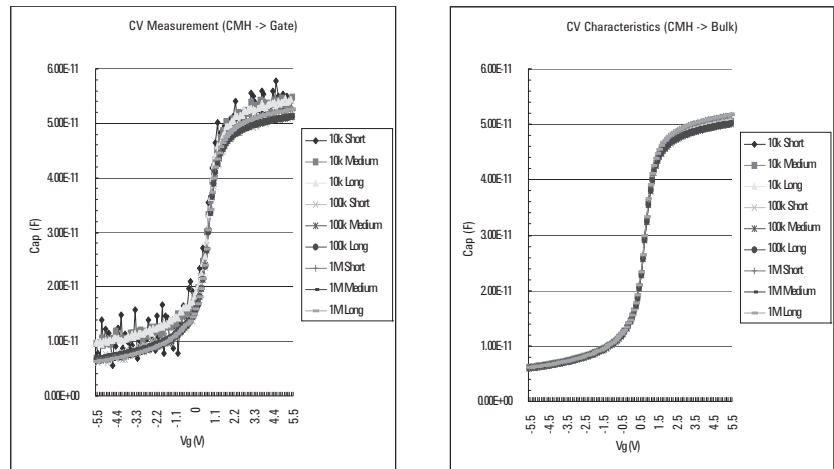
CMH connected to gate | CMH connected to bulk (wafer chuck)

Figure 8.19. Measurement data comparing the results obtained connecting the CMH terminal to the gate versus connecting the CMH terminal to the wafer chuck (bulk).

161

Almost all of the B1500A's EasyEXPERT application tests for capacitance measurement are designed to have the CML terminals contact the DUT and the CMH terminals contact the wafer chuck. The nice feature of these application tests is that they automatically take care of inverting all of the signs on the measurement inputs so that you do not have to think about this.
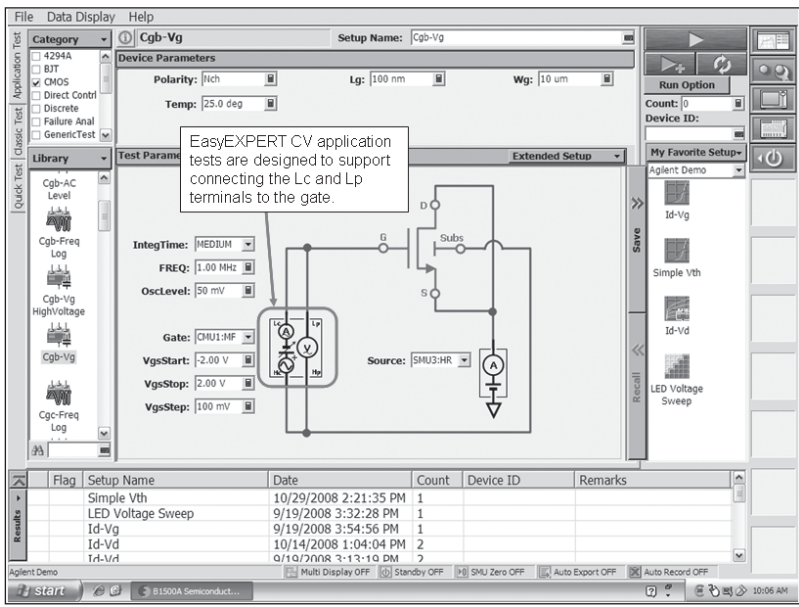
Figure 8.20. The terminal connections on a B1500A EasyEXPERT CV application test.

High frequency (> 5 MHz) capacitance measurement

The choice of 5 MHz as a dividing point between "low-frequency" and "high-frequency" measurement is not arbitrary. It is actually a natural dividing line between these two measurement regimes. Above 5 MHz, new measurement techniques and schemes need to be employed in order to obtain satisfactory measurement results. For example, above 5 MHz it becomes increasingly difficult or impossible to obtain good measurement results with standard DC probes; instead RF probes that require a ground-signal (GS) or ground-signal-ground (GSG) pad layout become mandatory. Also, in addition to having a pad structure that supports these types of RF probes, the actual layout of the device takes on critical importance. In fact, the most common reason for failure to achieve satisfactory capacitance measurement results above 5 MHz is not the measurement equipment or cabling, but improper test structure design. Therefore, it is important that you read and understand this section before you undertake the design of your capacitance test structures.

Thin gate dielectric fundamentals
Many years ago as transistor gate dielectrics became extremely thin (~25 angstroms or less) the issue of gate leakage current due to electron tunneling moved to the forefront of parametric test. While this is not a difficult issue to deal with for standard current versus voltage (IV) parametric measurement, it presents some challenges for capacitance versus voltage measurement (CV). There are three essential requirements for the successful measurement of thin gate (leaky) dielectric MOSFET capacitors:

1. A high measurement frequency (> 5 MHz)
2. A device model suitable for thin-gate dielectrics
3. The elimination of wafer chuck effects

In reality achieving all three of these simultaneously is not easy, making on-wafer measurement of thin gate dielectrics one of the most challenging areas of capacitance measurement. We will examine each of these requirements in-turn.

To start with please consider the basic two-element capacitance model shown below.

If Ir is large → Rp is small
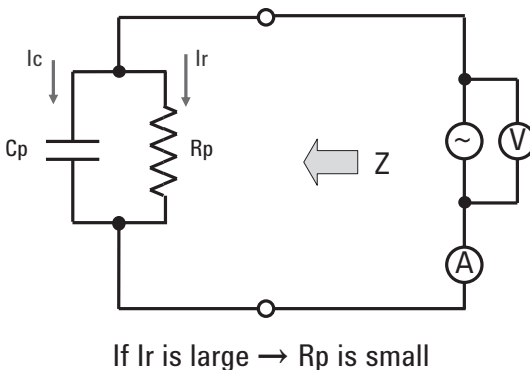
Figure 8.21. The two-element capacitance model.

The equivalent impedance of this circuit is given by:

$$Z = \frac{\left| \dfrac{1}{C_p j\omega} \cdot R_p \right|}{\dfrac{1}{C_p j\omega} + R_p} = \frac{R_p}{1 + R_p C_p j\omega} = \frac{R_p - R_p^2 C_p j\omega}{1 + R_p^2 C_p^2 \omega^2} \qquad \textit{(Equation 8.11)}$$

We can separate out the real and imaginary parts of this impedance as:

$$Re\{Z\} = \frac{R_p}{1 + R_p^2 C_p^2 \omega^2} \qquad\qquad Im\{Z\} = \frac{-R_p^2 C_p \omega}{1 + R_p^2 C_p^2 \omega^2}$$

The dissipation factor (D) is given by the ratio of the real part to the imaginary part:

$$D = \frac{Re\{Z\}}{Im\{Z\}} = \frac{\dfrac{R_p}{1 + R_p^2 C_p^2 \omega^2}}{\dfrac{-R_p^2 C_p \omega}{1 + R_p^2 C_p^2 \omega^2}} = -\frac{1}{R_p C_p \omega} \qquad \textit{(Equation 8.12)}$$

For thin gate dielectrics, large leakage currents imply a small value for the equivalent parallel resistor. This means that the current flowing through the capacitor is decreased, making it difficult to measure accurately. To increase the current flow through the capacitor (decrease the D), a higher measurement frequency is required to reduce the impedance of the capacitor (which is equal to $1/Cj\omega$). Looking at this in the impedance plane, a smaller Rp implies a larger D (dissipation). This means that the accuracy of the capacitance measurement is degraded and that the influence from external noise is increased. For these reasons, thin gate oxide measurement requires higher measurement frequencies. The relationship between Rp, Cp and D is shown below.
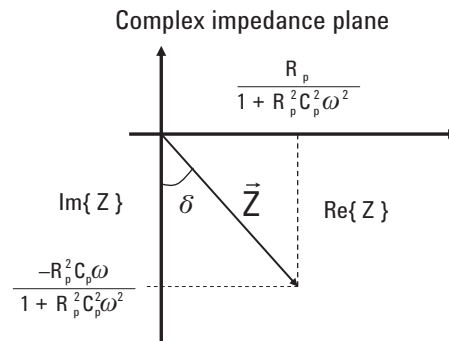
$$D = \tan \delta = \frac{Re\{Z\}}{Im\{Z\}} = -\frac{1}{R_p C_p \omega}$$

Rp is small $\rightarrow$ D is large

Figure 8.22. The relationship between parallel resistance (Rp) and dissipation (D).

Note: The above results imply that the dissipation can be made arbitrarily small by increasing the frequency. However, as we will see later when we discuss the three-element capacitance model this is not necessarily true for thin gate dielectric devices.

Most capacitance meters or impedance analyzers have the ability to plot the value of D in real time. For the case of thick gate dielectrics, it is quite common to see values of D much less than 1 (indicating that most of the measurement current is flowing through the capacitor rather than the parasitic resistor). However, for thin gate dielectrics the user's expectations for an acceptable value of D need to be adjusted. Even increasing the frequency often does not reduce the value of D to a value less than 1. It is difficult to come up with a clear "rule of thumb" for an acceptable dissipation factor for thin gate dielectrics, but usually a value of less than 5 is sufficient to achieve a satisfactory capacitance measurement.

Even though it is clear that we need to go to higher frequencies to measure thin gate dielectric transistor gate capacitance, the two element model cannot explain the results observed at higher frequencies.
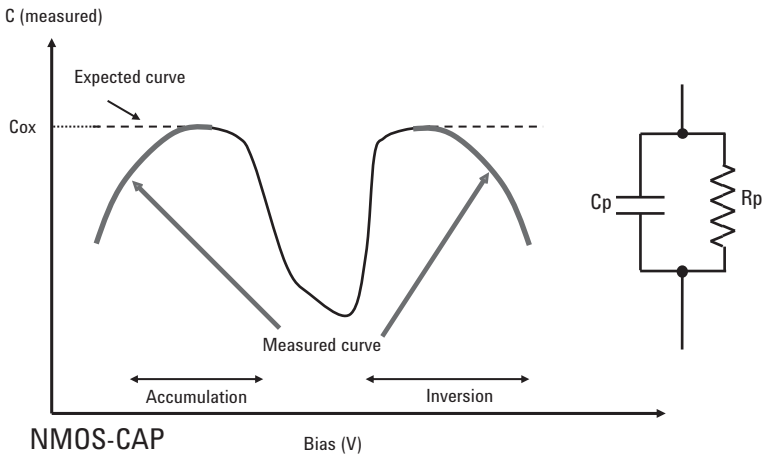
Figure 8.23. The two element capacitor model cannot explain the results seen when measuring the gate capacitance on thin gate dielectric transistors at higher frequencies.

In order to properly characterize thin-gate transistor dielectrics we need to move to the three element capacitor model shown below that includes contact and substrate resistance.
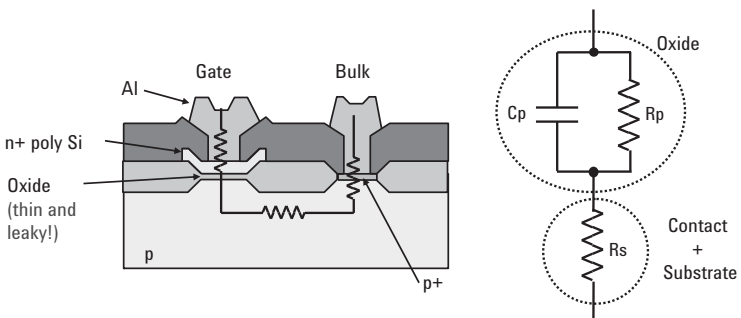
Figure 8.24. The three element capacitor model (which includes contact and substrate resistance).

165

For the conventional thick gate dielectrics, the equivalent parallel resistance R_p is much larger than the series resistance R_s. This means that R_s can be ignored when measuring at low frequency ranges under 1 MHz. This explains why the parallel model was valid for the thicker gate dielectrics of the past.

Many engineers noticed what seemed to be a "negative capacitance" when they started to measure thin gate oxides. An example of the observed measurement results is shown below.
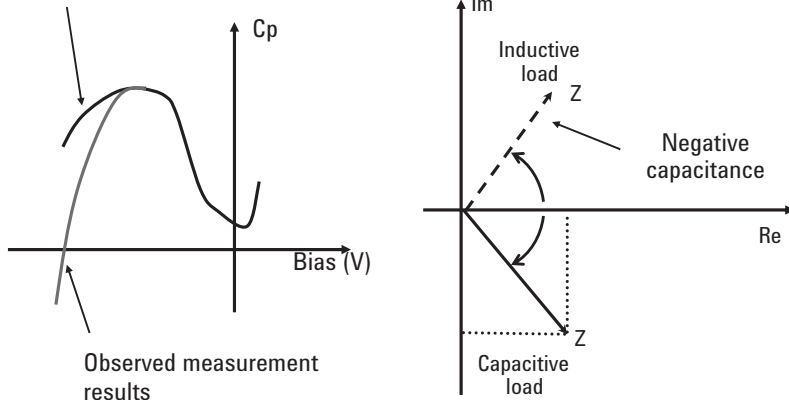
Figure 8.25. Illustration of negative capacitance effect observed when measuring thin gate dielectric capacitors.

Negative capacitance results when the measured impedance has a positive value for its imaginary part. In general, positive imaginary impedance implies inductance. This led many people to try to explain the observed negative capacitance as an inductance caused by wiring in the wafer prober or even by the generation and recombination of holes and electrons. However, none of these approaches were successful.

Agilent found that the negative capacitance comes from the four terminal pair measurement technique used for on-wafer measurement, not from the device itself. The following circuit diagram shows the whole measurement system including the wafer chuck.
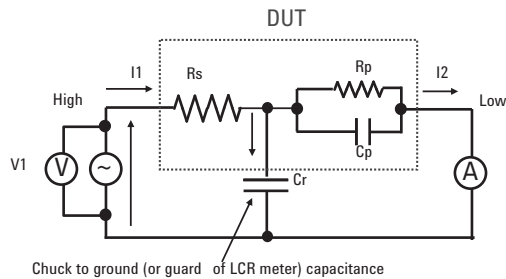
$$Zmeas = \frac{V1}{I2} = \frac{Rp + Rs + \omega^2 Cp^2 Rp^2 Rs + \omega^2 Cp\ CrRp^2\ Rs - \boxed{j\omega Rp(CpRp - CrRs)}}{1 + \omega^2 Cp^2 Rp^2}$$

The imaginary part of Zmeas becomes positive when CpRp < CrRs.
This gives the appearance of a negative capacitance!

Figure 8.26. The explanation of the negative capacitance effect seen in thin gate dielectrics.

The key point is that capacitive coupling occurs between the wafer chuck and the return path of the signal (outer shield of the BNC cable or Earth). When using a four terminal pair, part of the signal flowing through the DUT leaks through the wafer chuck via capacitive coupling. In this case the measured impedance is described by the equation shown above. From this equation, if the product of Cp and Rp is smaller than the product of Cr and Rs, the imaginary part of the measured impedance becomes positive. Because Rp comes from the direct tunneling of electrons through the gate dielectric, Rp rapidly becomes smaller with increasing gate bias. This means that a negative capacitance is observed when the gate bias goes over a value that satisfies the condition of Cp × Rp < Cr × Rs. This analysis explains the observed measurement results when performing the CV measurements on thin gate oxide devices. Note: The best way to eliminate this effect is through structure design (minimizing Rs), rather than through improved measurement techniques. More information on this can be found in a paper by Y. Okawa, et al, from ICMTS 2003 "The negative capacitance effect on the C-V measurement of ultra thin gate dielectrics induced by the stray capacitance of the measurement system".

