Vrije Universiteit Brussel

Vrije Universiteit Brussel
Faculteit Ingenieurswetenschappen
Vakgroep ELEC
Pleinlaan 2, B-1050 Brussels, Belgium

Contributions to Large-Signal Network Analysis

Proefschrift ingediend tot het behalen van de academische graad van doctor in de ingenieurswetenschappen

Frans Verbeyst

September 2006

Promotor: Prof. Dr. ir. Yves Rolain

Vrije Universiteit Brussel

Vrije Universiteit Brussel
Faculteit Ingenieurswetenschappen
Vakgroep ELEC
Pleinlaan 2, B-1050 Brussels, Belgium

Contributions to Large-Signal Network Analysis

Proefschrift ingediend tot het behalen van de academische graad van doctor in de ingenieurswetenschappen

Frans Verbeyst

Voorzitter:
Prof. Dr. ir. Jacques De Ruyck (Vrije Universiteit Brussel)

Vice-voorzitter:
Prof. Dr. ir. Jean Vereecken (Vrije Universiteit Brussel)

Promotor:
Prof. Dr. ir. Yves Rolain (Vrije Universiteit Brussel)

Secretaris:
Prof. Dr. ir. Alain Barel (Vrije Universiteit Brussel)

Jury:
Prof. Dr. ir. Don DeGroot (CCNi Measurement Services, Andrews University, Michigan, USA)
Prof. Dr. ir. Rik Pintelon (Vrije Universiteit Brussel)
Prof. Dr. ir. Roger Pollard (University of Leeds, UK)
Prof. Dr. ir. Johan Schoukens (Vrije Universiteit Brussel)
Prof. Dr. ir. Steve Vanlanduit (Vrije Universiteit Brussel)

What lies behind us
and what lies in front of us
is nothing compared to
what lies within us.

Table of contents

CHAPTER 5: System identification approach applied to drift estimation.

Preface

Although scattering S-parameters have been around longer than I have, today, some of us still manage to incompletely and therefore incorrectly define them as[1] $S_{ij} = \frac{b_i}{a_j}$.

The other problem with S-parameters is that they got so ingrained that many people believe that they are omnipotent when it comes to solving microwave problems. Don't get me wrong, S-parameter theory and the associated instrumentation has served and still serves the RF community extremely well. Use it to its full power ... without denying its basic assumptions: the superposition theorem, and therefore linear behaviour, must hold. S-parameters are all around, because they are technology-independent, because they can be measured and because they model the reflected and transmitted waves as function of the incident waves and, as such, can be used during simulations.

Although its community is steadily growing, "Large-Signal Network Analysis" is still in its infancy and I guess not too many people can give a sufficiently correct description of what LSNA stands for. Without claiming to hold the ultimate knowledge, LSNA has three major cornerstones. First, the device under test - referred to as the network - is put in (as) realistic (as possible) large-signal[2] operating conditions, not only with respect to input power levels, but also with respect to spectral content and mismatch conditions. Next, the behaviour of that DUT is completely and accurately characterized in order to be analyzed. Because the basic quantities (voltage and current) are measured, there is a natural link to make this data available in simulators. The data can be made available as is or through the use of a behavioural model. Finally, it is technology-independent, usable from the device up to the system level. It's the engine of a unified approach ... "beyond S-parameters". The behaviour of the device can be studied in the domain and in the format which is most convenient for the user. Some people prefer the time domain, other the frequency domain. Some prefer voltage and current, others travelling voltage waves.

This work contains humble contributions to different aspects of "Large-Signal Network Analysis", which started more than 10 years ago.

Accurate measurements require both reliable hardware and software. I am the last person on earth to claim that building reliable hardware at microwave frequencies is a piece of cake. However, the hardware doesn't have to be perfect. That's because there are clever persons conceiving clever calibration algorithms. The software which is used to control the hardware and to collect the raw data however must be perfect

1. b_i corresponds to the transmitted or reflected voltage wave at port i and a_j represents the incident voltage wave at port j
2. large-signal refers to the fact that the stimulus becomes significant compared to the operating range of the device

and must withstand the tooth of time. The chapter on “Software Architecture” shortly describes the basic principles used when designing and implementing the hardware abstraction layer of a Large-Signal Network Analyzer.

Hardware at RF and microwave frequencies is never perfect. Fortunately, this can be compensated for, using a set of proper calibration elements and ditto calibration techniques. Calibration of a Large-Signal Network Analyzer is somewhat more complex than the calibration of a classical Vector Network Analyzer (VNA). On top of a relative calibration, the LSNA also requires an absolute power and phase calibration. The power calibration is performed using a power sensor, which is traceable up to NIST. The phase calibration however requires a new calibration element, which is referred to as the Harmonic Phase Reference (HPR). The latter is a pulse generator, which itself must be calibrated. This is done using a high-frequency sampling oscilloscope. Unfortunately, this one isn’t perfect either and it seems to become a never-ending story. Luckily, the imperfections of the sampling oscilloscope can be compensated for using a “nose-to-nose” calibration technique. The basics, the individual imperfections, their estimation and compensation are described in detail in another PhD. In order to be really useful, additional work is required. The chapter “Enhancements to the nose-to-nose calibration technique” shortly describes the streamlined implementation of the calibration technique and the replacement of some of the techniques by other techniques which were published and were proven to be superior either in quality or in speed. I’m convinced that this additional work has been essential in the adoption of the technique both by the people at NIST and by the calibration lab of Agilent Technologies in Santa Rosa.

The nose-to-nose calibration and its application as a part of the calibration of the Large-Signal Network Analyzer, fuelled the research at NIST related to their Electro-Optic Sampling (EOS) system. This system allows to characterize a photodiode up to 110 GHz. By measuring the (known) impulse response of this photodiode using a high-frequency sampling oscilloscope, an alternative method does exist to determine both the amplitude and the phase distortion introduced by this oscilloscope. The discrepancies that were reported by NIST are verified in the chapter “Comparison of the nose-to-nose and EOS-based calibration technique”. It is not the ambition of the chapter to find the reason of this discrepancy nor to eliminate it. Based on the work described below, some of the required processing is performed in a different way, after the initial processing. As part of this additional verification, an exact expression is found for the variance of different realizations of a sine wave in the presence of normally distributed jitter noise and additive noise. Implementations are realized both in the absence and the presence of time base distortion.

The above verification has been the trigger for some recent research with respect to jitter estimation. Existing literature demonstrates that jitter, which has a symmetrical probability density function, does not introduce any phase distortion. As such, jitter estimation was not given too much attention as a part of the nose-to-nose calibration technique. Because jitter does have a low-pass effect on the amplitude, it becomes important when verifying the discrepancy that was reported for the amplitude characteristic. The main motivation for additional research is the observation that the

generally accepted first order model to describe the sample variance of repeated measurements in the presence of both jitter noise and additive noise cannot explain some of the measurement results that are obtained during the nose-to-nose and the EOS-based calibration. Backed up by the system identification knowledge, which is available at the department, the existing model is extended and different estimators are implemented. Several years earlier, I presented some of my behavioural modelling work based on neural networks to people of the department and I was asked if the residual error was small or large. At that time, I didn't understand the rationale of that question and therefore I could not answer it. Fortunately, things have changed and I feel the urge to ask that question too, each time others tell me how well their model works. Anyway, the power of a stochastical framework is demonstrated once more when the excellent results that are obtained based on simulations are in strong contrast to the results based on measurements. It is found that the problem is caused by overlooking the effect of time base drift compensation on the sample variance. The solution that is found for that problem also properly deals with time base distortion. The results are more than satisfactory. The rather extended "System identification approach applied to jitter estimation" chapter describes the work in detail and differs from the previous chapters by providing an avalanche of uncertainty intervals.

During the review of the paper that describes the above research on jitter estimation, an enhanced version of the time base drift estimation has been proposed by one of the co-authors. This proposal and its implementation trigger new insights, especially when jitter is present. Study of the covariance matrix of the spectral noise in the frequency domain allows proper weighting of the contribution of the individual spectral components to the cost function. This not only provides a relevant value for the cost function, but it also reduces the uncertainty on the estimated drift in the presence of a realistic quantity of additive and jitter noise by a factor of 2. This work is described in the "System identification approach applied to drift estimation" chapter.

The last chapter reports on one of several contributions to another aspect of Large-Signal Network Analysis: closing the loop between measurements and simulations. Volterra-based behavioural modelling work has been performed in the early days. The idea of predistorting a base-band signal using an inverse Volterra model is believed to be original at that time and still today it seems to be alive and kicking. I vividly remember that I was asked to write a C program to generate all unique combinations of the spectral components at the inputs of a MIMO[1] system, given the degrees of nonlinearity of each output. The resulting model is referred to in literature as VIOMAP[2]. After performing some magic with pointers and recursion, I proudly presented benchmarks for an increasing number of frequency components and an increasing maximum degree of nonlinearity. History has taught me to be less ambitious. Nevertheless, original results are presented at the IMS 1995 conference[3].