Optimizing CV measurement results through proper structure design

In order to improve measurement performance, we want the contact resistance (Rs) to be as small as possible. Remember from the previous discussion that problems first arose when Rs started to become comparable to Rp. Therefore, reducing Rs through prudent layout design should help to improve our measurement results. Also remember that when Cp × Rp < Cr × Rs a "negative capacitance" will be observed. Therefore, by minimizing Rs we can eliminate this effect. The following figure is a good example of a test structure designed to minimize Rs.
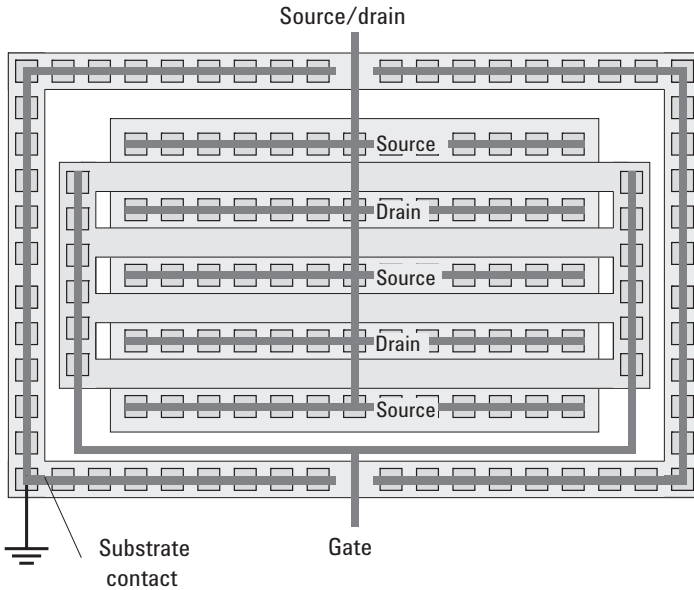
Figure 8.27. A capacitance measurement structure designed to minimize Rs and eliminate the negative capacitance effect.

The details of this structure are not important (and indeed the details are not shown). The important points to remember are:

1. Make the structure symmetrical
2. Provide lots of contacts to minimize contact and substrate resistance
3. Use a ground-signal (GS) or ground-signal-ground (GSG) layout to connect to the pads

Note: The use of these types of test structures to minimize Rs was first proposed by J. Schmitz, et al, at ICMTS 2003: "Test structure design considerations for RF-CV measurements on leaky dielectrics".

Pad layout considerations

As the previous discussions have shown, a dedicated MOSFET capacitance structure is needed to get good capacitance measurement results at high frequencies. In addition to the structure design, the actual layout of the pads is also important. Putting some forethought into the pad layout can save you many measurement headaches later. Although high-frequency capacitance measurements require RF probes for optimal results, this does not necessarily mean that ground-signal-ground probe tips are always necessary. The following illustration shows two basic MOSFET capacitor layouts (NMOS and PMOS) that only require a ground-signal probe tip in order to make the capacitance measurement.
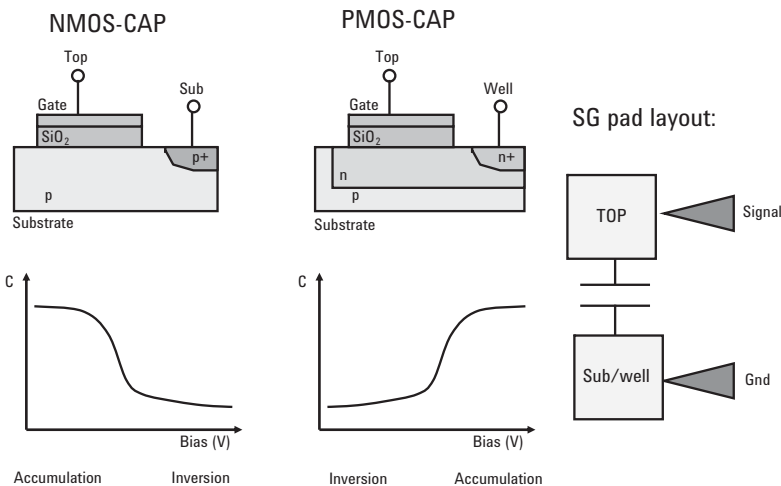
Figure 8.28. Using a signal-ground (SG) style pad layout you can perform simple CV sweeps (gate to substrate) on both NMOS and PMOS capacitors.

If you want to measure both gate to substrate and gate to source/drain capacitance, then the following (GSG) scheme meets this need and only uses two probes at one time.
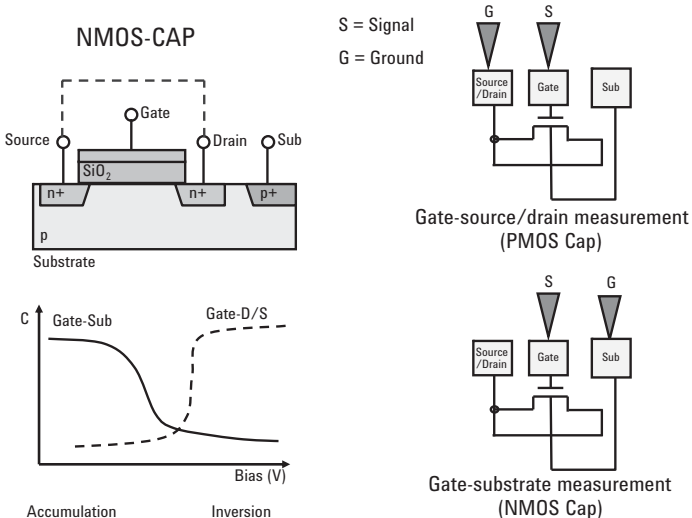
Figure 8.29. Using a carefully designed three terminal (GSG) style pad layout you can measure both gate to substrate and gate to drain/source capacitance.

Using the above scheme can save you some money, since GS probe tips are less costly than GSG probe tips.

OPEN/SHORT/LOAD calibration

As previously mentioned, above 5 MHz simple OPEN/SHORT calibration is no longer sufficient to guarantee measurement accuracy. The following plot compares the measurement results obtained using the 4TP cabling technique with an Agilent 4294A Impedance Analyzer (which has a frequency range up to 110 MHz) for both OPEN/SHORT and OPEN/SHORT/LOAD calibration.
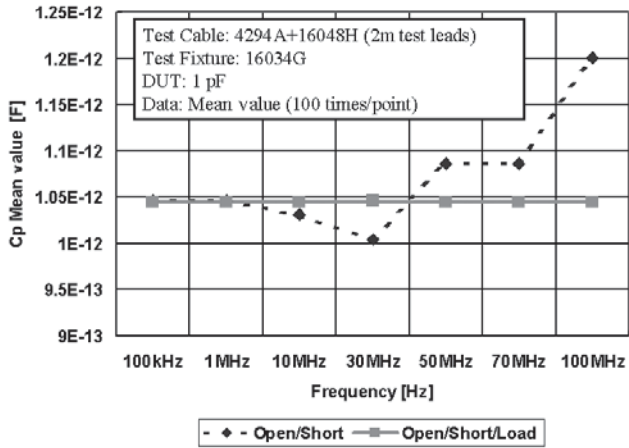
Figure 8.30. The measurement error that occurs at higher frequencies unless a load calibration is performed.

Please refer to Agilent application note 4294-3 (publication number 5988-5102EN) for more information. As this measurement data shows, significant measurement error occurs above 5 MHz unless the LOAD compensation is performed in addition to the basic OPEN/SHORT calibration.

The OPEN/SHORT compensation scheme previously discussed was an over-simplification. The actual OPEN/SHORT/LOAD scheme treats the cabling as an unknown 2-terminal pair circuit, and it determines the true value of the DUT impedance as shown in the diagram shown below.

Unknown 2-terminal pair circuit
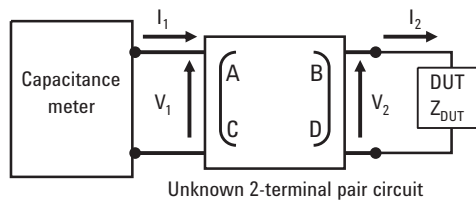
$$Z_{DUT} = \frac{Z_{Std}(Z_0 - Z_{SM})(Z_{XM} - Z_S)}{(Z_{SM} - Z_S)(Z_0 - Z_{XM})}$$

Z_0 : Measured value of OPEN
Z_{SM} : Measured value of standard DUT
Z_{XM} : Measured value of the DUT

Z_S : Measured value of SHORT
Z_{Std} : True value of the standard DUT
Z_{DUT} : Corrected value of the DUT

Figure 8.31. The OPEN/SHORT/LOAD compensation scheme actually performs an analysis on the unknown network connecting the capacitance meter with the DUT.

It cannot be over-emphasized that the LOAD portion of the calibration requires a well-calibrated resistance standard, and that this calibration has to occur at the ends of your probe tips. The accuracy of ALL of your subsequent high-frequency measurements depends upon the accuracy of the load standard, so it is very important to use a standard with known characteristics. Analytical wafer prober companies that support RF probing can supply standards for this purpose. In addition, you must make sure that the calibration plane for the OPEN/SHORT/ LOAD calibrations is always the same or you will get inaccurate measurement data. For example, when using probe cards one common mistake is to perform OPEN/SHORT compensation at the probe tips and then perform the LOAD compensation using a special load standard built into the probe card. However, since the calibration plane is not the same in all three cases the result is inaccurate measurement at high frequencies.

Calculating the values of the three element capacitor model
The three element capacitor model obviously requires the determination of three parameters: Cp, Rp and Rs. While this might seem to require three measurement points this is actually not the case. Remember that each measurement of imped-ance actually supplies two pieces of data: magnitude and phase. This means that two impedance measurements actually supply four pieces of data, which is more than sufficient to determine the three unknown elements. This concept can be applied to the three element capacitor model by measuring capacitance and dissipation at two different frequencies and using the equations shown below.
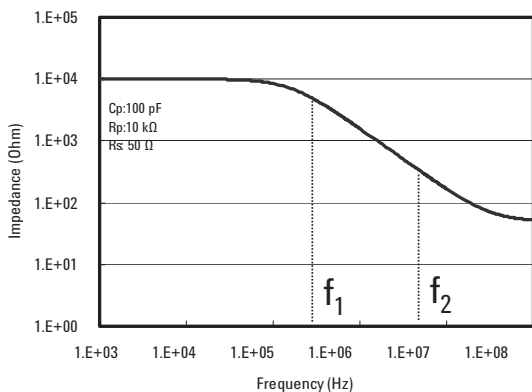
$$Cp = \frac{f_1^2 C_1 (1 + D_1^2) - f_2^2 C_2 (1 + D_2^2)}{f_1^2 - f_2^2}$$

$$Rp = \frac{1}{\sqrt{\omega_1^2 C_1 Cp \,(1 + D_1^2) - \omega_1^2 Cp^2}} = \frac{1}{\sqrt{\omega_2^2 C_2 Cp \,(1 + D_2^2) - \omega_2^2 Cp^2}}$$

$$Rs = \frac{D_1}{\omega_1 C_1 (1 + D_1^2)} - \frac{Rp}{1 + \omega_1^2 Cp^2 Rp^2} = \frac{D_2}{\omega_2 C_2 (1 + D_2^2)} - \frac{Rp}{1 + \omega_2^2 Cp^2 Rp^2}$$

C_1: Measured capacitance by parallel model at f_1
D_1: Measured dissipation by parallel model at f_1
C_2: Measured capacitance by parallel model at f_2
D_2: Measured dissipation by parallel model at f_1

Figure 8.32. The two frequency method for calculating the parameters for the three element capacitor model for thin gate oxides.

Unfortunately, the two frequency method has been shown to be unreliable for calculating the three element capacitor model for thin gate oxides (please refer to paper by R. Clerc, et al, at INFOS 2001). Therefore, an alternate method that involves plotting the impedance and phase across frequency and then using data obtained at the point of minimum phase angle was proposed. This method is shown below.

$$R_s = |Z_0| \frac{1 + \sin\theta_0}{\cos\theta_0}$$

$$R_p = -\, 2|Z_0|\tan\theta_0$$

$$C = \frac{|Z_0|}{\omega_0 R_p R_s} = \frac{1}{4\pi f_0 |Z_0|}\left(\frac{\sin\theta_0 - 1}{\sin\theta_0}\right)$$

f_0: frequency at phase minimum
Z_0: impedance at phase minimum
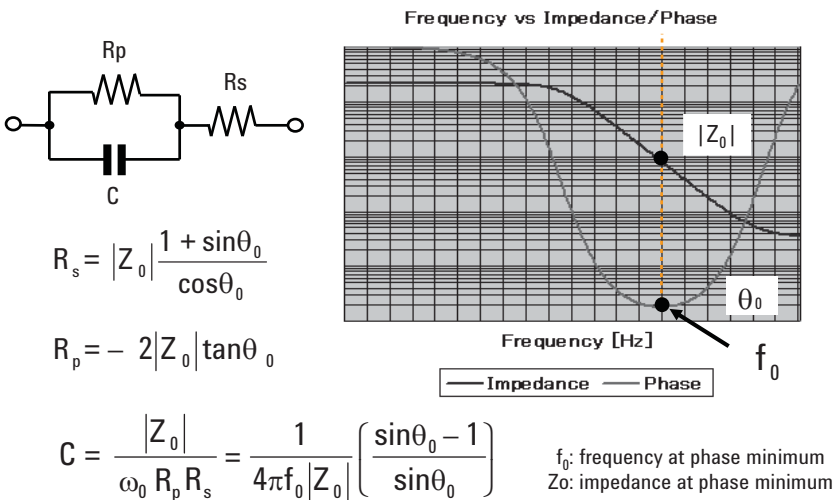
Figure 8.33. The minimum phase method (shown above) is the preferred method for calculating the parameters for the three element capacitor model.

Please refer to Agilent application note 4294-3 (publication number 5988-5102EN) for more information.

An interesting point regarding the three element capacitor model is that applying a higher frequency does not necessarily result in a lower value for the dissipation factor (D). The equivalent impedance for the three element case consists of the parallel combination of Rp and Cp in series with Rs. Referring back to equations 8.11 and 8.12 we can see that for the three element model the equivalent impedance is given by:

$$Z = R_s + \frac{R_p - R_p^2 C_p j\omega}{1 + R_p^2 C_p^2 \omega^2} \qquad \text{(Equation 8.13)}$$

We can separate out the real and imaginary parts of this impedance as:

$$\text{Re}\{Z\} = R_s + \frac{R_p}{1 + R_p^2 C_p^2 \omega^2} \qquad\qquad \text{Im}\{Z\} = \frac{-R_p^2 C_p \omega}{1 + R_p^2 C_p^2 \omega^2}$$

As before, the dissipation factor (D) is given by the ratio of the real part to the imaginary part:

$$\text{(Equation 8.14)}$$

$$D = \frac{\text{Re}\{Z\}}{\text{Im}\{Z\}} = \frac{\dfrac{R_s(1 + R_p^2 C_p^2 \omega^2) + R_p}{1 + R_p^2 C_p^2 \omega^2}}{\dfrac{-R_p^2 C_p \omega}{1 + R_p^2 C_p^2 \omega^2}} = -\frac{R_s(1 + R_p^2 C_p^2 \omega^2) + R_p}{R_p^2 C_p \omega}$$

Unlike the two element model case, dissipation does not continuously decrease with increasing frequency but actually has a minimum value. The following plot of dissipation versus frequency for a three element capacitor illustrates this point.

Three-element equivalent circuit: D vs. frequency
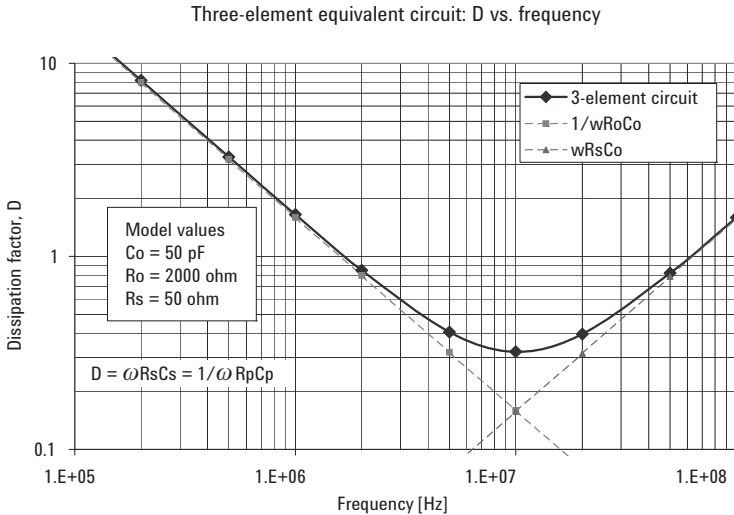
Figure 8.34. This three element capacitor model plot shows that arbitrarily increasing the frequency does not necessarily result in a lower value for dissipation.

Therefore, in practice it may take some trial and error to determine the optimal frequency at which to perform measurements. ***Note:*** This behavior was first documented by G. Brown, at ICMTS 2005: "Capacitance Characterization in Integrated Circuit Development: The Intimate Relationship of Test Structure Design, Equivalent Circuit, and Measurement Methodology".