1. Multiple-Input-Multiple-Output.
2. Volterra input-output map.
3. IMS stands for International Microwave Symposium

A revised copy of the unpublished IMS paper can be found in the “Volterra-based behavioural modelling” chapter. It demonstrates that one must be very careful when using a model which is extracted using one class of excitation signals and then used to predict the output of the system for another class of excitation signals. Unfortunately, there are still a lot of colleagues out there who need additional education. The work also demonstrates that one should not overlook the effect of the biasing circuitry. Other applications of the VIOMAP were targeted to bridge the gap between this Volterra-based technique and existing techniques. It is demonstrated that the VIOMAP can be used as an alternative to load-pull measurements. The VIOMAP is also a natural extension of S-parameters for weakly nonlinear devices. This statement is emphasized by extracting the VIOMAP for two different amplifiers and by predicting the behaviour of the cascade of both amplifiers by cascading their individual VIOMAPs.

Another application is the characterization at the fundamental frequency of the nonlinear behaviour of a device under test in a near 50 Ω environment. The combined idea of linearizing the behavioural model with respect to the incident wave at the output and the use of readily available components like a load, open, short, adapters and attenuators to synthesize different loads to extract such a model results in a US Patent Application Publication, No. US 2003/0057963 A1. This patent is filed as employee of Agilent Technologies.

Recently, some new alternatives have been published to existing fundamental source-pull and load-pull techniques. Also, fundamental-only measurement-based behavioural models are introduced using either active injection in combination with a manual load tuner or using an electro-mechanical load tuner. This work has been performed as employee of NMDG Engineering and is referenced in the publication list.

Acknowledgements

When I graduated in 1986, I was sure about one thing: I would never get involved in nonlinear stuff and especially not at microwave frequencies. At that time, both seemed much too complicated.

This PhD is yet another proof that there are no certainties in life. Fortunately, the latter is not one hundred percent true. As such, I first want to thank my wife Hilde for supporting me in what I do and for the sacrifices, which seemed to converge to a local maximum lately. It's been a bumpy ride and it hasn't been easy, but we made it !

Eli, thanks for being a great son, although there was much more mum than dad lately. This reminds me of being a son myself. I want to thank both my parents and my parents-in-law for giving me the opportunity to become what I am.

Next, I want to thank my promotor Yves Rolain, not only for his huge "open source" theoretical and practical knowledge, but also for his exceptional moral support. Thank you for being there as a scientist and as a friend.

This PhD would never have been started nor finalized without the support of the department ELEC. In particular, I want to thank Alain Barel, Rik Pintelon, Johan Schoukens and my promotor to provide the opportunity to "finally finalize" my PhD on a part-time basis at their department. I want to thank each of them for their enthusiasm during the many discussions we had.

The same holds for Marc Vanden Bossche, who hired me as initial team member of the NMDG group of Hewlett - Packard and who allowed me to trade my full-time job at NMDG Engineering for a part-time job until "I was done". I'm honoured to be part of his never-ending effort to make large-signal network analysis a success for each of us, meanwhile maintaining his attitude to live up to the spirit of Bill and Dave.

I had the privilege to work with several great people while I was with Hewlett - Packard (and later with Agilent Technologies). Doug Rytting was one of them and if there is one person, who is able to "promote" large-signal network analysis, it is surely him.

I want to say thanks to everyone at the department, including the technicians and the secretaries, to my "room mates" Wendy and Salua, and to the new generation of researchers for their "energy" and for accepting an "older bloke".

I want to thank all my co-authors too and all people worldwide whom I got in touch with one way or another along the challenging large-signal network analysis road. A special thanks to Tracy Clement at NIST for providing the measurement data that are used in the jitter and drift estimation chapters.

Finally, I want to apologize to all other people who specifically contributed to this work and who are not explicitly mentioned here.

Merchtem, 06/06/2006

Abbreviations

ADC A/D	analog-to-digital (converter)
API	application programming interface
ARFTG	Automatic Radio Frequency Techniques Group
AWG	arbitrary waveform generator
CLR	common language runtime
CW	continuous wave(form)
dB	decibel
dBm	decibels referenced to one milliwatt
DC	direct current
DCA	digital communication analyzer
DLL	dynamic-link library
DUT	device under test
EOS	electro-optic sampling
FFT	fast Fourier transform
GUI	graphical user interface
HP	Hewlett - Packard
HPR	harmonic phase reference (see also REFGEN)
i.i.d.	independent and identically distributed
IMS	International Microwave Symposium
IMTC	Instrumentation and Measurement Technology Conference
IQ	in phase and quadrature phase
ISA	industry standard architecture (bus)

IVI	interchangeable virtual instrument
KIS	keep it simple
LF	low frequency
LOST	load, open, short, thru
LS	least squares
LSNA	Large-Signal Network Analysis or Large-Signal Network Analyzer
MIMO	multiple input multiple output
MLE ML	maximum likelihood (estimator)
NIST	National Institute of Standards and Technology
NMDG	Network Measurement and Description Group
O/E	opto-electrical (converter)
PC104	(ISA-based) personal computer (bus) utilizing 104 pins
PCI	peripheral component interconnect (bus)
PDF	probability density function
PISPO	periodic in same period out
QAM	quadrature amplitude modulation
REFGEN	reference generator (see also HPR)
RF	radio frequency
RMS	root mean square
RTTI	run-time type information
SNR S/N	signal to noise (ratio)
SPMT	single-pole multiple throw
TBDn	time base distortion
var	variance

VIOMAP	Volterra input output map
VME	versa module eurocard
VNA	vector(ial) network analyzer
VXI	VME extensions for instrumentation
WLS	weighted least squares

CHAPTER 1 Software Architecture

Abstract

Characterization of the large-signal behaviour of a high-frequency active component under realistic conditions requires a measurement system, which is very versatile with respect to both the applied excitations and mismatch conditions.

Given the disruptive character of such measurement systems, both the underlying hardware and software must be kept as flexible as possible. A typical architecture of a *Large-Signal Network Analyzer* (LSNA) is shown in figure 1-1.

The device under test is connected to the LSNA hardware. The hardware abstraction layer allows the LSNA core software to communicate with this hardware in an abstract way. The design goal of this layer is to be able to replace part of the hardware by other hardware with similar capabilities, without impacting the core software. The latter mainly takes care of the data collection and the calibration of the LSNA. The measured data is provided in different domains and formats and can be both uncalibrated and calibrated. A user can interact with the system through a graphical user interface (GUI) and create applications running on top of the LSNA core functionality, using the LSNA application programming interface (API).

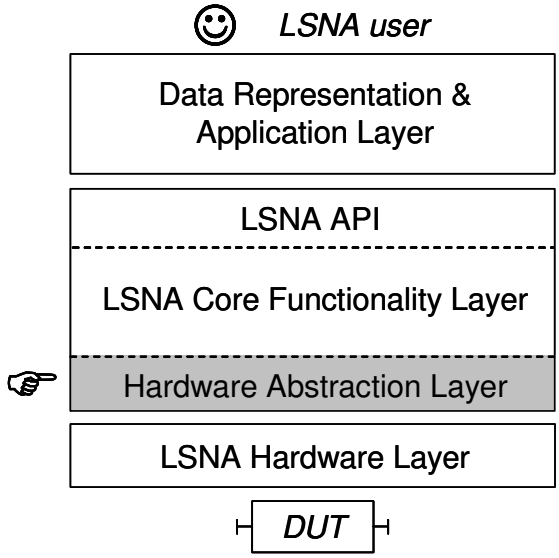

Figure 1-1. Typical architecture of a *Large-Signal Network Analyzer*.

This chapter describes the hardware abstraction layer, which must be both robust, fast and flexible. Based on these requirements, this layer is written in C++. Furthermore, the necessary foundations are added to provide robustness and flexibility. The initial implementation dates from the early 90's, before similar idioms and patterns were described in literature [1], [2]. Almost fifteen years later, the same software has survived the transition from VXI to both PC104 and PCI and provides the heartbeat of the *MT4463 Large-Signal Network Analyzer*, commercialized by Maury Microwaves and NMDG Engineering BVBA.

Introduction to object-oriented programming

First, fundamental concepts of object-oriented programming, such as classes, objects and abstraction are shortly introduced, based on an example: the LSNA test set.

The main functionality of an LSNA test set is to simplify its calibration by routing the signal, which is applied to its input, to one of its outputs. Meanwhile other outputs are terminated, typically into 50 Ohms.

The set of possible test set modes is defined, as described in figure 1-2.