The advanced IV method
The 4TP cabling technique has limitations. As frequency increases not only chuck capacitance but also the residual chuck inductance starts to impact the measurement results. The best means to minimize these effects is to ground the wafer chuck, but as has been mentioned previously the 4TP method requires that the wafer chuck be floating. For this reason, Agilent developed an alternative fixturing (cabling) solution for making high-frequency measurements on-wafer known as the advanced IV method. An illustration of this method is shown below.
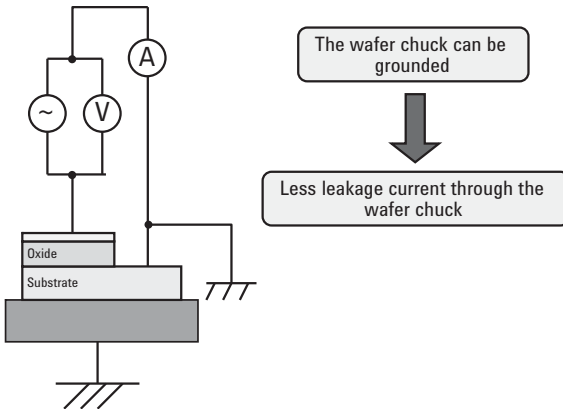
Figure 8.35. The advanced IV method allows the wafer chuck to be grounded, resulting in less leakage current through the wafer chuck due to residual capacitance and inductance.

The advanced IV method has the following advantages:

1. The low side of the current meter is grounded.
2. The voltage source and the voltage meter are floating.
3. Frequency measurements up to 110 MHz can be supported.

The advanced IV method requires the Agilent 4294A Precision Impedance Analyzer with the Agilent 42941A Impedance Probe. You cannot use the 42941A probe with the E4980A LCR meter. A picture of this instrument and the probe are shown below.

Figure 8.36. The 4294A precision impedance analyzer and the 42941A impedance probe are both required to perform the advanced IV method of CV measurement.

The 4294A has certain idiosyncratic features of which all users should be aware. One of these is a feature known as the "cable terminated auto-balancing bridge method." Essentially, what this means is that the 4294A inserts 50 Ω termination resistors to reduce reflected waveforms when the frequency reaches 5 MHz or above whenever 1 or 2 meter cables are used with the instrument (which of course is always the case when making on-wafer measurements). The figure below illustrates this feature.
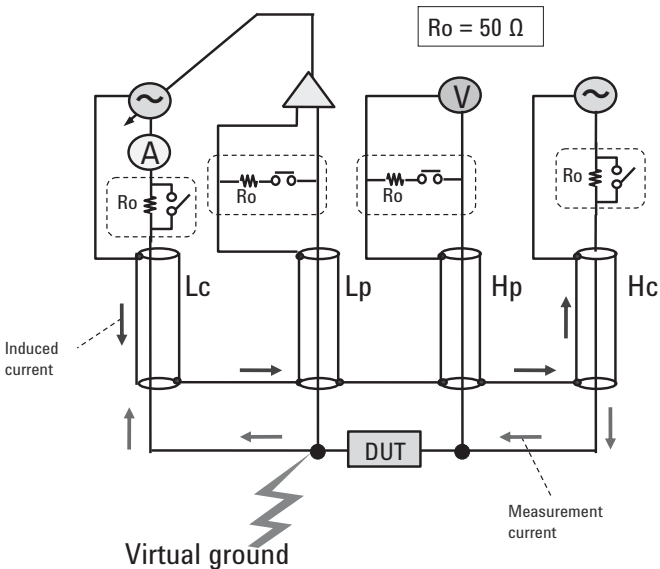
Figure 8.37. The 4294A switches in 50 Ω resistors to reduce waveform reflections at frequencies of 5 MHz and above when 1 and 2 meter cable extensions are used.

Unless a load calibration is performed at 5 MHz, a discontinuity in the capacitance measurement will be observed at 5 MHz. Therefore, during the 4294A load compensation it is essential to make sure the 5 MHz point is included in the frequency compensation list.

Extremely thin gate oxides usually require both the advanced IV method and careful test structure layout in order to obtain satisfactory measurement results. The following graph shows a conventional test structure design with a ground-signal-ground (GSG) pad layout measured using the advanced IV method.
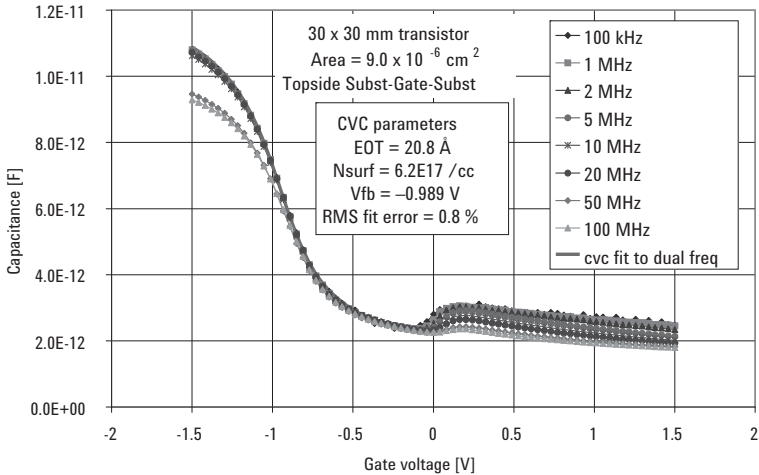
Figure 8.38. The measurement results obtained across frequency using the advanced IV method on a conventional test structure design (GSG pad layout). [Note: Data courtesy of SEMATECH]

As this plot shows, there are measurement variations across frequency and the device does not exhibit QSCV behavior in the inversion region. In contrast, the following graph shows a test structure with a ground-signal-ground (GSG) pad layout measured using the advanced IV method that follows the design guidelines previously discussed to minimize the series resistance (Rs).

Figure 8.39 - The measurement results obtained across frequency using the advanced IV method on a test structure designed to minimize the series resistance (GSG pad layout). [Note: Data courtesy of SEMATECH]

175

This measurement shows excellent correlation across measurement frequency as well as a QSCV response in the inversion region, and the electrical oxide thickness calculates out to 2 nanometers! Again, this result reinforces the importance of proper structure design when making capacitance measurements on extremely thin gate oxides.

Controlling the 4294A Using EasyEXPERT
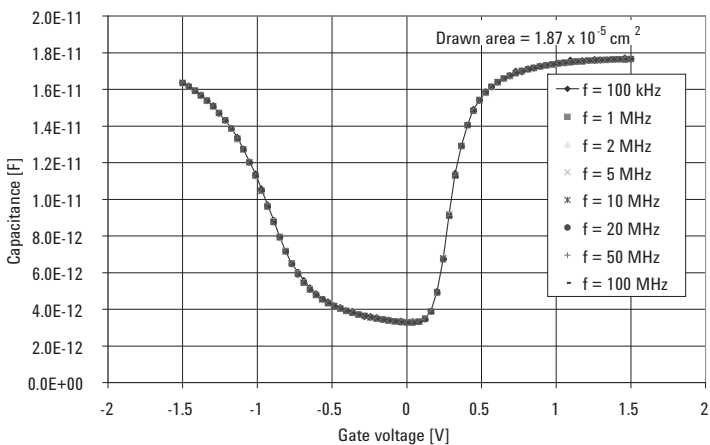The front panel user interface of the 4294A is complex with many sub-menus. In addition, the 4294A has a couple of features that sometimes cause users to make inadvertent errors. The first of these is that, unlike the B1500A/B1505A and E4980A, the 4294A does not retain cabling compensation information during power-down. One solution is to always perform open/short/load compensation each time the 4294A is powered up, although this can be a bit tedious and impractical. Fortunately, the 4294A does have a built-in Flash memory that permits compensation data to be saved into a file and retrieved after power-up; however, the user has to remember to do this. The second issue with the 4294A is that many of its default settings are not appropriate for semiconductor capacitance measurement. Again, it is a little tedious to change all of the default setting each time the instrument is powered-up and it is also easy to miss something or make a mistake.

To help with these issues, Agilent has made available EasyEXPERT application tests that can control the 4294A. One of these application tests simply guides the user through the 4294A open, short and (optionally) load calibration procedures. After the calibration procedures are completed the application test stores the calibration data into files within the 4294A's Flash memory. The other application tests permit the user to perform capacitance versus dc voltage (CV) sweeps and capacitance versus frequency (C-f) sweeps. Importantly, both the CV and C-f application tests first check for the presence of the stored calibration files created by the calibration application test. If these files are not present then the application test issues a message informing the user that they must first perform cable compensation and the application test terminates. Of course, if the calibration information is present then the CV and C-f application tests automatically load the calibration data into the instrument. In addition, both the CV and C-f application tests take care of changing the 4294A default settings into values appropriate for semiconductor capacitance measurement. Therefore, these application tests make it almost impossible to take invalid data due to calibration or setup issues. Finally, these application tests make it easy for the user to store and export the test results into PC-based formats, which is not possible from the front panel of the 4294A.

A screen capture of the EasyEXPERT 4294A CV application test is shown below.
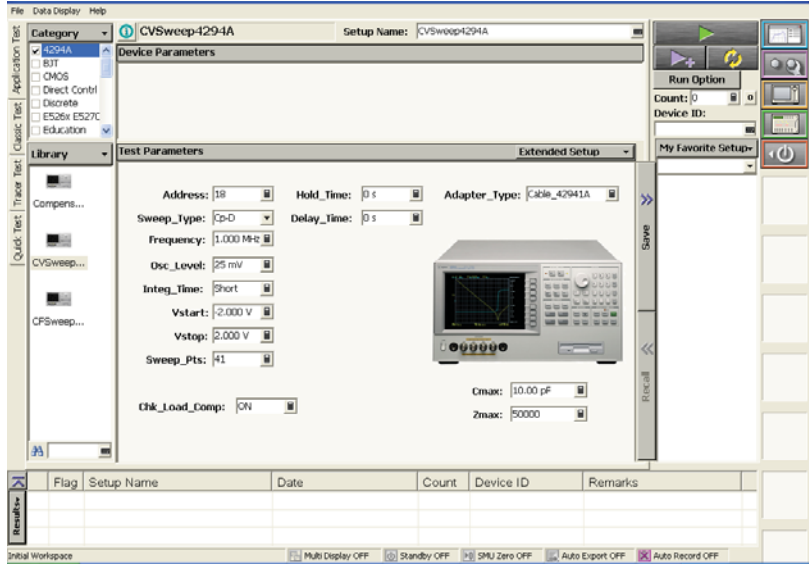
Figure 8.40. The EasyEXPERT application test to perform capacitance versus dc voltage sweeps using the 4294A.

This application test provides control of all of the important CV sweep parameters for the 4294A. If you cannot perform or do not want to perform LOAD compensation, then you can change the "Chk_Load_Comp" input to "OFF", and the application test will not check the stored calibration file to determine if the LOAD compensation was performed during calibration. Note: The "Integ_Time" parameter is actually a simplification of the 4294A's settings, as the 4294A does not actually support integration time in the same sense as (for example) the B1500A. However, for the vast majority of users this feature is an improvement over the innate settings of the 4294A.

A screen capture of the EasyEXPERT C-f application test is shown below.
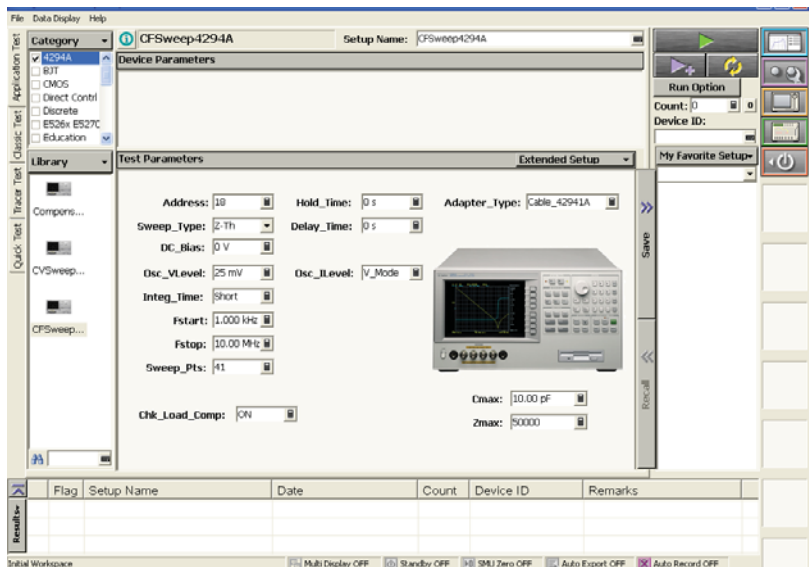
Figure 8.41. The EasyEXPERT application test to perform capacitance versus frequency sweeps using the 4294A.

The key features of this application test are similar to those of the CV sweep application test. One difference is that this application tests permits the user to set the ac oscillator into either voltage or current mode. A numeric value of "0" should be entered into the oscillator level input that is not being used (a symbol name such as "V_Mode" or "I_Mode" will automatically appear). Another difference is that, although the capacitance parameters are plotted on a linear scale, the frequency values are plotted on a logarithmic scale.

Summary of thin gate dielectric best practices

In summary, the following are the key factors to keep in mind when attempting to measure capacitance on thin gate dielectric transistors:

1. Perform an OPEN/SHORT/LOAD calibration (after first performing PHASE compensation if it is supported) before making any measurements.

2. Use a frequency high enough to minimize dissipation and maintain accuracy.

3. Use the 3 element device model to extract the device parameters.

4. Account appropriately for the effects of the wafer chuck through careful layout   of the test structure (minimize Rs).

5. If the 4TP cabling method does not yield satisfactory results, consider using the advanced IV method.

Measurement method	Measurement probe contact point	Instrumentation	Comments
4TP	Wafer or chuck (chuck preferred)	E4980A, B1500A, B1505A, 4294A	• Complex setup • Max freq ~30 MHz
Advanced IV	Wafer	4294A + 42941A	• Easy setup • Max freq to 110 MHz • Requires RF probes

Figure 8.42. Table summarizing the 4TP and Advanced IV measurement methods.

Making capacitance measurements through a switching matrix

One long-running issue in parametric test has been called "the CV-IV measurement dilemma." The dilemma referred to has to do with the different cabling needs of capacitance meters and SMUs. Capacitance meters use BNC cables and SMUs use triaxial cables, yet most parametric testing requires both CV and IV measurements. Manually switching from triaxial to BNC cables (and back again) is extremely tedious, and it also does not permit any sort of test automation. Obviously, switching matrices (discussed in Chapter 4) represent one solution to this problem. However, by introducing a switching matrix into the CV measurement path to take care of the BNC to triaxial conversion we solve one issue and create several new ones. In this section we will discuss the proper way to make capacitance measurement through a switching matrix.

Switching matrix capacitance measurement issues
When we discussed making on-wafer measurements in Chapter 4 we covered Agilent's positioner-based CV-IV switching solutions. As discussed in that chapter, Agilent's positioner-based solutions have both built-in capacitance compensation (to the outputs of the ASU and SCUU) and the means to short the guards together during capacitance measurement to stabilize the series inductance of the cables. However, it should be clear that performing these same operations on a switching matrix is much more difficult. Positioner-based solutions have the advantage of a fixed cable length; switching matrices must contend with signal path lengths and loadings that change each time that an input-to-output relay connection is changed.

The figure shown below outlines the three sections that must be considered when making capacitance measurements through a switching matrix.
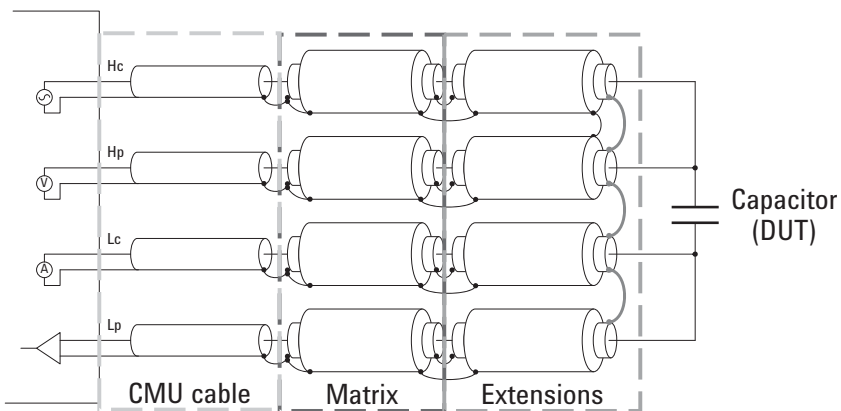
Figure 8.43. When measuring capacitance through a switching matrix there are three sections to consider: the CMU cable, the matrix, and the extensions.