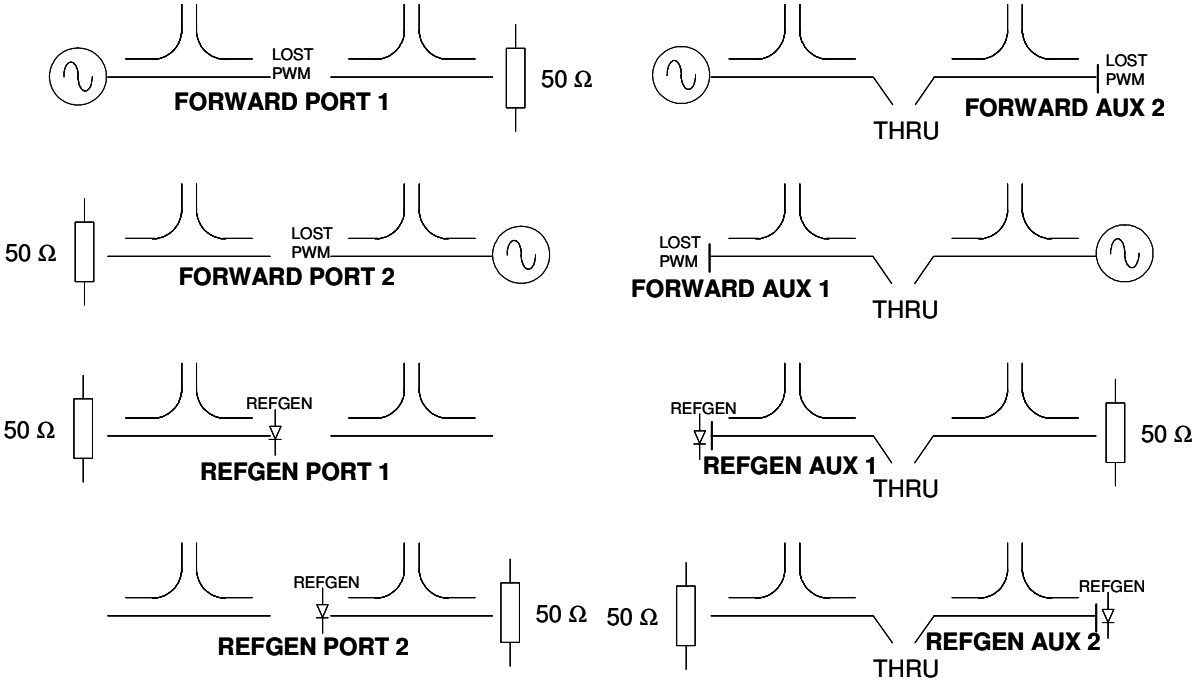

Figure 1-2. Example modes of a test set used as part of the *Large-Signal Network Analyzer*.

The name convention used for these modes answers the following simple question in a consistent way: “where is the calibration element connected?” or “what is the primary calibration port?”.

LOST refers to the connection of a Load, Open, Short or Thru calibration element. PWM corresponds to the connection of a power sensor to allow absolute amplitude calibration. REFGEN refers to the phase reference generator, which is in fact a pulse generator and used as part of the phase calibration of the LSNA.

A test set is able to indicate if it supports any of these modes, to put itself into the specified mode and to return its actual mode. How this is done is of no concern to the users of this test set. This in fact defines the abstract interface of a test set.

Classes describe the common characteristics and functionality of a group of similar objects. Test sets can be realized using different hardware, typically using a set of single-pole-multiple-throw (SPMT) switches and/or transfer switches. Each of these realizations typically is given a model number and represents a concrete class.

Objects correspond to unique instantiations of a class. Each individual test set, its uniqueness being represented by a unique serial number, corresponds to an object.

Patterns for increased robustness

Handles and smart pointers

In the C and C++ programming language, the programmer is responsible for the dynamic memory management. The proper amount of memory must be allocated at runtime, when the object is constructed. When the object is no longer needed, the associated dynamic memory must be freed. This must be done “just in time”.

If memory is released too soon, other objects may still hold a pointer to that memory. Such a pointer is referred to as a dangling pointer and results in undesired behaviour of the software and even system failure.

When the object is no longer needed and the programmer forgets to free the corresponding memory, this memory cannot be reused later on. This situation is referred to as memory leakage. Although memory is cheap these days and plenty of it is available, eventually the program will run out of memory and fail.

In order to avoid these problems and automate the release of memory, the C++ layer of the LSNA uses handles instead of pointers. Handles are objects which refer to smart pointers. The latter keep track of how many times the underlying dynamic memory is referenced. Because handles are objects, a programmer can rely on the C++ compiler and the correct implementation of the copy constructor, the assignment operator and the destructor of the Handle class. The required code is implemented once, either as a macro or as a template. If this code is properly implemented, one is assured that the underlying dynamic memory is freed only if, and as soon as, no one else references this memory.

Finally, the handles should be implemented such that they allow late binding. The latter is a powerful mechanism to be used in combination with abstraction, i.e. the actual implementation of the abstract interface is determined at run-time instead of at compile time.

Singletons

Sometimes it is important to make sure to have one and only one instance of a class. Typically this is the case for “manager” classes. The C++ layer of the LSNA uses two such classes, being the class manager and the handle manager. The former keeps track of all LSNA-related classes being available to the software, the latter doing the same for LSNA-related handles. It is essential for the robustness to make sure that all objects communicate with “the” class manager and “the” handle manager. According to [1], this is referred to as the Singleton pattern. The need for both a class and handle manager as part of the LSNA C++ software is described in the next paragraph.

Patterns for increased flexibility

Given the disruptive nature of the *Large-Signal Network Analyzer*, both hardware and software must be flexible. Over time the *Large-Signal Network Analyzer* has evolved from a VXI-based instrument to a PC104-based instrument. The original calibration module was replaced by a test set, while four one-channel VXI-based ADC cards were replaced by one four-channel PCI-based ADC card.

Below, several patterns are shortly described, providing the required flexibility at the hardware abstraction level, such that the same core LSNA software can be used on all existing systems.

For the remainder of this chapter, a handle to an object will be referred to as either a handle or an object.

Class and handle manager

Each software object, which is part of the LSNA, is referred to by a unique name. In general, each software object represents a hardware component. The functionality of these objects is implemented in classes and these too are referred to by a unique name.

The mapping of unique object names onto unique handles is taken care of by the handle manager. The same is done at the class level by the class manager.

All LSNA-related classes register themselves to the class manager. In the case of static libraries this is done automatically by applying the appropriate pattern. In the case of dynamic-link libraries (DLLs), one either needs to instantiate a dummy object of each required type in the main function of the application or one can force the necessary symbol references when linking the main application to the libraries.

Using names for both classes and objects, the actual hardware configuration can be stored in one or more configuration files. Each line of such a configuration file starts with the class name followed by the object name. At that moment both the class and handle manager come into play. First the handle manager verifies if an object with that name already exists. If this is the case and the object turns out to be of the correct type, a handle to that object is returned by the handle manager.

If the object does not yet exist, the handle manager uses the class manager to create an object, based on the specified class name. In fact each LSNA-related class implements two member functions which can be invoked by the class manager. The first function allows the interactive configuration of the object, while the other reads the remainder of the configuration of that object from file.

If the remainder of the configuration refers to other objects, the handle manager is used once more. When an object with the specified name already exists and if the object turns out to be of the correct type, the handle manager returns a handle to that object. If the object does not yet exist and interactive configuration is allowed, a dialog is initiated with the user through the console window to configure the missing object.

Finally, a mechanism is implemented allowing to find out class information at run-time. This allows to verify if an object is an instantiation of a specified class or an instantiation of a concrete subclass, in the case the specified class is an abstract class. It also allows to find all concrete subclasses of a specified class. Again, this is a powerful feature when correctly used in combination with abstraction. After the initial implementation as part of the LSNA, this mechanism was added to the C++ standard and is referred to as RTTI (Run-Time Type Information).

“Role” interface

Instrument drivers can be either very specific or rather abstract. Specific drivers have the advantage that they allow to exploit the full power of the instrument, while abstract drivers promote reusability and interchangeability. The hardware abstraction layer of the *Large-Signal Network Analyzer* contains both. However, only the abstract version of the driver is made available to the outside world. This interface is based upon the functionality required by the *Large-Signal Network Analyzer*.

Recently, this concept is described as part of IVI (Interchangeable Virtual Instruments) [3] and more specifically as part of the Measurement and Stimulus Subsystems Specification. The latter describes the concept of a Role Control Module, which maps an instrument interface on a “role” interface. This role interface corresponds to the required functionality when the instrument or hardware is used as part of the *Large-Signal Network Analyzer*.

In [1] this concept is described as the Adapter pattern.

The IVI driver architecture also specifies different operational modes. One of them is simulation mode, which allows to write software based on the instrument driver before the physical instrument itself is available. This concept was used as part of the hardware abstraction layer way before it was published by the VXI Consortium.

Delegation versus inheritance

C++ promotes code reuse through inheritance. However, inheritance tends to be overused, resulting in an explosion of the number of classes. An alternative to inheritance is delegation.

Proper combination of inheritance, delegation and abstraction is important when dealing with hardware which can be controlled in different ways. A good example is a step attenuator which can be controlled using either an Agilent E1339A Digital Output/

Relay Driver VXI interface card, an Agilent 11713A Attenuator/Switch Driver or dedicated PC104 hardware. It is necessary to separate the functionality of a step attenuator from the way it is controlled.

Typical functionality of a step attenuator is to realize a specified attenuation as faithfully as possible and to return the actually realized attenuation. In case the S-parameters are measured in a certain frequency range, the step attenuator can also return its S-parameter values for the actual attenuation as a function of frequency. This is a typical situation where subclassing is used: adding S-parameter capability to the basic functionality of a step attenuator.

With respect to the control part of the step attenuator, first an abstract “role” interface is defined. Step attenuators are typically controlled by activating and deactivating one or more sections. Each concrete implementation of this control API is written on top of the specific E1339A and 11713A driver and as such allows easy replacement of the control hardware.

Finally a handle to this abstract “role” interface is defined as a part of the step attenuator class, resulting in an orthogonal solution. The functionality of the step attenuator itself can grow by proper subclassing. Meanwhile the control portion is delegated through the handle. New control hardware can be added by proper implementation of the “role” interface.

Template Method pattern

Another pattern, which can be found in the hardware abstraction layer, is the Template Method pattern.

This design pattern is mainly used to define the skeleton of a complex algorithm in the base class and have the subclasses implement different versions for each part of the algorithm.

The abstract step attenuator driver base class does not know how to activate or deactivate a step attenuator section. However it keeps track of which section is activated, such that the concrete subclasses are relieved from this burden. As such the abstract driver class defines a public activateSection () member function which provides the template. This template correctly keeps track of the activated sections and leaves the physical activation or deactivation to the concrete subclass by calling the appropriate private virtual doActivateSection () member function. Defining a member function as private, makes sure that this function cannot be invoked from outside the abstract base class. Virtual functions are used in C++ to support the late binding, as explained on page 1-5.

Conclusion

We conclude with two well-known principles [4] in the object-oriented design world, which help to create more scalable, robust and reusable applications.

The open/close principle

Bertrand Meyer stated in 1996 that: “Software entities (classes, modules, etc.) should be open for extension, but closed for modification.”

In plain English this means that software modules should be designed such that their behaviour can be modified (open) without making source code modifications (closed), but by adding new code.

The Liskov substitution principle

Barbara Liskov stated 8 years earlier that: “Derived classes must be usable through the base class interface without the need for the user to know the difference.”

Both principles emphasis the proper usage of abstraction. Given the fact that the LSNA hardware abstraction layer software was conceived in the early nineties and is successfully used today in both legacy (VXI-based) and new (PC104/PCI-based) LSNA systems, without causing any frustration to both conceiver and users, it can be claimed that the used approach “simply works”.

References

[1] E. Gamma, R. Helm, R. Johnson, J. Vlissides, “Design Patterns. Elements of Reusable Object-Oriented Software,” *Addison-Wesley*, 1995.
[2] D. Box, “Essential COM,” *Addison-Wesley*, 1997.
[3] “IVI-3.10: Measurement and Stimulus Subsystem (IVI-MSS) Specification,” 2001, http://www.ivifoundation.org.
[4] R. Martin, “Agile Software Development Principles, Patterns, and Practices,” *Prentice Hall*, 2002.

CHAPTER 2

Enhancements to the nose-to-nose calibration technique.

Abstract

A *Large-Signal Network Analyzer* is conceived to analyze the behaviour of nonlinear devices. This implies that the superposition principle no longer holds for these devices. Even when applying a pure sine wave at the input of such a device and terminating it into 50 Ohm, the reflected and transmitted voltage wave will no longer be a pure sine wave. As such, in order to accurately measure incident and scattered voltage waves, not only a relative calibration like Short - Open - Load - Thru is required, but also a power and phase calibration. The power calibration is performed using a calibrated power sensor, while phase calibration is performed using a calibrated harmonic phase reference. The latter is a pulse generator which has a sufficiently rich harmonic content for a sufficiently broad range of fundamental frequencies. In order to use this pulse generator as an additional calibration standard, one needs to know the exact phase relationship between the different spectral components of the pulse at each fundamental frequency[1].

Therefore the harmonic phase reference is measured using a high-frequency sampling oscilloscope. Unfortunately this sampling oscilloscope also introduces distortions. These are mainly caused by the non-ideal time base and the non-ideal impulse response of the sampling oscilloscope, resulting in both amplitude and phase distortion. Nonlinear distortions are avoided by limiting the amplitude of the measured signals, while offset and gain errors are removed by performing a vertical calibration.

The list of distortions introduced by a high-frequency sampling oscilloscope were first studied and described separately in [1]. The characterization and compensation of these distortions are referred to as the “nose-to-nose” calibration technique. In order to be really useful, the problems and solutions described in [1] needed additional effort. The first contribution to the enhancement of the nose-to-nose calibration technique is the implementation of a streamlined process, such that the calibration can be performed in a repetitive way and, if necessary, by a technician. Furthermore, it allowed to share this procedure in detail with people from NIST for thorough crossverification. As a result, numerous articles were published by NIST related to the “nose-to-nose” calibration technique [2]-[7]. Finally, the procedure was transferred into the calibration lab of Agilent Technologies at Santa Rosa in order to allow the phase calibration of the Agilent 86030A 50 GHz *Lightwave Component Analyzer*, which gave it a unique competitive advantage and contributed significantly to its success. The second contribution is the replacement of techniques described in [1] by other techniques, which are superior either in quality or in speed.

Starting from a nose-to-nose calibrated oscilloscope, one can characterize either the harmonic phase reference to be used as part of the *Large-Signal Network Analyzer* or the opto-electrical subsystem of the *Lightwave Component Analyzer*.

1. This fundamental frequency corresponds to the repetition frequency of the pulse.

Streamlined implementation

In order to be able to perform nose-to-nose calibrations on a regular basis and in order to allow a trained technician to perform these calibrations, a streamlined process is implemented in *VEEtest™* from Agilent Technologies. The full calibration process typically takes 4 hours.

First, a time base distortion measurement of the scope is performed. This requires the measurement of sine waves at two or more non-harmonically related frequencies. At each frequency, the sine wave must be measured twice, where ideally the second measurement has a phase shift of 90° with respect to the first measurement. Typically, this is taken care of by using two channels and adding a delay line in the second channel, resulting in a phase shift of approximately 90° at discrete frequencies. The usage of different frequencies allows the estimation method [8] to discriminate between a harmonic due to nonlinear behaviour and due to time base distortion. The detection of the time base distortion is insensitive when the slope of the applied signal is small. Using a sine wave, the slope is minimal in the extrema and maximal in the zero crossings. Applying a 90° phase shift, the delayed sine wave has a maximal slope whenever the original sine wave has a minimal slope and vice versa. Once these measurements are performed, an estimate for the time base distortion and its uncertainty is obtained. Figure 2-1 shows the estimated time base distortion, defined as $TBDn[i] = t_i - i \cdot \Delta t$, the non-equidistant time stamps being represented by t_i; i corresponds to the time sample index and Δt represents the assumed constant equivalent-time sample rate. For an ideal time base, all values $TBDn[i]$ equal zero.

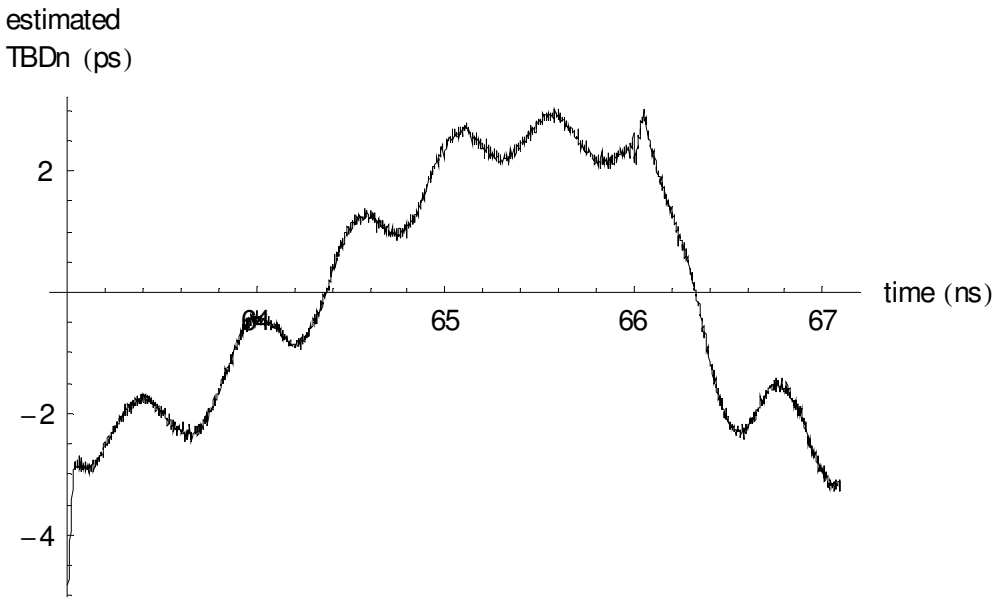

Figure 2-1. Estimated time base distortion as part of the nose-to-nose calibration.