Most of the errors that occur when making capacitance measurements through a switching matrix can be attributed to one or more of the following factors:

1. Non-standard CMU cables or an unsupported cable length are used to connect the capacitance meter to the switching matrix.
2. The guards are not shorted together close to the DUT.
3. The cable extensions create cable impedance mismatch.
4. The matrix does not support any sort of compensation scheme that takes into account changes in loading and path length due to the opening and closing of input-to-output connections.

In the following discussion we will examine the best means to mitigate these factors.

CMU cable compensation
Although we have already discussed the issue of capacitance meter cable compensation quite extensively, it is worthwhile to repeat some of the important points again since many times these seem to be forgotten when a switching matrix is involved. One of the key points is the length of the capacitance meter cables. Proper cable length is critical for the bridge of the capacitance meter's measurement circuitry to balance. If an improper cable length is used, then the capacitance meter may not be able to balance its bridge and you will get an error during the compensation process.
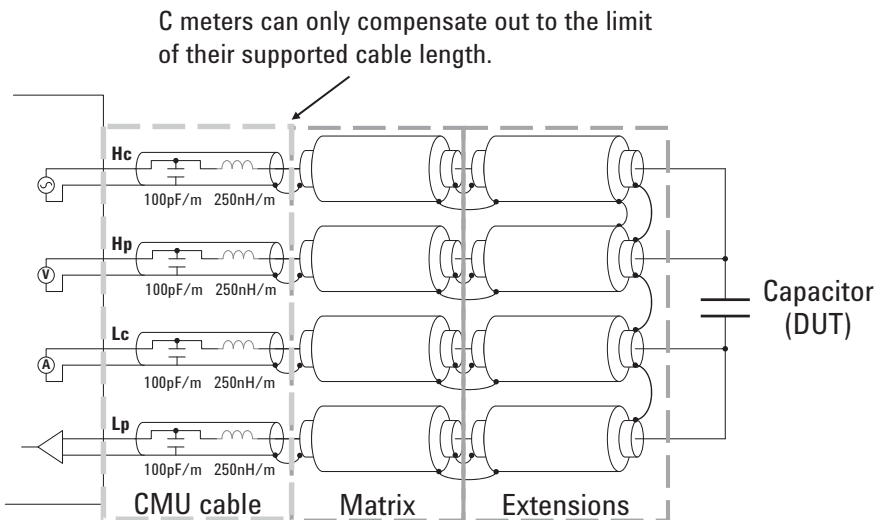
Figures 8.44. Using an unsupported CMU cable or an unsupported cable length are common reasons for poor measurement results when making capacitance measurements through a switching matrix.

The figure shown below describes the correct procedure to compensate the cabling going from the capacitance meter to the switching matrix.
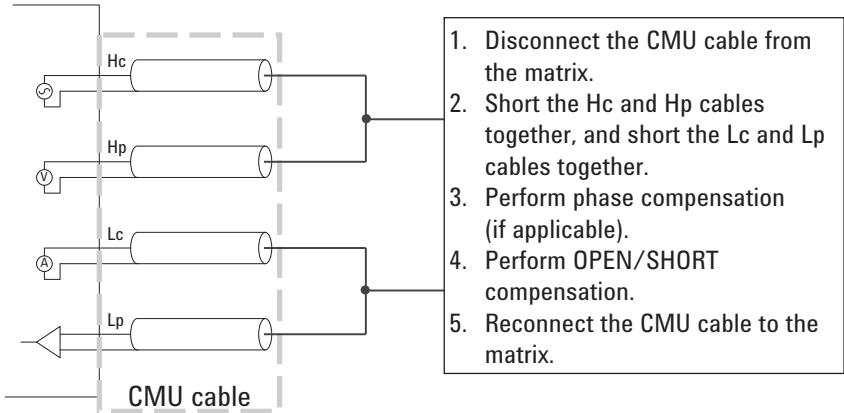
Figure 8.45. The proper procedure to compensate the cable going from the capacitance meter to the switching matrix.

It is also very important to use Agilent-supplied cables to connect to the matrix. Not only are the Agilent cables at the correct cable lengths for proper compensation, but they are manufactured and tested to known quality standards and they have known characteristic impedance. If you try to use off-the-shelf BNC cables then Agilent cannot guarantee anything about the accuracy of your measurements even if you perform compensation and everything seems to pass OK. The Agilent E4980A supports cable lengths of 1 m, 2 m, and 4 m. The Agilent B1500A MFCMU supports cable lengths of 1.5 m and 3 m. You should try to use as short a cable as possible to connect to your matrix.

Shorting the guards

Although the reasons for shorting the guards together during capacitance measurement were discussed in Chapter 4, we will briefly review them again. The outer shield of the BNC cables coming from a capacitance meter are not at ground potential, but are "virtual grounds". The reason for shorting the driven guards together during capacitance measurement is to stabilize the series inductance of the measurement path. If the guards are not shorted then the series inductance can change from 250 nH/m to more than 400 nH/m. This much variation in the series inductance of the measurement path makes it virtually impossible to get stable capacitance compensation.
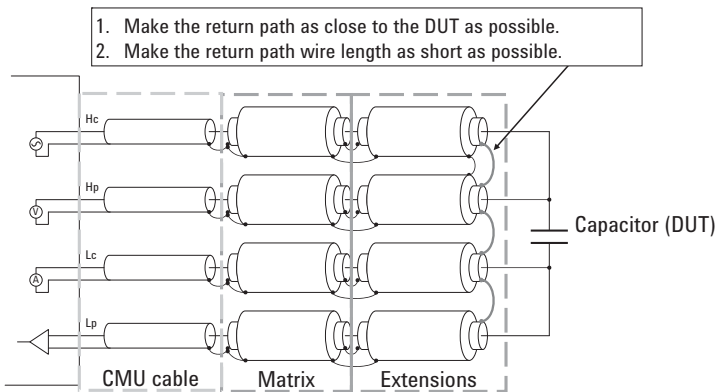
Figure 8.46. Important considerations when shorting the guards together during capacitance measurement.

Impedance mismatch

Capacitance meters expect a 50 Ohm cable environment when they perform OPEN/SHORT/LOAD compensation. Unfortunately, neither the switching matrix nor the cable extensions have characteristic impedances of 50 Ohms. This affects the accuracy of the compensation.
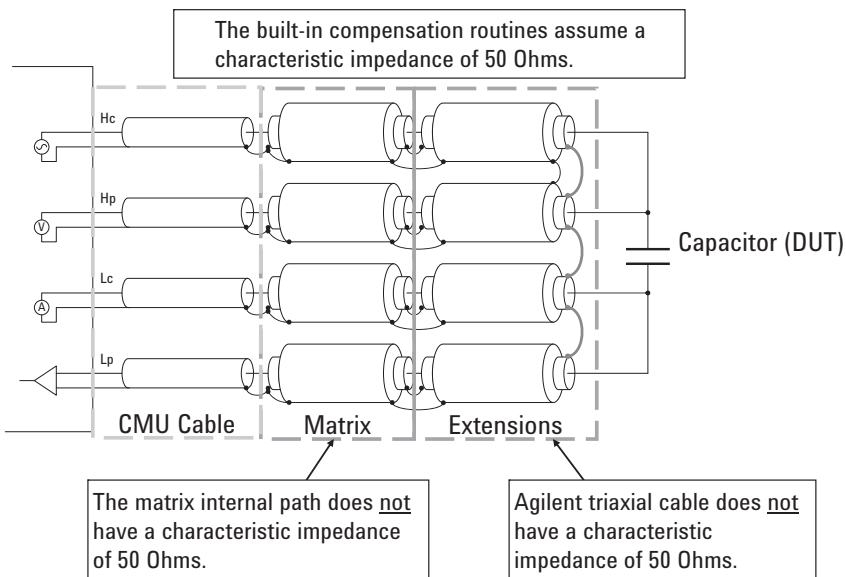
Figure 8.47. One major challenge is that the matrix path and cable extensions do not have a characteristic impedance of 50 Ohms.

Ideally the LOAD impedance should be in the same range as the device impedance that you want to measure. However, this is never the case for semiconductor devices. Unfortunately, the typical input impedance of a semiconductor device is very much higher than 50 Ohms. This affects the accuracy of the LOAD compensation.
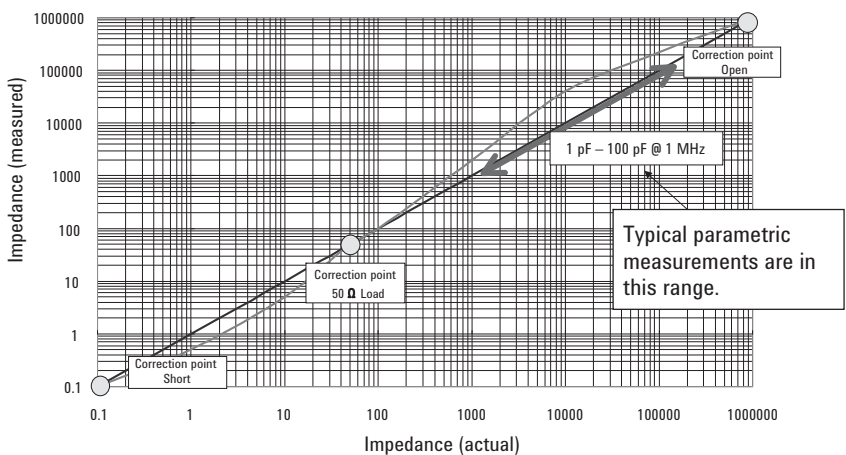
Figure 8.48. This graph illustrates why semiconductor load impedance in not ideal for making capacitance measurements.

Agilent's switching matrix solution

Agilent can supply a complete CV-IV solution that compensates for the cabling and fixturing all the way down to the probe card tips. Besides improving CV measurement, this solution also yields excellent low-current measurement performance.

Return path is connected in the B2220A probe card I/F.

The B1500A EasyEXPERT software supports B2200A and B2201A compensation.

Connection cables supported by the compensation routine.

Optimized CV inputs in the switching matrix.
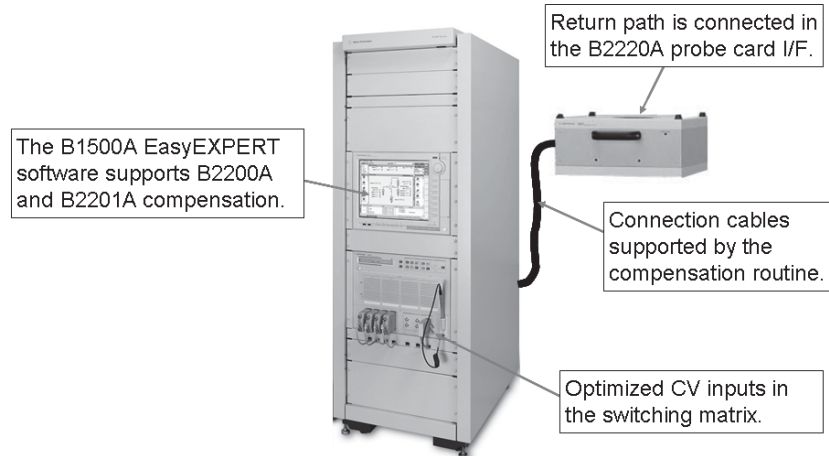
Figure 8.49. Agilent's integrated switching matrix CV-IV capacitance measurement solution.

The details of this solution are shown below.

To Agilent B1500A MFCMU

Perform OPEN compensation at this plane. SHORT compensation is optional.

Agilent N1300A – 001/002

Agilent B2220A (C2H/C2L Fixed)

To CML   To CMH

Case 1: Probe card interface

Agilent 16494A/C cable (C1 fixed)

Probe card, etc.; C3H/C3L need update

Agilent 16495F/G

Case 2: Connector plate

Agilent 16494A/B Cable (C1 fixed)

Agilent B2200A or B2201A

B2200A/B2201A outputs

Extended coaxial cables with connector plate; C3H/C3L need update

C2H, C3H for CMH path
C2L, C3L for CML path

Extended triaxial cables with connector plate; C2H/C2L need update

Figure 8.50. The EasyEXPERT B2200A/B2201A matrix compensation utility has built-in compensation data for both Case 1 and Case 2 shown above.

The following information is provided as a reference for the figure shown above.

> B2220A-024/048: Probe card interface (24 pins/48 pins)
> N1300A-001/002: CMU cable for B1500A (1.5 m/3.0 m)
> 16494A-002/005: Triaxial cable (3.0 m/4.0 m)
> 16494B-001/002: Kelvin triaxial cable for 16495F/G (1.5 m/3.0 m)
> 16494C-002/005: Kelvin triaxial cable for B2220A (3.0 m/4.0 m)
> 16495F: Connector plate with 12 triaxial, interlock, and GNDU connectors
> 16495G: Connector plate with 24 triaxial, interlock, and GNDU connectors

EasyEXPERT also allows you to use connection options other than the B2220A Probe Card Interface or 16495F/G connector plate. However, you then have to calibrate your fixturing and create your own Compensation Data file. This process is outlined in the EasyEXPERT help menu.

The above solution works because the matrix and extension path is well-defined with known correction coefficients. EasyEXPERT will support this automated compensation up to frequencies of 1 MHz.

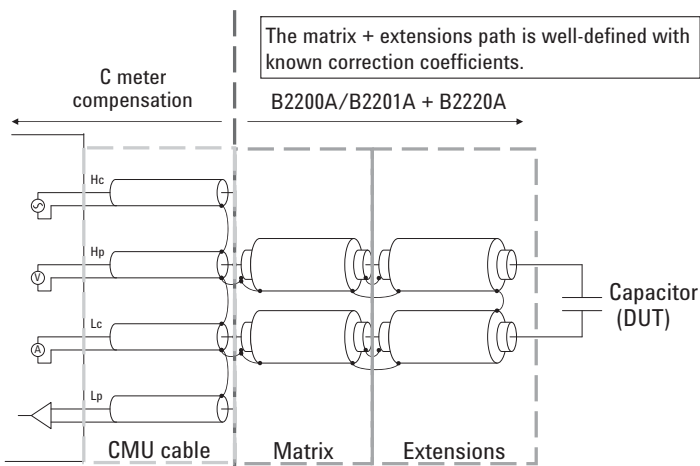Figure 8.51. Schematic overview of Agilent's switching matrix compensation scheme.

Tips if trying to go above 1 MHz

Before proceeding it is important to emphasize that Agilent does not offer any guarantees that you can make capacitance measurements through any of our switching matrix products at frequencies above 1 MHz. This is an unsupported measurement application. However, if you have a need to do this then we can offer some suggestions to maximize your chances for success. If you are trying to use a switching matrix above 1 MHz then here are the key points to keep in mind:

1. Your C meter has a maximum supported cable length. If the total length of your cabling (connection cable + matrix path + extensions) exceeds the length supported by your C meter, then your OPEN/SHORT/LOAD compensation will fail (you will get a "bridge unbalanced" message).

2   Every time you change your matrix settings, your path length will change. You need to perform OPEN/SHORT/LOAD compensation every time that you change the switch settings.

3. You must supply a current return path near the DUT.

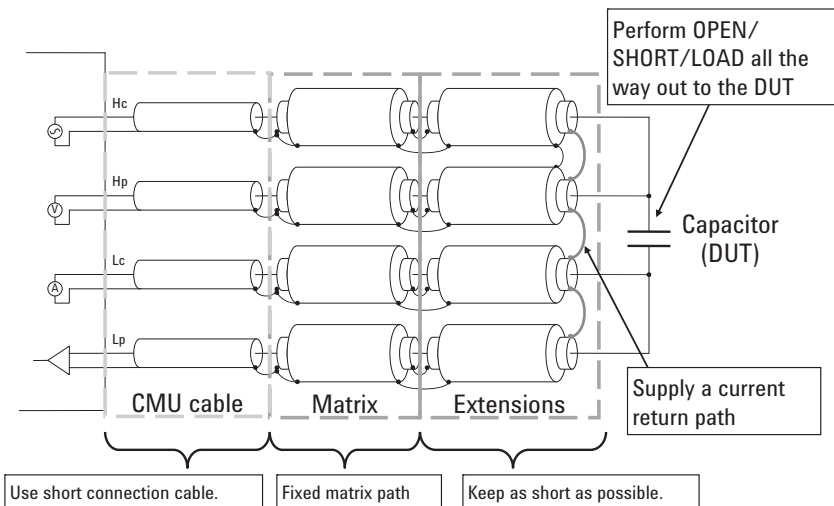*Figure 8.52. Follow this procedure if trying to go above 1 MHz when using a switching matrix.*

Even though you can have success making capacitance measurements above 1 MHz through a switching matrix, it needs to be emphasized that 5 MHz represents a hard upper limit on the achievable frequency. Above 5 MHz you should use direct connections in order to obtain satisfactory capacitance measurement results.

High DC bias capacitance measurements

In Chapter 4 we discussed making on-wafer high-voltage and high-current measurements. However, we did not have sufficient background on capacitance measurement theory to discuss high DC bias voltage capacitance measurement. The key to making these measurements is the high-voltage bias-T.

High-voltage bias-T
The high-voltage bias-T is designed to work with the MFCMU and HVSMU. A simplified circuit schematic of the high-voltage bias-T is shown below.
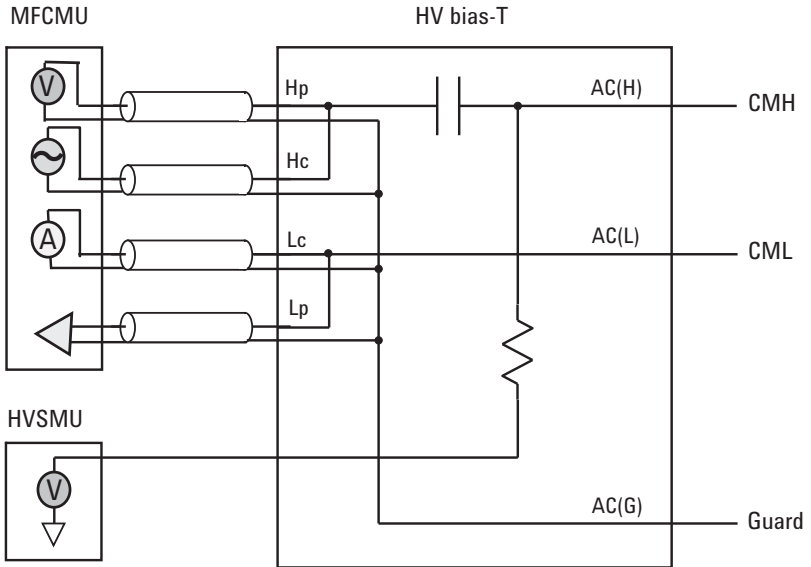
Figure 8.53. The high-voltage bias-T takes the outputs of the MFCMU and HVSMU and converts them into the familiar CMH and CML 4TP outputs.

The CMH and CML outputs of the N1260A high-voltage bias-T (wafer prober version) are SHV connectors. Note that the AC guard of the MFCMU is available, and in certain measurement situations it is necessary to connect the AC guard to one of the DUT terminals. However, to understand the use of the AC guard it is first necessary to review the basics of high-power MOSFET capacitance measurement.

High-Power MOSFET Capacitance Measurement

A conceptual diagram of a power n-channel MOSFET device is shown below.

Figure 8.54 – Power n-channel MOSFET diagram showing depletion regions and capacitance.

For power MOSFETs the drain is biased to a very high voltage, and both the drain-to-source capacitance (Cds) and the gate-to-source capacitance (Cgs) are dependent on the DC value of the drain voltage. The AC model of a MOSFET is shown below.
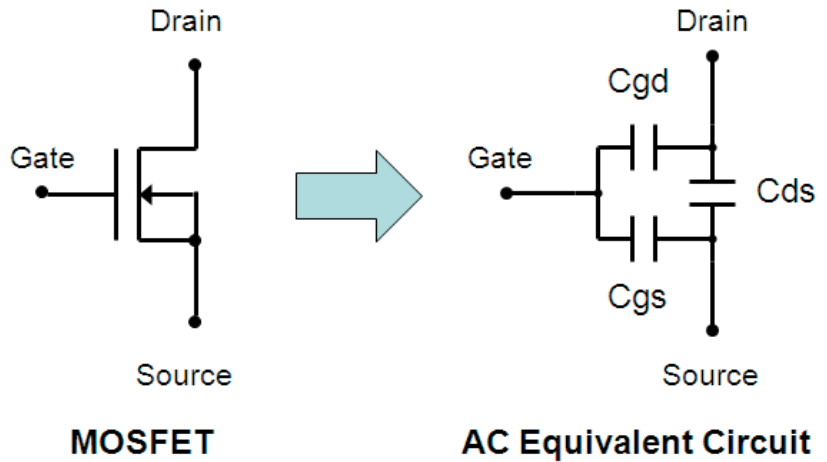
Figure 8.55 – AC model of a power MOSFET.

Let us now consider what happens when we try to measure any single one of these three capacitances using a capacitance meter. The following figure show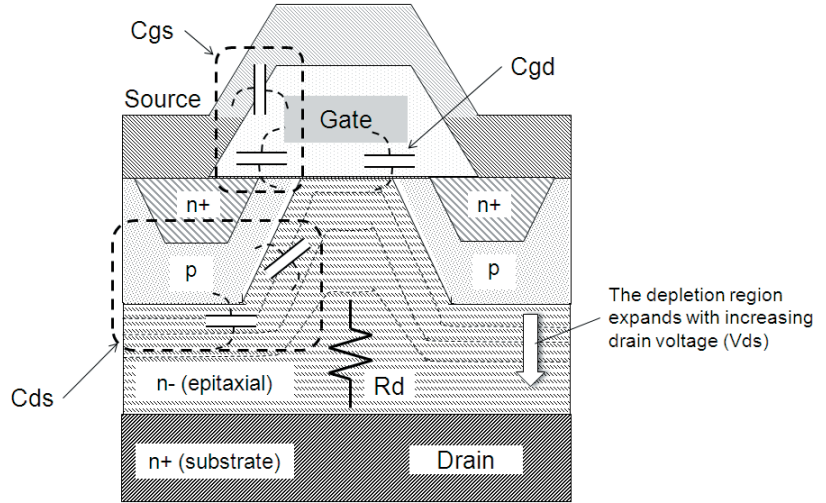s the situation when we try to measure an unknown capacitance (Cx) on a 3-terminal device without using the AC guard (i.e. unused terminal is floating).
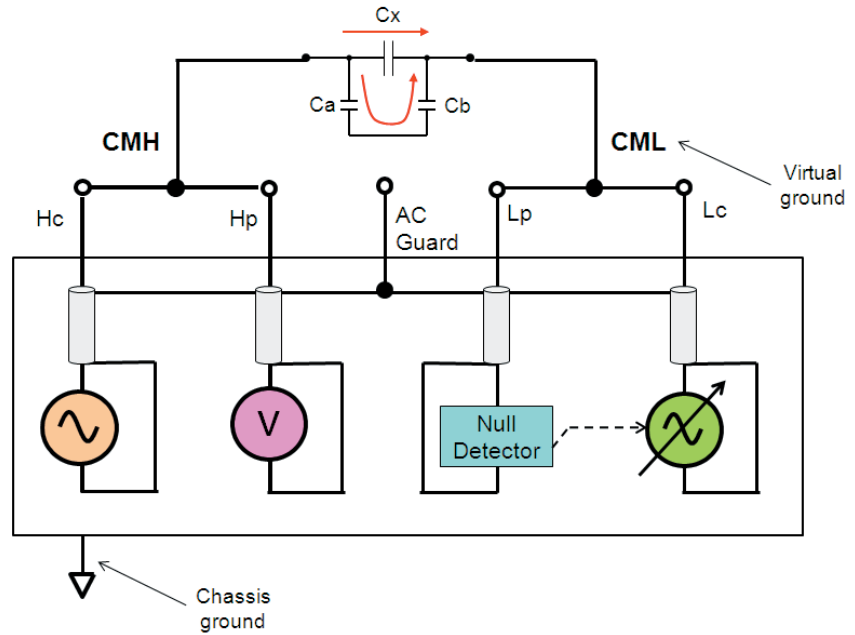
Figure 8.56 – Measuring capacitance on a 3-terminal device without using the AC guard.

This figure shows that when the unused terminal is floating current can flow through the other two capacitors resulting in erroneous measurement results. The best way to prevent this from occurring is to provide an alternative current path so that the current flowing through Ca does not flow back through Cb into the CML node. We can achieve this by connecting the unused terminal to the AC guard of the capacitance meter. Note: Before proceeding it is important to understand that the AC guard is the circuit common of the auto-balancing bridge and that it is connected to the shields of the four-terminal pair connectors. The AC guard is NOT the same as the ground terminal, which is connected to the chassis ground.

The next figure shows the benefit of connecting the AC guard to the unused measurement terminal.
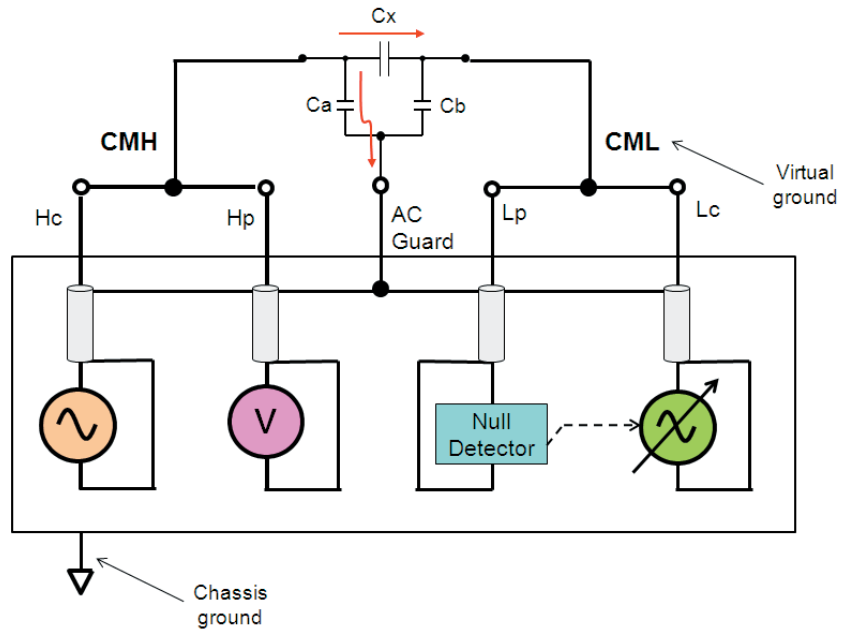
Figure 8.57 – Measuring capacitance on a 3-terminal device using the AC guard.

When the AC guard is connected to the third terminal the current flowing through the parasitic path (Ca) does not affect the accuracy of the measurement of the unknown capacitance (Cx), since the capacitance measurement is done through the CML node. Of course, this scheme assumes that the impedance of the AC guard node is much less than that of the parasitic path (Cb). Although this discussion did not include the use of the HVSMU and HV bias-T, we will explain how to incorporate them as we examine each capacitance measurement individually.

The following figure shows the correct way to measure Cgd using the HV bias-T.
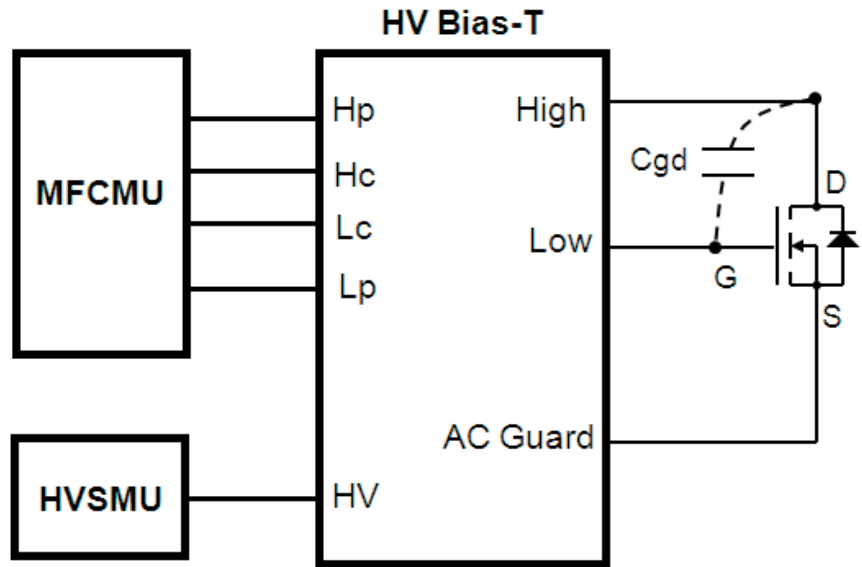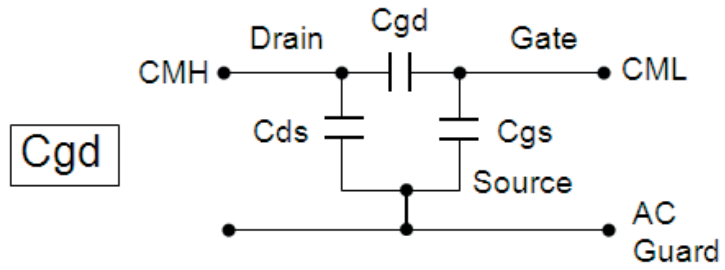
Figure 8.58 – Measuring the gate-to-drain (Cgd) capacitance of a MOSFET using the high-voltage bias-T and AC guard.

As can be seen, the measurement of Cgd using the HV bias-T is straightforward. Similarly, the measurement of Cds is also easy using the HV bias-T as shown in the following figure.
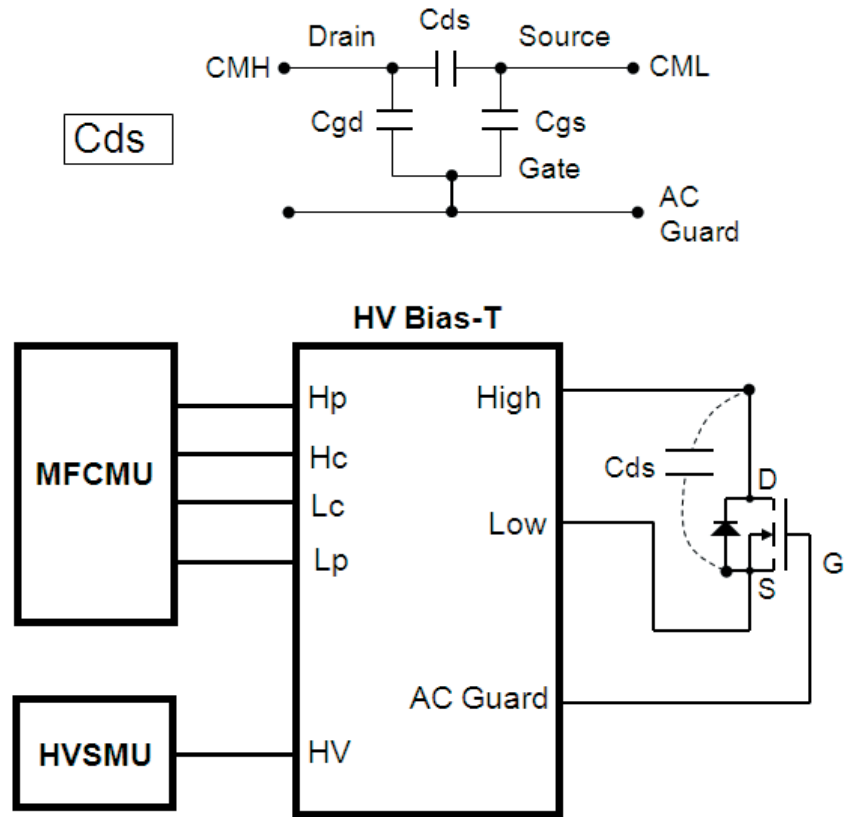
Figure 8.59 – Measuring the drain-to-source (Cds) capacitance of a MOSFET using the high-voltage bias-T and AC guard.