Acquiring 128 dual-channel traces of 2048 points at 12.4, 13.6 and 14.8 GHz, using

[8], the 95% confidence interval on the estimated time base distortion is approximately 0.075 ps.

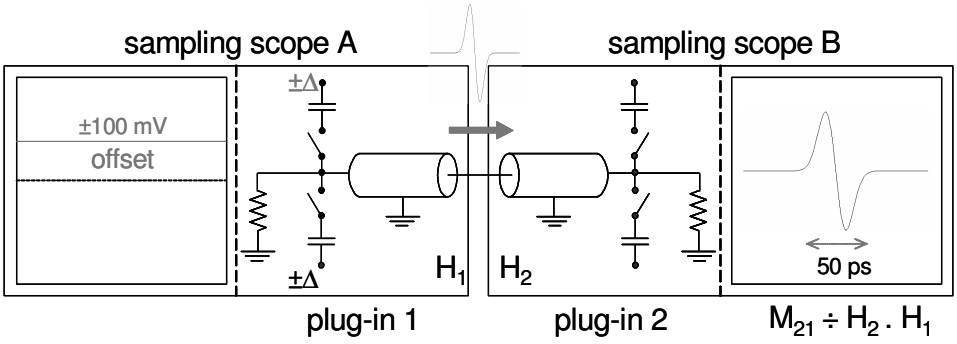

Figure 2-2. The “nose-to-nose” setup.

Next, the combined impulse response of two sampling oscilloscopes is measured using a “nose-to-nose” setup (figure 2-2). To do so, the inputs of two sampling oscilloscopes are connected to each other. A DC offset is applied to one oscilloscope. Whenever its samplers are closed, a pulse is fired from the internal samplers towards the input connector of the oscilloscope. This pulse is referred to as the kickout pulse. The nose-to-nose calibration technique is based on the assumption that the kickout pulse is proportional to the impulse response of the sampling oscilloscope and is described in detail in [1]. This kickout is then measured by the second sampling oscilloscope. The measured pulse is the convolution of the impulse response of both oscilloscopes. By measuring 3 sampling oscilloscope combinations, it is possible to retrieve the impulse response of each contributing sampling oscilloscope (Eq. 2-1). These combinations are referred to as M_{ij} measurements, where i represents the kickout-receiving oscilloscope and j represents the kickout-generating oscilloscope.

$$\begin{aligned} M_{12} &\div H_1 \cdot H_2 \\ M_{13} &\div H_1 \cdot H_3 \\ M_{23} &\div H_2 \cdot H_3 \end{aligned} \quad \Rightarrow \quad \sqrt{\frac{M_{12} \cdot M_{13}}{M_{23}}} \div H_1 \qquad \text{Eq. 2-1}$$

In order to estimate and reduce the uncertainty on each M_{ij} measurement, proper averaging is required. Since it is possible that the time base drifts, this phenomenon must be estimated and compensated before averaging. The drift is mainly caused by temperature variations in combination with air flow. During this step of the procedure, the time base drift is estimated and compensated before averaging. Also, based on the mean value and the variance of the pulse in the time domain, an estimate is given for the jitter standard deviation. More recent work on jitter estimation is described in a separate chapter: “System identification approach applied to jitter estimation”.

A portion of the strobe pulse, which fires the samplers, leaks through towards the output (the input connector). In order to eliminate this common mode portion of the kickout, two measurements are performed. First a positive DC offset and then a negative DC offset is applied to the kickout-generating sampling oscilloscope. The

inversion of polarity will cause a polarity reverse for the kickout pulse, while it will leave the strobe contribution untouched. Hence, after averaging both the positive and negative kickout and after proper alignment, one can subtract both kickouts to remove the common mode contribution of the strobe pulse.

After these three M_{ij} measurements, a second time base distortion measurement is performed to verify if the time base distortion has not changed during these M_{ij} measurements.

Finally, postprocessing is performed to correct the M_{ij} data, based on the estimated time base distortion [9]. In order to provide the amplitude and phase distortion of the sampling oscilloscope on a specified frequency grid, a chirp-z transform [10] is applied. Using a proper combination of the corrected and interpolated M_{ij} data, by taking the mismatch of each oscilloscope and the required adapter into account and after correcting for the low-pass effect of the jitter on the amplitude characteristic, one ends up with an estimate of the amplitude and phase characteristic of the three oscilloscope plug-ins separately.

This procedure and its implementation was shared with - and explained to - people from the Optoelectronics Division, the Statistical Engineering Division and the Radio-Frequency Technology Division within NIST. The procedure was re-implemented independently at NIST. No anomalies were found and the above procedure was described in detail [2].

Enhancement of the different parts

The implementation was performed using an Agilent 83480 Digital Communications Analyzer (DCA) and three dual-channel 50 GHz 83484A plug-ins. Later, the code was adapted to support also the Agilent 86100 DCA Oscilloscope. Both instruments use equivalent-time sampling.

Time base drift estimation

The best way to get rid of the time base drift is not to have it at all. Hence, one should always try to minimize sudden temperature variations and uncontrolled air flow when performing measurements using a high-frequency sampling oscilloscope.

The logarithmic averaging of the spectral data proposed and described in [1] has the disadvantage that a bias on the amplitude estimation is introduced which becomes significant when the S/N ratio approximates 0 dB[1]. Also, logarithmic averaging does not decrease the noise floor. This turned out to be a limitation during the measurement of the impulse response of the opto-electrical subsystem of the *Lightwave Component Analyzer*, given its very poor S/N ratio (< 0 dB). As such, it was decided to replace the logarithmic averaging by regular averaging after estimating and compensating for the time drift.

The time base drift is estimated by minimizing

$$V = \sum_{m=1}^{M} \left| X_{ref}(\omega_m) - e^{-j\omega_m \tau} \cdot X(\omega_m) \right|^2 \qquad \text{Eq. 2-2}$$

with respect to τ, within the bandwidth[2] of the signal. $X_{ref}(\omega_m)$ represents the measured spectral data of the reference signal at $\omega_m = 2\pi f_m$, while $X(\omega_m)$ represents the corresponding measured spectral data of the signal to be aligned with respect to the reference signal. $\omega_m = m \cdot \omega_0 = m \cdot \frac{2\pi}{T_0}$, where T_0 represents the width of the time window, used to capture the impulse response. A starting value for τ is obtained, either based on a crosscorrelation test or by calculating the value of the cost function for a limited range of τ values on a sufficiently dense grid.

1. Based on equation (3.6-6) of [1] the bias on the amplitude estimation for a S/N ratio of 0 dB equals 0.9 dB.
2. In the case of impulse responses, measured using the 50 GHz sampling oscilloscopes, frequencies up to the first transmission zero (70 GHz) of the sampling oscilloscope are taken into account.

In the actual implementation, the first measurement is taken as a reference and all successive measurements are aligned with respect to that reference. In [6], it is shown that a better alignment is possible when estimating the relative shifts of each measurement with respect to all other measurements or by using an adaptive reference. Because these methods require that one keeps track of all the realizations, these methods are not implemented as part of the streamlined implementation. Furthermore, a comparison of the estimated drifts under realistic conditions (figure 2-3, figure 2-4), shows a good correspondence between the actual implementation and the optimal implementation referred to in [6]. Furthermore, in the “System identification approach applied to jitter estimation” chapter, it is explained that time base jitter is interpreted as time base drift and smoothing of the estimated time base drift is proposed.

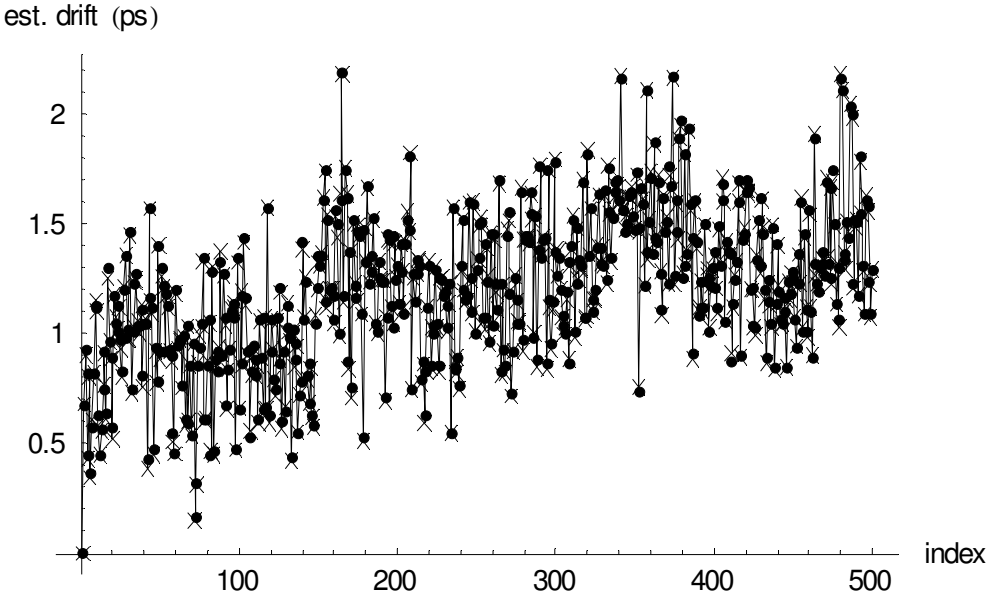

Figure 2-3. Estimated time base drift (x: actual implementation, dots: implementation proposed by NIST [6]) based on 500 impulse response measurements performed at NIST.

Due to the presence of the feedthrough of the strobe pulse, one must measure a positive and negative kickout to eliminate this unwanted response. To minimize the elapsed time between the measurement of a positive and a negative kickout, the measurement of a positive kickout is followed immediately by the measurement of a negative kickout. Using the above method, all positive kickouts can be aligned and averaged. The same is done for the negative kickouts.

Finally the mean value of the positive kickout and the mean value of the negative pulse must be aligned with respect to each other before they can be subtracted. Let τ_k^+ be the estimated time base drift of the k^{th} measurement of the positive kickout with respect to the first measurement, while τ_k^- is the equivalent for the negative kickout. It is observed that the overall shape of τ_k^- closely resembles that of τ_k^+ as

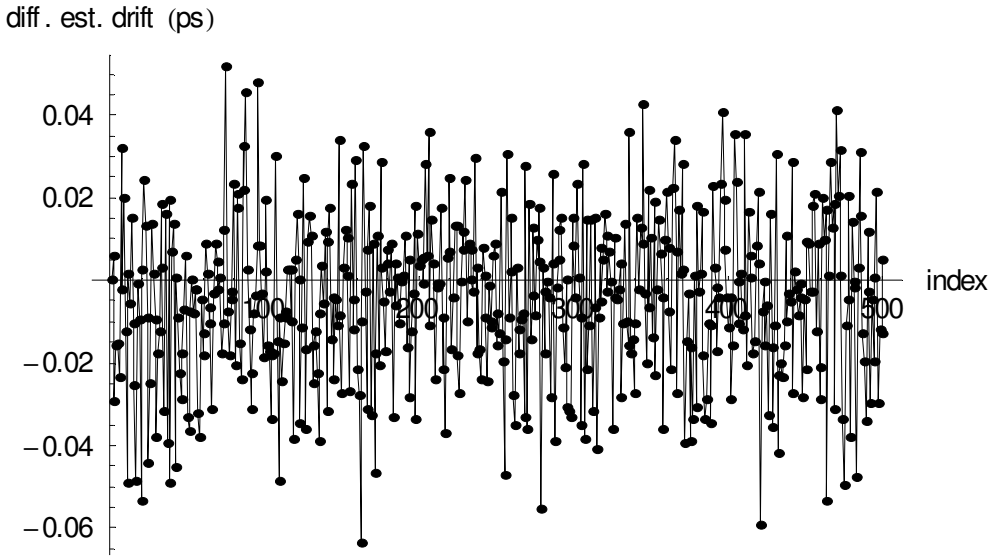

Figure 2-4. Difference between estimated time base drift based on the actual implementation and the implementation proposed by NIST [6].

function of k (figure 2-5), while the variation between τ_k^+ and τ_k^- can be large for each individual value of k (figure 2-6).

All positive kickouts are aligned with respect to the first positive kickout. All negative kickouts are aligned with respect to the first negative kickout. Based on the above, the mean drift of the positive kickout should equal the mean drift of the negative kickout[1]. Therefore, the actual difference between the mean drift of the positive kickouts and that of the negative kickouts can be used to align both averaged kickouts (figure 2-6).

Positioning and width of time window

The Agilent 83480 sampling oscilloscope has a limited data memory. When measuring an impulse response, one has to meet two boundary conditions. On one hand, one wants to keep the time window for the measurement as small as possible to allow high time resolution and reasonable S/N ratio. On the other hand, one must keep the window long enough to make sure that the impulse response is not truncated.

In the case of a nose-to-nose measurement, a kickout pulse is generated by one oscilloscope and measured by a second oscilloscope. However, due to the imperfect internal match of both oscilloscopes, some of the pulse is reflected back and forth between both sampling oscilloscopes. After averaging the kickout pulses, a first

1. This statement neglects the fact that the measurement of the negative kickouts is slightly offsetted in time with respect to the measurement of the positive kickouts.

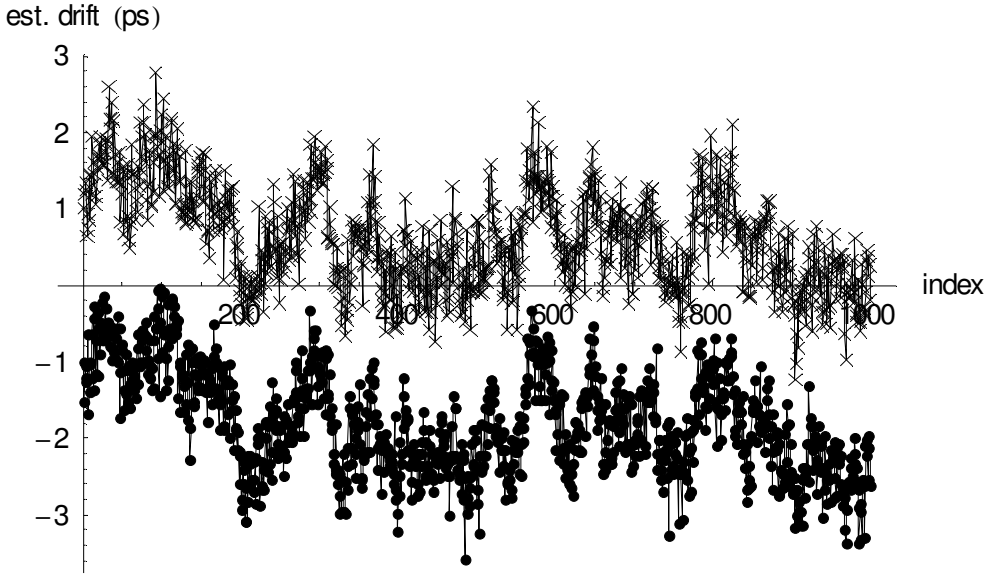

Figure 2-5. Estimated time base drift (x: positive kickout, dots: negative kickout) based on 1000 M_{ij} measurements. To ease comparison, the estimated time base drift of the positive kickout is shifted by 1 ps and that of the negative kickout by -1 ps.