Measuring the gate-to-source capacitance (Cgs) presents some challenges that are not present for the other two measurements. Although we need to short the AC guard to the drain, we also need to bias the drain to high-voltage. The solution for this is to connect the drain and AC guard via a large capacitor (much larger than Cgd or Cds) such that the impedance seen by the drain with respect to the AC guard is much smaller than the impedance that it sees to either the source or to the gate. Conversely, we need to connect the HVSMU up to the drain through a relatively large resistor to prevent the HVSMU from interfering with the AC signal coming from the MFCMU. The following figure illustrates the proper way to make this measurement.
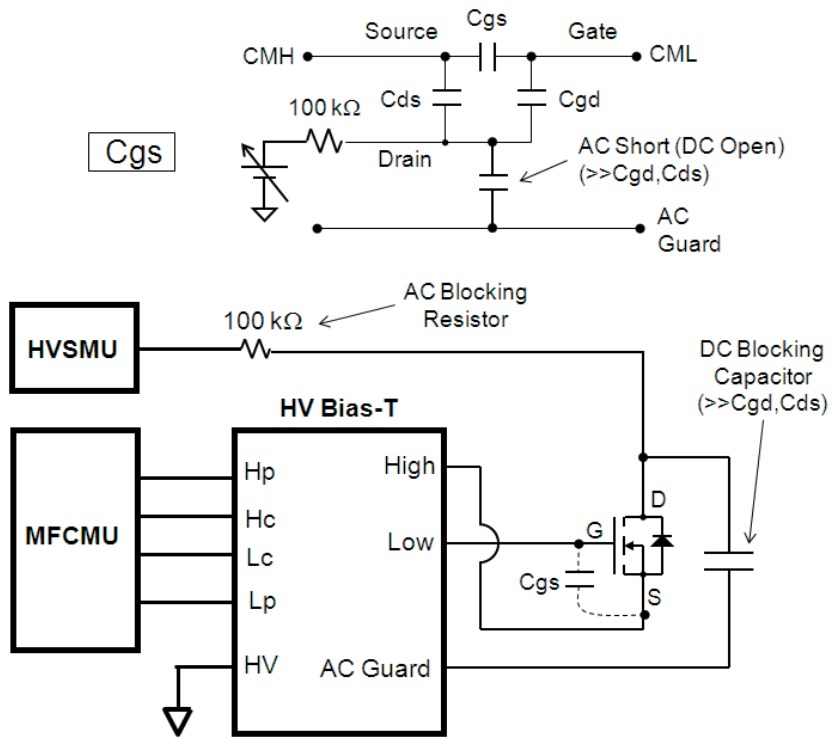
Figure 8.60 – Measuring the gate-to-source (Cgs) capacitance of a MOSFET using the high-voltage bias-T and AC guard.

Note: Although in this case the high-voltage bias-T is not actually required (and in fact is being by-passed), the connections are shown assuming that it is being used. This is simply due to the fact that the Cgd and Cgs measurements do require the high-voltage bias-T, so it is assumed that the user already has it in-place.

Power MOSFET data sheets specify MOSFET capacitance differently. The typical parameters specified are output capacitance (Coss), input capacitance (Ciss), and reverse transfer capacitance (Crss). These parameters can be calculated from Cds, Cgs, and Cgd using the equations shown below.

$$C_{oss} = C_{gd} + C_{ds}$$ *(Equation 8.15)*

$$C_{iss} = C_{gs} + C_{gd}$$ *(Equation 8.16)*

$$C_{rss} = C_{gd}$$ *(Equation 8.17)*

We have already discussed how to measure Crss (Cgd) previously. Ciss and Coss can also be measured directly using the B1505A's HVSMU, MFCMU and high-voltage bias-T, although the connections are different than for the cases of Cgd, Cds and Cgs. Since Coss is a simpler measurement to make we will discuss it first. To measure Coss we simply need to short the gate and source terminals using a wire as shown in the following figure.
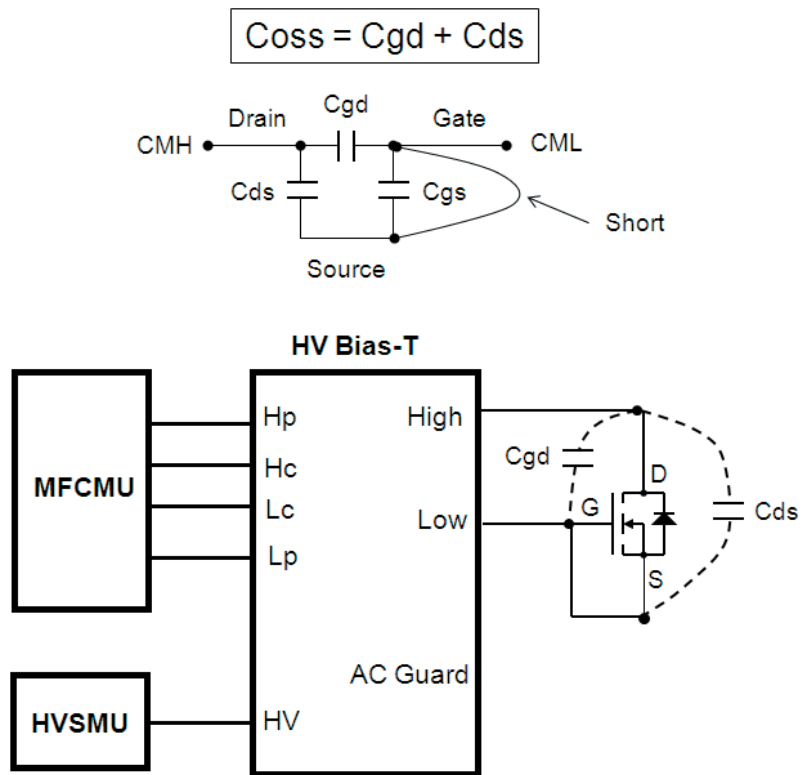
Figure 8.61 – Measuring Coss using the B1505A's HVSMU, MFCMU and high-voltage bias-T.

Note that when making this measurement we do not need to use the AC guard since we want to measure the current flowing through both Cgd and Cds. The following plot shows a direct capacitance measurement of Coss on a high power MOSFET made using the B1505A.

193
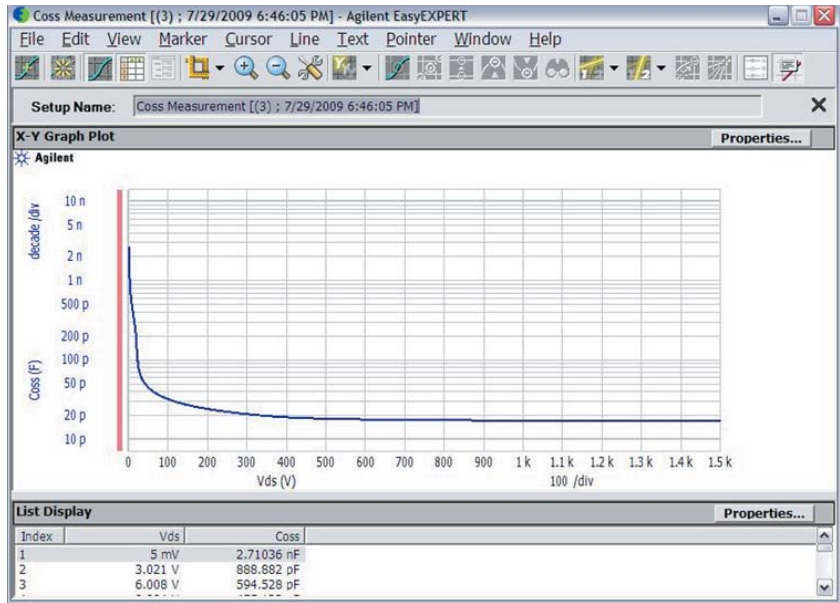
Figure 8.62 – Output capacitance (Coss) measured at 1500 V of DC bias.

The measurement of Ciss is more challenging because we need to connect the CMH terminal to the drain and also bias the drain to high-voltage. The solution is similar to the case of Cgs measurement in that we need to create an AC connection to the CMH terminal of the MFCMU and a DC connection to the HVSMU. The following figure illustrates this technique.
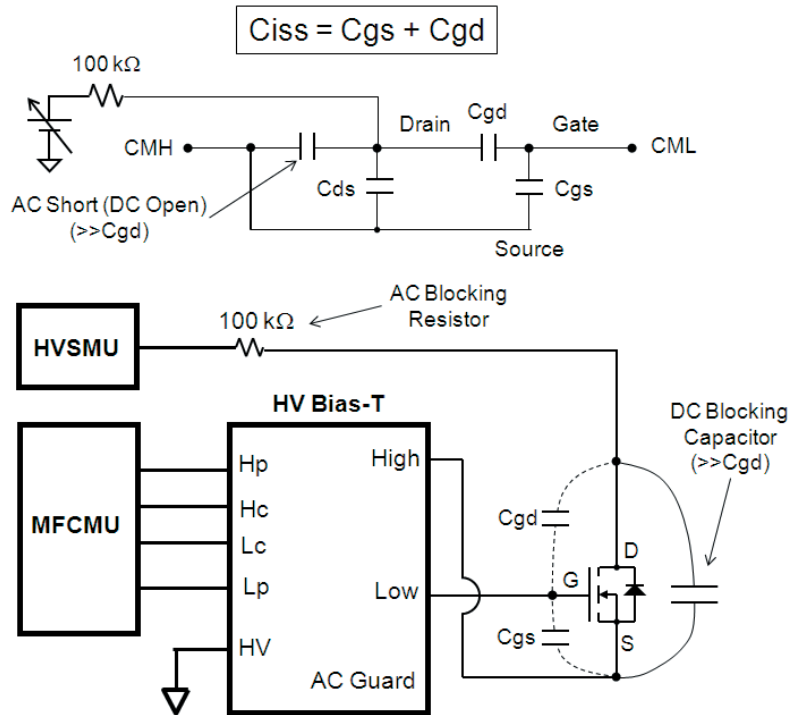
Figure 8.63 – Measuring Ciss using the B1505A's HVSMU, MFCMU and high-voltage bias-T.

A large capacitor is placed between the CMH and drain terminals to short out the drain-to-source capacitance (Cds) to AC signals. This DC blocking capacitor has to be much larger than the gate-to-drain capacitance (Cgd) so that the effective impedance between the CMH and CML terminals consists only of the parallel combination of Cgs and Cgd. As in the case of the Cgs measurement previously discussed, the HVSMU must be connected to the drain node through a relatively large resistor to prevent it from interfering with the AC capacitance measurement. The following is an illustration showing these connections for a packaged MOSFET using the N1259A test fixture.
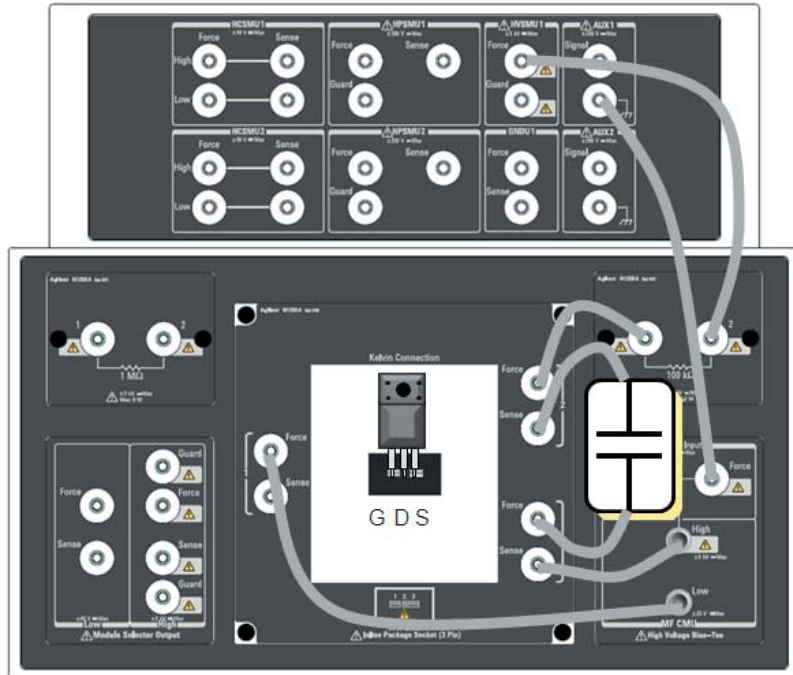
Figure 8.64 – Measuring Ciss on a MOSFET using the N1259A test fixture.

The following plot shows a direct capacitance measurement of Ciss on a high power MOSFET made using the B1505A.
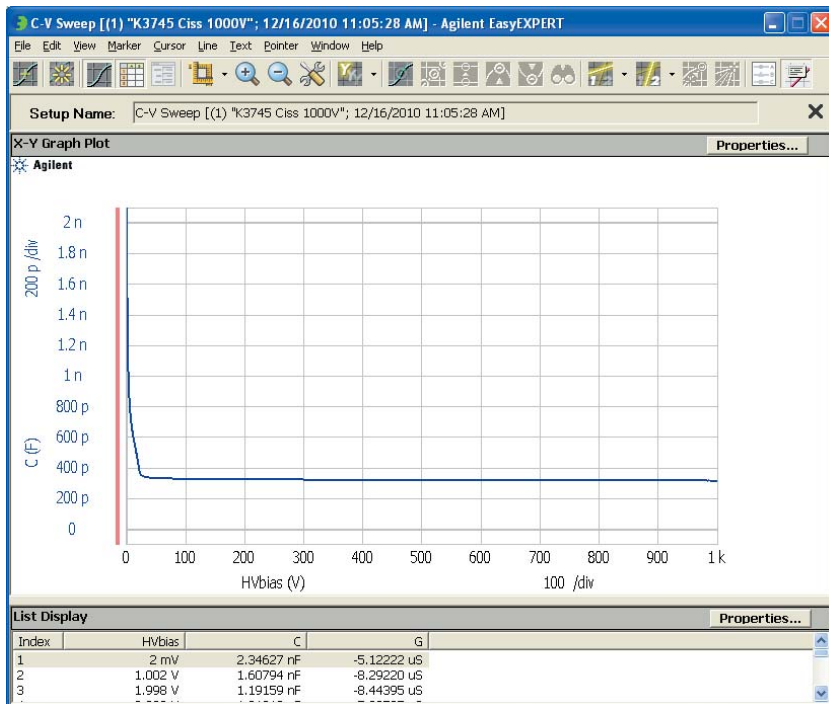
Figure 8.65 – Input capacitance (Ciss) measured at 1000 V of DC bias.

For low-power capacitance measurements load compensation generally only needs to be performed for measurements above 5 MHz. However, when using the high-voltage bias-T and making high-power capacitance measurements, load compensation needs to be performed at much lower frequencies to insure accurate measurement results. Most data sheets specify the Ciss, Coss and Crss at a frequency of 1 MHz. However, the following data shows a plot of gate to source (Cgs) capacitance (Cp-G) versus frequency using the high-voltage bias-T after performing open/short capacitance compensation. Note that the measured conductance becomes negative as the frequency increases beyond 100 kHz.
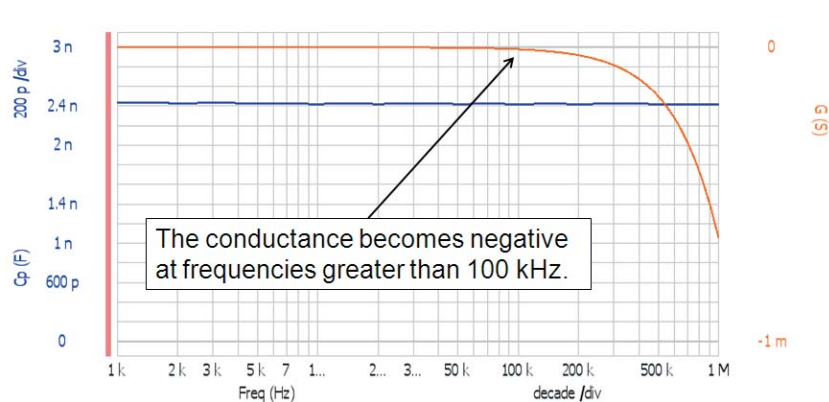
Figure 8.66 – Plot of Cp-G versus frequency using the high-voltage bias-T.

While this effect can be eliminated by performing a load capacitance compensation, it is not very practical to do a load compensation in most high-power device wafer probing environments. Therefore, if the load compensation cannot be performed it is best to measure the MOSFET capacitance parameters at frequencies no greater than 100 kHz. The following table summarizes the relative measurement accuracy of a Cp-G measurement on a power MOSFET as a function of compensation performed for the both 100 kHz and 1 MHz (when using the high-voltage bias-T).

	100 kHz		1 MHz	
	Cp	G	Cp	G
Open	Small error (1%)	OK	Large error	Large error
Open/Short	OK	OK	OK	Large error
Open/Short/Load	OK	OK	OK	OK

Figure 8.67 – Table comparing relative accuracy of Cp-G measurements at 100 kHz and 1 MHz for different levels of capacitance compensation when using the high-voltage bias-T.

Note: This table is based on the relatively large junction capacitances (on the order of nanofarads) that are typical of power MOSFETs.

Using the techniques just explained it is possible to directly measure power MOSFET capacitances at voltages up to 3000V. This represents considerable improvement over the conventional methods to measure these same parameters, which typically consist of homemade test setups that involve measuring a device's step response and extracting the value of capacitance from the resulting RC time constant.

High DC bias voltage wafer prober connections

There are generally two different types of devices on which we will want to perform CV measurements when using a wafer prober: lateral and vertical. On lateral device the connections are a little simpler since we can use an HPSMU to bias the wafer chuck (substrate connection). The two most common capacitance measurements are drain-to-source (Cds) and gate-to-drain (Cgd). In terms of the cabling and connectors needed to make these measurements there are no differences; however, the connections have to be manually changed as the following two figures illustrate. The following two figures illustrate the connections for lateral devices.

Figure 8.68. A B1505A configuration to perform gate-to-drain capacitance (Cgd) measurement at high DC voltage bias on a lateral device.

Figure 8.69. A B1505A configuration to perform drain-to-source capacitance (Cds) measurement at high DC voltage bias on a lateral device.

For vertical devices we need to energize the wafer chuck to high voltages. To do this you need an SHV to high-voltage triaxial adapter, and this adapter will have to be supplied by either your wafer prober vendor or by some other third party that makes adapters of this type. As in the case of the lateral device, the cabling and connectors are the same for the two different measurements (Cgd and Cds) but the connections need to be manually changed. The following two figures illustrate the connections for these cases.

Figure 8.70. A B1505A configuration to perform gate-to-drain capacitance (Cgd) measurement at high DC voltage bias on a lateral device.
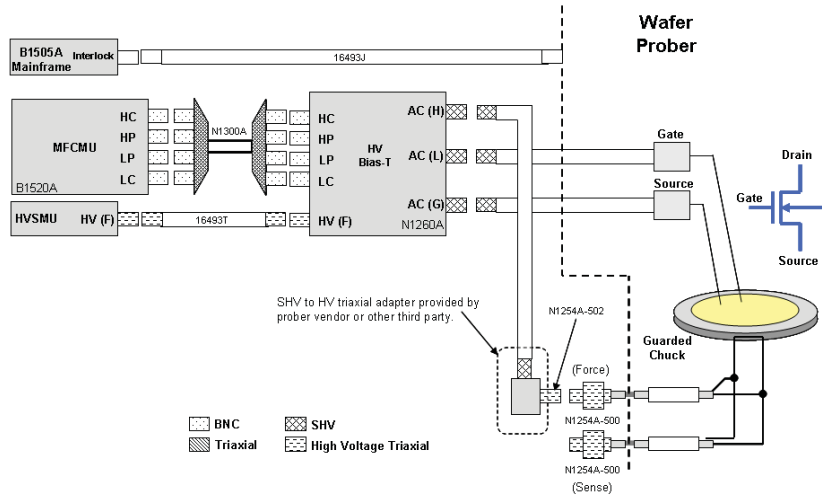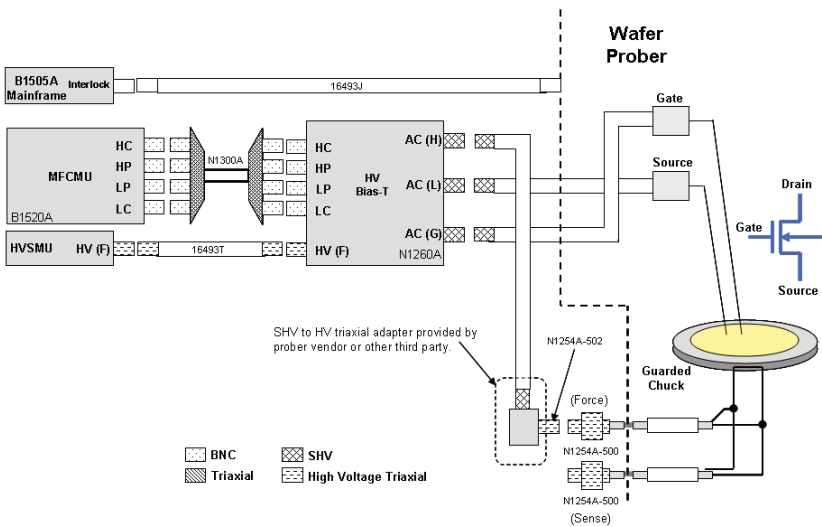

Figure 8.71. A B1505A configuration to perform drain-to-source capacitance (Cds) measurement at high DC voltage bias on a vertical device.

As you can see, it is important to understand exactly what types of devices you will be measuring when configuring a wafer prober for high DC bias voltage capacitance measurements.

Appendix A: Agilent Technologies' Parametric Measurement Solutions

The following parametric measurement instruments are available from Agilent Technologies. Only some of the key measurement capabilities are listed. For more information, please refer to the technical data sheet or web page for each instrument.

Semiconductor parameter/device analyzer

Agilent B1500A Semiconductor Device Analyzer

· Modular configuration: 10 slots
· MS Windows-based instrument with EasyEXPERT software
· Furnished with > 230 application tests
· Current measurement resolution: 0.1 fA
· Voltage measurement resolution: 0.5 μV
· ±200 V and ±1 A HPSMU available
· 1 kHz to 5 MHz capacitance module available
· ±40 V PGU module available (with ALWG capability)
· Waveform generator/fast measurement unit (with ALWG capability) supports fast IV and pulsed IV with 5 ns sampling rate

Power device analyzers/curve tracers

Agilent B1505A Power Device Analyzer/Curve Tracer
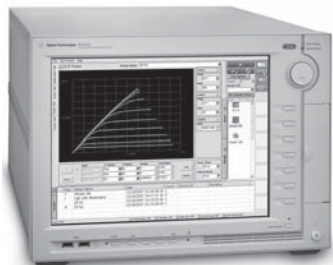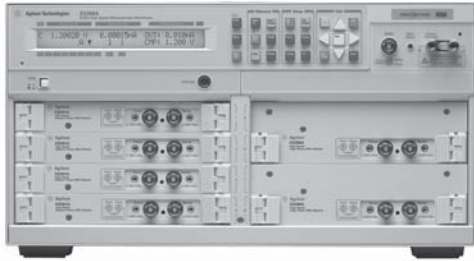

· Modular configuration: 10 slots
· MS Windows-based instrument with EasyEXPERT software
· Curve tracer mode with knob sweep support
· Current measurement resolution: 10 fA
· Voltage measurement resolution: 0.2 μV
· ±20 A high-current SMU available
· ±3000 V high-voltage SMU available
· 1 kHz to 5 MHz capacitance module available with support for CV measurements at 3000 V DC bias

Modular source/monitor units

Agilent E5260A 8-Slot High Speed Measurement Mainframe

- · Modular configuration: 8 slots
- · Basic user interface for spot measurements and program debug
- · Code compatible with the 4142B
- · SMUs measure several times faster than 4142B SMUs
- · Output up to 4 A (4 × HPSMUs) at once
- · 8 programmable trigger in/out lines in addition to BNC trigger in/out connectors
- · 4.0 A ground unit

Agilent E5262A 2-channel (medium power, medium power) Source Monitor Unit

- · Fixed configuration: 2 × MPSMUs
- · Basic user interface for spot measurements and program debug
- · Code compatible with the 4142B
- · SMUs measure several times faster than 4142B SMUs
- · 8 programmable trigger in/out lines in addition to BNC trigger in/out connectors
- · 2.2 A ground unit

Agilent E5263A 2-channel (medium power, high power) Source Monitor Unit

- · Fixed configuration: 1 × MPSMU, 1 × HPSMU
- · Basic user interface for spot measurements and program debug
- · Code compatible with the 4142B
- · SMUs measure several times faster than 4142B SMUs
- · 8 programmable trigger in/out lines in addition to BNC trigger in/out connectors
- · 2.2 A ground unit

Agilent E5270B 8-slot Precision Measurement Mainframe

- · Modular configuration: 8 slots
- · Basic user interface for spot measurements and program debug
- · Code compatible with the 4142B
- · Supports HRSMU with 1 fA and 0.5 μV measurement resolution
- · HRSMU supports ASU for 0.1 fA measurement resolution
- · Output up to 4 A (4 × HPSMUs) at once
- · 8 programmable trigger in/out lines in addition to BNC trigger in/out connectors
- · 4.0 A ground unit
- · TIS commands supported for both Basic and C so that algorithms can be transported to the 4070 or 4080

Benchtop source/measure units

Agilent B2901A Precision Source Measure Unit – 100 fA 1ch

Agilent B2902A Precision Source Measure Unit – 100 fA 2ch

· Single and dual SMU channel instruments
· 4.3" LCD-based GUI display
· Set/measure range: ±210 V, ±3 A (DC), ±10.5 A (Pulsed)
· Setting resolution: 1 pA & 1 μV
· Measurement resolution: 100 fA & 100 nV
· Supports 4-wire (Kelvin) measurements
· USB port for storing data, graphs and measurement setups via Flash drive
· Supports web-browser control (fully compliant with LXI class C specifications)
· Fully controllable from a PC via GPIB, LAN or USB ports

Agilent B2911A Precision Source Measure Unit – 10 fA 1ch

Agilent B2912A Precision Source Measure Unit – 10 fA 2ch

· Single and dual SMU channel instruments
· 4.3" LCD-based GUI display
· Set/measure range: ±210 V, ±3 A (DC), ±10.5 A (Pulsed)
· Setting resolution: 10 fA & 100 nV
· Measurement resolution: 10 fA & 100 nV
· Supports 4-wire (Kelvin) measurements
· USB port for storing data, graphs and measurement setups via Flash drive
· Supports web-browser control (fully compliant with LXI class C specifications)
· Fully controllable from a PC via GPIB, LAN or USB ports

Switching matrices

Agilent E5250A Low Leakage Switch

· Triaxial inputs: 6
· Coaxial inputs: 4
· Internal paths: 6
· Low-current paths: 2
· Maximum output ports: 48
· Effective current measurement resolution: 20 fA
· Bandwidth: 10 MHz

Agilent B2200A Femtoamp Low Leakage Switch
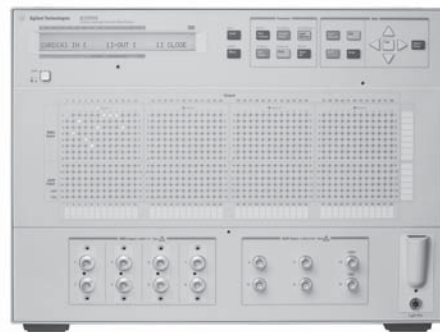
· Triaxial inputs: 8
· Coaxial inputs: 6
· Internal paths: 14
· Low-current paths: 8
· Maximum output ports: 48
· Effective current measurement resolution: 1 fA
· Bandwidth: 30 MHz

Agilent B2201A 14-channel Low Leakage Switch
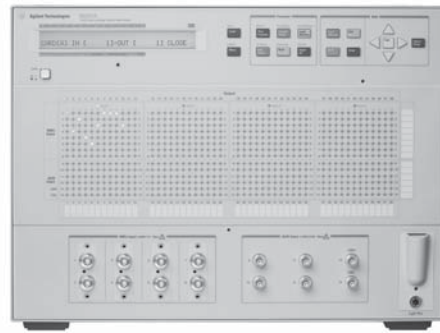
· Triaxial inputs: 8
· Coaxial inputs: 6
· Internal paths: 14
· Low-current paths: 8
· Maximum output ports: 48
· Effective current measurement resolution: 10 fA
· Bandwidth: 30 MHz

Impedance analyzers
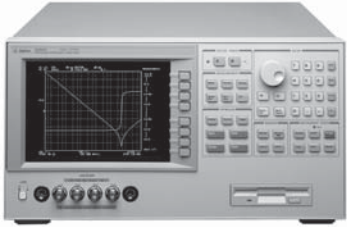
Agilent 4294A Precision Impedance Analyzer
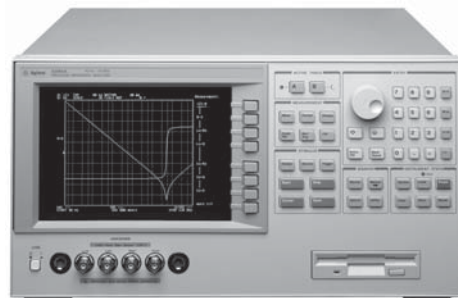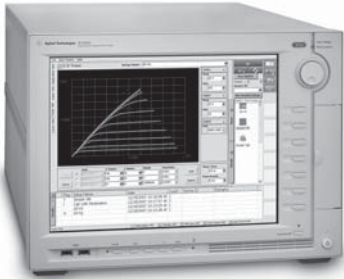
· Frequency range: 40 Hz to 110 MHz
· Test voltage signal range: 5 mV to 1 V (rms)
· Test current signal range: 200 μA to 20 mA (rms)
· DC voltage bias range: 0 V to ±40 V
· DC current bias range: 0 mA to ±100 mA

Appendix B: Agilent On-Wafer Capacitance Measurement Solutions

Agilent Technologies can supply an integrated suite of capacitance measurement solutions that span the range from quasi-static CV to high-frequency CV (110 MHz). In all cases, Agilent EasyEXPERT software provides an easy-to-use user interface that both simplifies the capacitance measurement process and makes it easy to export the data into PC-based work environments for analysis and reporting. A summary of these solutions for low DC bias (100 V or less) is shown below.

Agilent B1500A Semiconductor Device Analyzer		*Agilent 4294A Precision Impedance Analyzer*
QSCV (SMUs)	1 kHz to 5 MHz (MFCMU)	Up to 110 MHz
Agilent EasyEXPERT software		

In addition, the Agilent B1505A Power Device Analyzer/Curve Tracer can measure capacitance with up to 3000 V of DC bias.

Agilent B1505A Power Device Analyzer/ Curve Tracer	*Agilent N1260A High-Voltage Bias-T*
1 kHz to 5 MHz at up to 3000 V DC bias	
Agilent EasyEXPERT software	

Appendix C: Application Note Reference

The following is a listing of Agilent Technologies application and product notes relevant to parametric test.