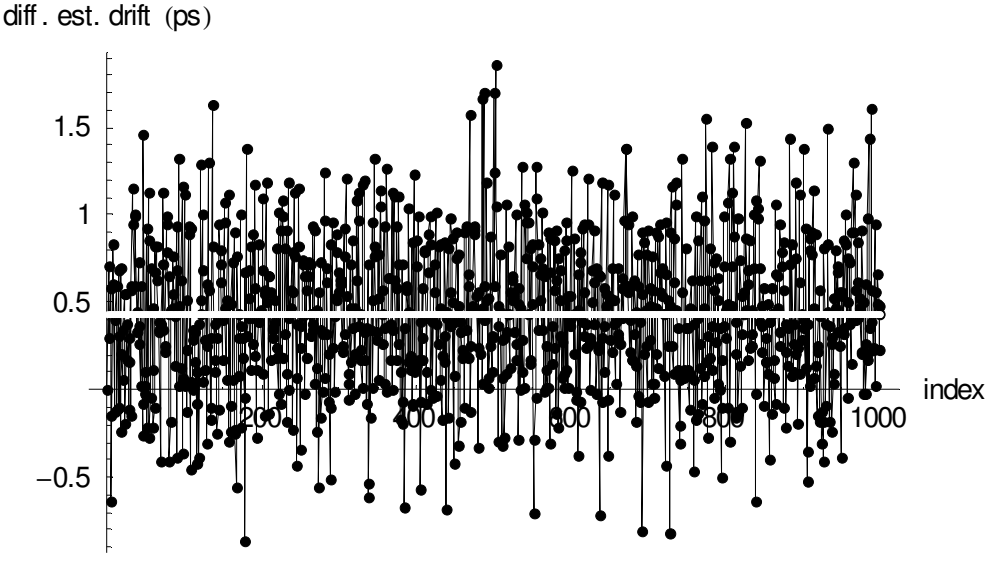

Figure 2-6. Difference of the estimated time base drift of the positive and negative kickout (white line: mean difference, used to align the averaged positive and negative kickout pulse).

reflection is clearly visible at approx. 1 ns delay with respect to the main pulse. By comparing the spectrum after averaging for different time window lengths, it was found experimentally that one can measure up to the third reflection. The time window for the measurement was therefore set to approx. 4 ns.

It was found that the actual distortion of certain portions of the time base of the 50 GHz sampling oscilloscopes varies both with the trigger frequency and the selected time step. This means that the time base distortion measurement must be performed using the same trigger and time step settings as the ones that are used during the actual measurement. Due to practical limitations of the sampling oscilloscope, the trigger frequency during nose-to-nose can not be increased above 2.5 kHz. Due to other trigger hardware limitations, the smallest achievable trigger frequency during the measurement of the time base distortion is approx. 5 kHz. It was found experimentally that the impacted regions of the time base are located at the beginning of the time window and after each discontinuity of the time base. For the Agilent 83480A Digital Communication Analyzer, the position of these discontinuities is known to be located at 22 ns + k.4 ns, $k \in \mathrm{N}$. The discontinuity of the time base is due to the usage of a 250 MHz restartable oscillator in combination with a fine ramp of 4 ns to create the time base. The impacted regions were found to span up to several tenths of nanoseconds and the spans seem to increase with temperature. Figure 2-7 up to figure 2-10 show the impact of changing the trigger repetition rate on the time base distortion.

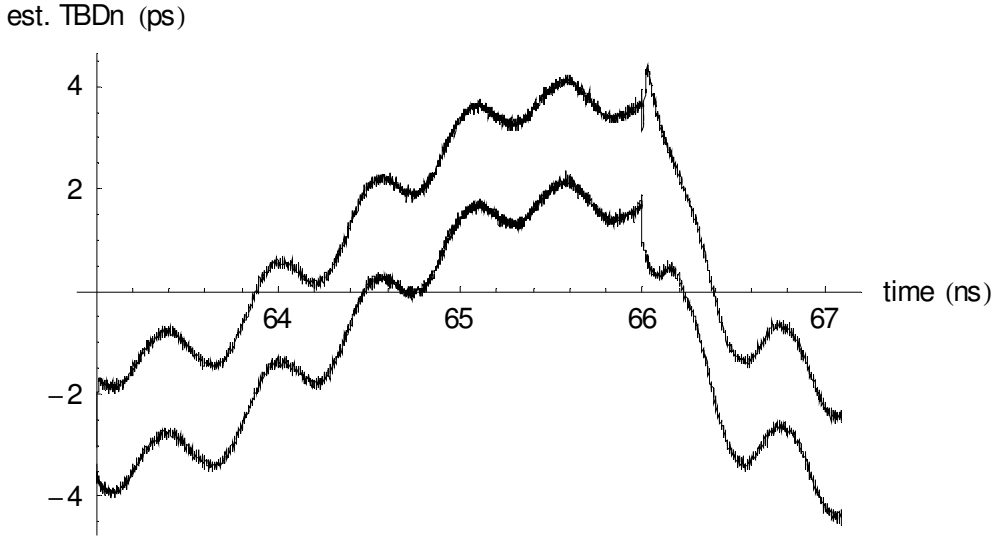

Figure 2-7. Estimated time base distortion (trigger rep. rate of 4.7 kHz and 27 kHz). To ease comparison, the estimated time base distortion using the 4.7 kHz trigger is shifted by 1 ps (upper curve) and that of the 27 kHz trigger is shifted by -1 ps (lower curve).

As such, it is important that the main pulse and its main reflections are not located in this region. Therefore the time window is set to start at 63 ns (1 ns after the 62 ns discontinuity) and the main pulse is located at 0.5 ns delay with respect to the left edge of the time window.

The shape of the difference of the estimated time base distortion (figure 2-10) deserves some additional attention. Because of the equivalent-time sampling, the physical time between two successive sampling instants does not correspond to the

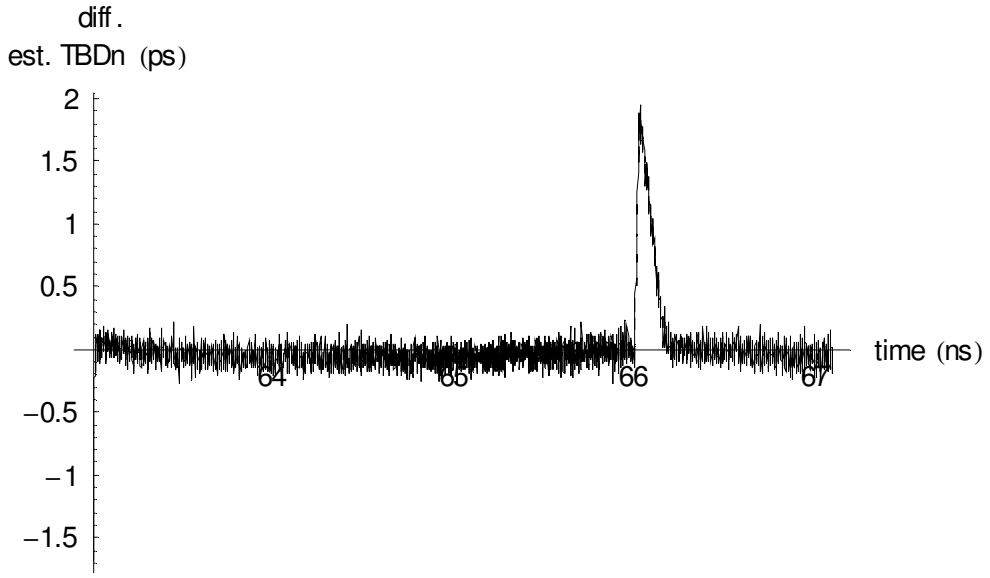

Figure 2-8. Difference of estimated time base distortion (trigger rep. rate of 4.7 kHz versus 27 kHz).

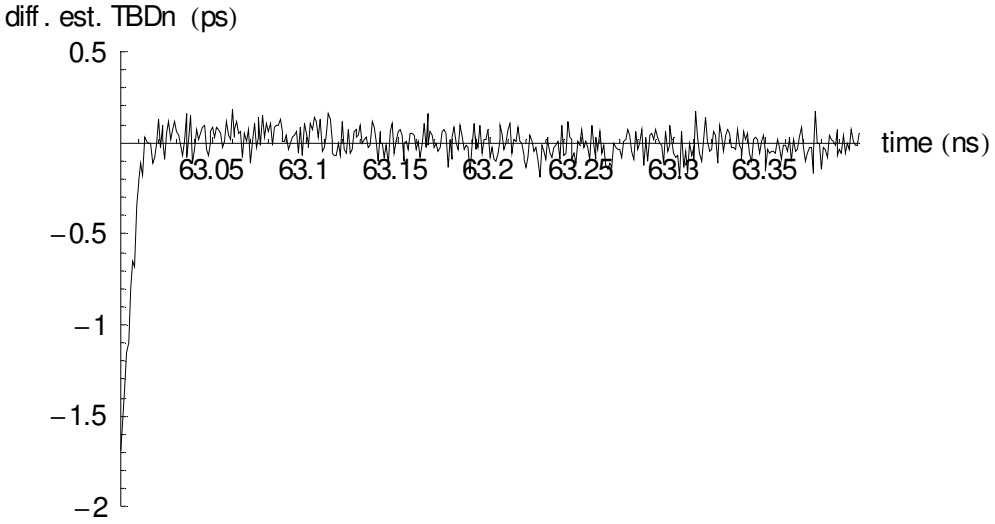

Figure 2-9. Difference of estimated time base distortion (trigger rep. rate of 4.7 kHz versus 27 kHz). Zooming in to the start of the window.

specified time step, but is determined by the trigger period[1]. As such, using a trigger repetition rate of 27 kHz and using an equivalent-time step of 1 ps, a time window of 0.17 ns corresponds to 170 samples and a physical time of 6.3 ms. Using a trigger repetition rate of 4.7 kHz, the same physical time of 6.3 ms corresponds to

1. For the 83480 DCA, the internal sample frequency equals the trigger repetition rate, if this one is smaller than or equal to 40 kHz. Otherwise the internal sample frequency is limited to 40 kHz.