Application/ product note	Title	Publication number
PN 4156-1	Agilent 4155B/4156B New Semiconductor Parameter Analyzers – Differences from Agilent 4155A/4156A	5965-9846E
PN 4156-2	Agilent 4155B/4156B New Semiconductor Parameter Analyzers – Configuring an NFS and Print Server for Network Capabilities	5966-4184E
PN 4156-3	Agilent 4155B/4156B Semiconductor Parameter Analyzers – Prober Connection Guide	5966-4185E
PN 4156-4	Agilent 4155B/4156B Semiconductor Parameter Analyzer – Advantages over the Agilent 4145A/4145B	5968-6681E
AN 4156-1	Ultra Low Current dc Characterization of MOSFETs at the Wafer Level	5963-2014E
AN 4156-2	Automated Extraction of Semiconductor Parameters using the Agilent 4155C/4156C	5963-1249E
AN 4156-3	Evaluation of Hot Carrier Induced Degradation of MOSFET Devices	5963-1111E
AN 4156-4	Evaluation of Flash Memory Cells	5965-5657E
AN 4156-5	Measurement of Power Devices Using External DC Power Supply	5963-2365E
AN 4156-6	Optimizing the Incoming Inspection of Semiconductor Devices	5963-2364E
AN 4156-7	Evaluation of Electromigration Using the SWEAT Procedure	5963-1110E
AN 4156-8	Evaluation of Oxide Reliability Using V-Ramp and J-Ramp Tests	5963-1248E
AN 4156-9	Evaluation of the Surface State Using Charge Pumping Methods	5964-2195E
AN 4156-10	Evaluation of Gate Oxides Using a Voltage Step Quasi-Static CV Method	5988-1025EN
AN 4156-11	Precision Measurement of Metal Line Width in Sub-quarter Micron Interconnect Systems	5988-2695EN
AN E5250A-1	Low Current Measurement with Agilent E5250A Switch Mainframe	5964-9112E
AN E5250A-2	Evaluation of Hot Carrier Induced Degradation of MOSFET Devices	5964-9113E
AN E5250A-3	Accurate and Efficient C-V Measurements	5965-5658E
AN B1500A-1	Measuring CNT FETs and CNT SETs Using the Agilent B1500A	5989-2842EN
AN B1500A-2	Migrating from the Agilent 4155C and 4156C to the Agilent B1500A	5989-2851EN
AN B1500A-3	IV and CV Measurement Using the Agilent B1500A MFCMU and SCUU	5989-3608EN
AN B1500A-4	Customizing Agilent B1500A EasyEXPERT Application Tests	5989-5167EN

Appendix C: Application Note Reference *continued...*

Application/ product note	Title	Publication number
AN B1500A-5	Creating a Test Sequence Using Agilent B1500A EasyEXPERT Software	5989-5524EN
AN B1500A-6	Accurate NBTI Characterization Using Timing-on-the-fly Sampling Mode	5989-5711EN
AN B1500A-7	Using the Agilent B1500A with a Nanoprober to Perform Failure Analysis	5989-5927EN
AN B1500A-8	Multi-Channel Parallel Timing-on-the-fly NBTI Characterization Using Agilent B1500A	5989-6526EN
AN B1500A-9	Improving Flash Memory Cell Characterization Using the Agilent B1500A	5989-8654EN
AN B1500A-10	Ultra-Fast 1 ms NBTI Characterization Using the Agilent B1500A's WGFMU Module	5989-9963EN
AN B1500A - 11	Characterizing Random Noise in CMOS Image Sensors	5990-3705EN
AN B1500A - 12	1 ms IV Characterization of Flash Memory Cells Using the Agilent B1530A	5990-3636EN
AN B1500A - 13	Measuring Pulsed/Transient Electrical Properties of OTFTs	5990-3983EN
AN B1500A - 14	IV and CV Characterization of Solar/Photovoltaic Cells Using B1500A	5990-4428EN
AN B1500A -15	Ultra Low Current DC MOSFET Characterization at the Wafer Level	5990-5547EN
AN B1500A -16	The B1500A: A Complete Solution for CMOS Device Electrical Characterization	5990-5566EN
AN B1500A -17	A Complete CMOS Reliability Test Solution	5990-5896EN
SE B1500A - 1	Agilent Pulsed IV Parametric Test Solution Selection Guide	5990-3672EN
AN B1505A - 1	Accurate and Efficient Characterization of Power Devices at 3000 V/20 A	5990-3749EN
AN B1505A - 2	Curve Tracer Mode Supports Simple and Easy Failure Analysis	5990-3813EN
AN B1505A - 3	Precision Power Device Evaluation at 40 Amps	5990-5432EN
AN-B1505A - 4	Direct Power MOSFET Capacitance Measurement at 3000 V	5990-7145EN
PN B1505A - 1	Creating Custom Socket Modules for Packaged Power Devices	5990-4654EN
AN 4294-3	Agilent Evaluation of MOS Capacitor Oxide C-V Characteristics Using the Agilent 4294A	5988-5102EN
AN 1369-3	Agilent Accurate Impedance Measurement with Cascade Microtech Probe System	5988-3279EN

Appendix C: Application Note Reference *continued...*

Application/ product note	Title	Publication number
AN B2900 - 1	Diode Evaluation Using the Agilent B2900A Series	5990-5671EN
AN B2900 - 2	Resistance Measurements Using the Agilent B2900A Series	5990-6633EN
AN B2900 - 3	Field Effect Transistor Characterization Using the Agilent B2900A Series	5990-6587EN
AN B2900 - 4	Diode Production Test Using the Agilent B2900A Series	5990-6993EN
AN B2900 - 5	IV Characterizations of Photovoltaic Cells Using the Agilent B2900A Series	5990-6660EN
AN B2900 - 6	Fast fT-Ic Measurement Using the Agilent B2900A Series	5990-6895EN
AN B2900 - 7	IV characterization of OLEDs Using the Agilent B2900A Series	5990-6994EN
AN B2900 - 8	Thermistor Evaluation Using the Agilent B2900A Series	5990-7118EN
AN B2900 - 9	Characterization of Bipolar Transistors Using the Agilent B2900A Series	5990-6995EN
AN B2900 - 10	Thermistor Production Test Using the Agilent B2900A Series	5990-7117EN
AN B2900 - 11	LED IV Measurement Using the Agilent B2900A Series	5990-6559EN
AN B2900 - 12	Varistor Evaluation Using the Agilent B2900A Series	5990-7116EN
AN B2900 - 13	LIV Test of Laser Diode Using the Agilent B2900A Series	5990-7115EN
AN B2900 - 14	Varistor Production Test Using the Agilent B2900A Series	5990-7114EN
AN B2900 - 15	Resistor Production Test Using the Agilent B2900A Series	5990-7141EN
AN B2900 - 16	LED Production Test Using the Agilent B2900A Series	5990-7140EN

Glossary

Accuracy – The degree of conformity of a measured or calculated quantity to its actual (true) value.

ALWG – The abbreviation of <u>A</u>rbitrary <u>L</u>inear <u>W</u>aveform <u>G</u>eneration. ALWG is a waveform generation technique in which a waveform is synthesized from a concatenated set of user-defined line segments.

Analog-to-digital converter (ADC) – A circuit used to convert an analog signal into the digital equivalent of that signal.

Atto-sense and switch unit (ASU) – A special module that works with a high-resolution SMU to obtain current measurement resolution of 0.1 fA. The ASU also has a switching capability that allows it to switch between its SMU connection and other instrumentation connected to it via coaxial inputs.

Auto ranging – An instrument measurement feature whereby the measurement circuitry will automatically switch between measurement ranges until the optimal measurement range for the quantity under test is reached.

Bipolar – An alternative transistor technology to MOS, bipolar transistors differ from MOS transistors in that they are current-controlled devices and they rely upon minority carriers for their operation. Bipolar transistors are most often used in RF and high-power applications.

BNC cable – See coaxial cable.

Capacitance – The amount of electric charge that can be stored for a given electric potential. The most common capacitor type is the two-plate capacitor, where the formula for capacitance is given by $C = Q/V$.

CDA – The abbreviation of <u>C</u>lean <u>D</u>ry <u>A</u>ir. This is air that has been filtered and dehumidified to remove particulates and moisture so that it can be safely used in pneumatic devices and systems.

CMU – The abbreviation of <u>C</u>apacitance <u>M</u>easurement <u>U</u>nit. This is a very generic term covering any sort of hardware that can perform capacitance measurement. It could mean a benchtop instrument such as the E4980A or 4294A, or a plug-in module such as the B1500A's MFCMU.

Coaxial cable – Also known as a "BNC" cable. A coaxial cable consists of an inner conductor surrounded by a tubular insulating layer typically made from a flexible material with a high dielectric constant, all of which is then surrounded by another conductive layer, and then finally covered again with a thin insulating layer on the outside.

Common mode rejection ratio (CMRR) – The ability of a circuit to reject input signals common to both input leads. CMRR is measured in positive decibels, and is defined as 20 times the log in base 10 of the ratio of the differential gain divided by the common-mode gain.

Common mode voltage – The voltage difference between the internal reference of an instrument and true earth ground.

Debye length – In semiconductor physics the solution of Poisson's equation in the presence of free charge always leads to a characteristic Debye length, which provides a qualitative measure of the extent of the space charge region.

Depletion approximation – In semiconductor physics, the assumption that a semiconductor junction can be divided into distinct regions which are either completely neutral or completely depleted of mobile carriers

Dew Point – The dew point is the temperature at which a given volume of air must be cooled (at constant barometric pressure) for water vapor to condense into water.

Drift – A relatively slow change in the measurement reading of an instrument over time that occurs even when the measurement environment and the quantity being measured are stable.

DUT – The abbreviation of Device Under Test. DUT is a very general term that can be used for simple components such as resistors, capacitors and transistors up to and including entire integrated circuits (ICs).

Faraday cage – An enclosure formed of conductive material (or a mesh of such material) to shield the inside of the enclosure from the effects of the surrounding electromagnetic fields.

Fixed ranging – An instrument measurement feature in which the measurement range (as the name implies) never changes from the specified value. Note: If the quantity being measured exceeds the maximum value supported by the selected measurement range then the measurement will fail and an error message will be generated.

Flat band capacitance – The capacitance of a MOS capacitor under flat band voltage conditions.

Flat band voltage – A non-thermal equilibrium condition in a MOS capacitor in which the voltage applied at the gate is set to a value that exactly compensates the difference in the work functions of the gate material and the semiconductor. In this condition, the energy bands in the silicon are flat in the surface regions as well as in the bulk (hence the name flat band voltage).

Ground loop – An undesirable measurement condition in which two or more conductive planes or surfaces are connected to ground through multiple points, thereby allowing current to flow between them due to common mode voltage differences.

Ground unit – A special type of SMU that has no measurement capabilities and is always set at zero volts.

GSWU – The abbreviation of Guard Switch Unit. The GSWU is used with the SCUU to automatically short the triaxial guards together during capacitance measurements and stabilize the series inductance.

Guarding – The technique of surrounding a signal line with an actively driven conductor maintained at the same voltage potential as that of the signal to eliminate leakage currents.

Impedance – The total opposition a device or circuit offers to the flow of an alternating current (AC) at a given frequency. It is represented as a complex quantity that can be graphically represented on a vector plane where the x-axis represents the real part (resistance, R) and the y-axis represents the imaginary part (reactance, X).

Input impedance – The effective impedance seen at the inputs to an electronic circuit.

Kelvin measurement – Also known as a "4-wire" measurement, the Kelvin measurement technique uses four terminals to make a resistor measurement: two to force current and two to measure voltage. By using this technique the effects of cable resistance on the measurement can be eliminated.

Limited auto ranging – An instrument measurement feature similar to auto ranging except that the measurement circuitry will stop going into any lower measurement ranges once the specified range limit has been reached.

Measurement range – The instrument measurement state (usually specified in decade increments) in which the measurement circuitry is internally configured to produce the optimum measurement result for a given range of values.

MOS – The abbreviation of Metal Oxide Semiconductor. The most common type of transistor structure, MOS transistors consist of some sort of gate material (metal or polysilicon) separated from a conducting (semiconductor) channel by an insulating material (usually silicon dioxide). MOS transistors are either n-channel type (NMOS) or p-channel type (PMOS), and circuits containing both NMOS and PMOS transistors are known as complimentary MOS (CMOS) technologies.

Multi-frequency capacitance measurement Unit (MFCMU) – A module for the B1500A and B1505A that can measure capacitance and that has a frequency range from 1 kHz to 5 MHz.

National institute of standards and technology (NIST) – A non-regulatory agency of the United States Department of Commerce, NIST supplies standard reference materials of the highest quality and metrological value. All Agilent instruments are calibrated to NIST traceable standards.

Noise (electronic) – The random variations in current and voltage that occur in an electronic circuit. There are many different types of noise that can occur in electronic circuits.

PLC – The abbreviation of Power Line Cycle. As the name implies, this is the length of time it takes the AC power supply to complete one cycle (20 ms for 50 Hz systems and 16.7 ms for 60 Hz systems).

Pulse generator unit (PGU) – An instrument or instrument module capable of generating extremely fast voltage and/or current pulses (in the nanosecond range or below).

Output impedance – The effective impedance seen at the outputs of an electronic circuit.

Repeatability – The degree to which repeated measurements or calculations show the same or similar results.

Resolution – The lowest resolvable quantity of data that an instrument can accurately measure.

SCFM – The abbreviation of Standard Cubic Feet per Minute. In the English system, this is the volumetric flow rate of a gas corrected to standardized conditions of temperature, pressure and relative humidity. Note: Be careful as the "standardized" conditions can vary between definitions. In the SI metric system, the term Normal Cubic Meter (Nm^3) is used.

SHV – The abbreviation of Safe High Voltage. The SHV is a type of connector used with BNC cables for safe connection to high voltage sources. The connector uses a bayonet mount similar to those of BNC connectors, but it is easily distinguished due to its very thick and protruding insulator. The insulation makes SHV connectors safer for handling high voltages by preventing accidental contact with the live conductor in an unmated connector or plug. The connector is also designed such that when it is being disconnected from a plug the high voltage contact is broken before the ground contact, thereby preventing accidental shocks.

Single point grounding – The preferred technique to avoid ground loops, whereby all ground planes are each connected to a common ground through a single point.

SMU CMU unify unit (SCUU) – A switching module designed to work with the B1500A's MFCMU and two of either the high-resolution or medium power SMUs. The SCUU form-factor is designed to accept the four BNC outputs of the MFCMU and the two pairs of Kelvin triaxial outputs from the two HRSMUs and/or MPSMUs. The SCUU provides the capability to switch between the MFCMU and the HRSMU/MPSMU pair automatically, and its outputs are two pairs of triaxial outputs (force and sense).

Source/monitor unit (SMU) – Also sometimes designated as "Source/Measure Unit." The SMU is a measurement module that can force voltage or current and simultaneously measure voltage and/or current. Although often thought of as a DC measurement unit, the SMU does have some time sampling and pulsing capability.

Triaxial cable – A cable designed for ultra-low (<1 fA) current measurements. The triaxial cable consists of an inner conductor surrounded by a tubular insulating layer typically made from a flexible material with a high dielectric constant, which is surrounded by a conducting guard layer that is actively driven to be at the same potential as the center conductor. The guard layer is then surrounded by another conductive layer (the shield), and covered again with a thin insulating layer on the outside.

Waveform generator/fast measurement unit (WGFMU) – The WGFMU is a measurement module that can generate arbitrary voltage waveforms (with 10 ns programmable resolution) and measure both current and voltage extremely fast (5 ns sampling rates).

Index

About the Author

Alan Wadsworth is the North and South American Market Development Manager for Agilent's semiconductor test division. Alan holds bachelors and masters degrees in electrical engineering from the Massachusetts Institute of Technology and an MBA from Santa Clara University. Alan has over 25 years of experience in the semiconductor industry in both design and test.

Agilent Email Updates

www.agilent.com/find/emailupdates
Get the latest information on the products and applications you select.

LXI

www.lxistandard.org
LAN eXtensions for Instruments puts the power of Ethernet and the Web inside your test systems. Agilent is a founding member of the LXI consortium.

Agilent Channel Partners

www.agilent.com/find/channelpartners
Get the best of both worlds: Agilent's measurement expertise and product breadth, combined with channel partner convenience.

Microsoft is a U.S. registered trademark of Microsoft Corporation.

AttoGuard and AttoFast are U.S. registered trademarks of Cascade Microtech, Inc.

Agilent
Advantage
Services

Agilent Advantage Services is committed to your success throughout your equipment's lifetime. To keep you competitive, we continually invest in tools and processes that speed up calibration and repair and reduce your cost of ownership. You can also use Infoline Web Services to manage equipment and services more effectively. By sharing our measurement and service expertise, we help you create the products that change our world.

www.agilent.com/find/advantageservices

Agilent Electronic Measurement Group
DEKRA Certified
ISO 9001:2008
Quality Management System
www.agilent.com/quality

www.agilent.com
www.agilent.com/find/parametric

For more information on Agilent Technologies' products, applications or services, please contact your local Agilent office. The complete list is available at:

www.agilent.com/find/contactus

Americas

Canada	(877) 894 4414
Brazil	(11) 4197 3600
Mexico	01800 5064 800
United States	(800) 829 4444

Asia Pacific

Australia	1 800 629 485
China	800 810 0189
Hong Kong	800 938 693
India	1 800 112 929
Japan	0120 (421) 345
Korea	080 769 0800
Malaysia	1 800 888 848
Singapore	1 800 375 8100
Taiwan	0800 047 866
Other AP Countries	(65) 375 8100

Europe & Middle East

Belgium	32 (0) 2 404 93 40
Denmark	45 45 80 12 15
Finland	358 (0) 10 855 2100
France	0825 010 700*
	*0.125 €/minute
Germany	49 (0) 7031 464 6333
Ireland	1890 924 204
Israel	972-3-9288-504/544
Italy	39 02 92 60 8484
Netherlands	31 (0) 20 547 2111
Spain	34 (91) 631 3300
Sweden	0200-88 22 55
United Kingdom	44 (0) 118 927 6201

For other unlisted countries:
www.agilent.com/find/contactus
Revised: January 6, 2012

Product specifications and descriptions in this document subject to change without notice.

© Agilent Technologies, Inc. 2012
Published in USA, May 3, 2012
5990-5278EN

Agilent Technologies