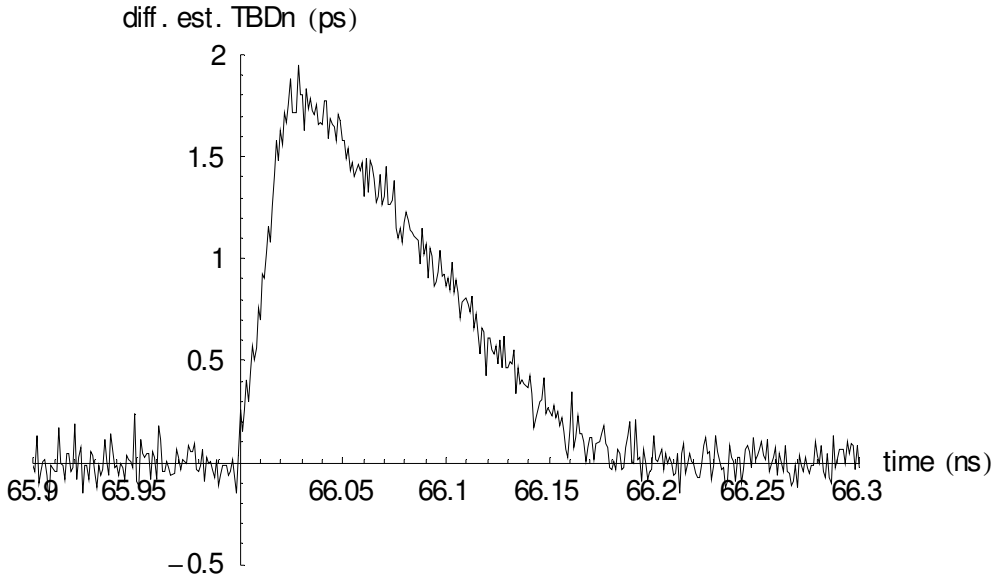

Figure 2-10. Difference of estimated time base distortion (trigger rep. rate of 4.7 kHz versus 27 kHz). Zooming in to the portion of the time base at the discontinuity of 66 ns.

30 samples and an equivalent-time of 0.03 ns. This means that a phenomenon with a given physical duration will manifest itself differently, depending on the applied trigger repetition rate. Figure 2-11 shows a time base distortion step of 2.2 ps, which linearly decreases as function of the physical time. As explained above, the impacted equivalent-time is different for a trigger repetition rate of 27 kHz and 4.7 kHz. Subtracting this effect results in a difference which is very similar to the one shown in figure 2-10.

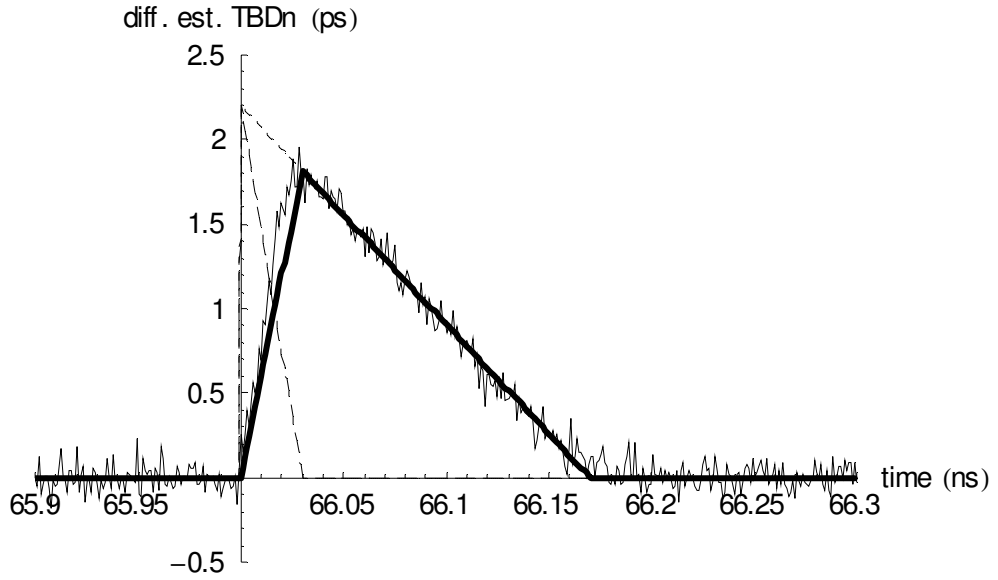

Figure 2-11. Simple model for the difference of the estimated time base distortion, corresponding to different trigger repetition rates. (long dashed line: 4.7 kHz trigger, short dashed line: 27 kHz trigger, thick line: difference of short and long dashed line).

Enhanced time base distortion estimation and faster correction

The original time base distortion estimation is replaced by a better technique, while the compensation is replaced by a faster technique. The latter allows to choose a simple local interpolator such that the systematic interpolation error remains below the noise floor of the reconstructed signal.

The original time base distortion estimation, as described in [1], basically performs a phase demodulation to extract the time base distortion. Systematic errors are introduced because of two reasons. First, the method assumes that the time base distortion can be represented by a band limited signal. This explains the modelling errors around the discontinuities of the time base. Also, the windowing which is performed in the time domain to reduce leakage, introduces large systematic errors at the boundaries of the time window.

The estimation is replaced by the maximum likelihood estimator (MLE), described in [8]. It combines the advantages of a non-parametric time base and the efficiency and robustness provided by the use of a statistical framework. In practice, the comparison of the actual value of the cost function and its expected value allows to verify for the presence of model errors. For instance, the method requires measurements at two or more non-harmonically related frequencies to distinguish between harmonics due to time base distortion and due to the (vertical) nonlinear behaviour of the oscilloscope. Harmonics can also be produced by the source. It was found that using certain combinations of frequencies, the actual cost was significantly larger than the expected one. It turned out that this was caused by the fact that the actual time base distortion varied with the applied frequency, while the method assumes that the time base distortion is independent of the applied frequency. The second advantage of the statistical framework is that uncertainty bounds are provided, which can be used to provide uncertainty bounds for the overall nose-to-nose method.

Once the time base distortion has been estimated, the next step is to compensate the M_{ij} measurements for this distortion. Although this may seem to be simple, the original time base distortion compensation, as described in [1], is very time consuming. The compensation is based on the construction of a least-squares estimator and requires the solution of a set of linear equations $y = A \cdot x$ in a least-square sense. A is a $N \times (2C + 1)$ real matrix, C represents the assumed number of spectral components and N corresponds to the number of non-equidistant measured time points. A typical value for N is 2048 while C is approx. 300 to 600. Using singular value decomposition[1] to calculate the solution, it was found that the calculation is too time consuming. Indeed, a typical solver requires $O(N^2 \cdot (2C + 1))$ operations to solve this set of equations. As such, it does not allow a Monte Carlo analysis to study

1. based on the implementation in C++ of a commercially available mathematical library M++.

the uncertainty after time base distortion correction, taking both the uncertainty on the estimated time base distortion and the uncertainty on the sample values into account.

To speed up the compensation, the solution of the set of equations is replaced by a KIS ("Keep It Simple") approach as described in [9]. The cubic interpolation method is very fast and requires only $O(N)$ operations. The error introduced by this interpolation increases with the relative bandwidth[1] of the signal to be interpolated. Fortunately, based on the inherent high oversampling rate of an equivalent-time sampling oscilloscope, the bandwidth of the measured signal relative to the sampling frequency is low: given a signal bandwidth of approx. 50 GHz, a 4 ns time window and 2000 points, the relative bandwidth is 0.1. In worst case, the systematic deviation from the equidistant time grid[2] equals $0.5\Delta t$: the sampling instance based on the distorted time base is located right in between the ideal equidistant sampling instances. For this relative bandwidth and systematic deviation, based on the simulations performed in [9], the mean squared error is approx. -40 dB relative to the root-mean-square (RMS) value of the signal.

A typical value for the jitter standard deviation is approx. 1 ps, which corresponds to $0.5\Delta t$. Applying an offset of 0.1 V to the kickout generator, an M_{ij} measurement typically has a signal-to-noise ratio (SNR) of approx. 20 dB. Given the above and based on the interpolator selection table in [9], it makes sense to use a simple cubic interpolation if time base distortion compensation is required before averaging. If one can increase the SNR first by averaging, it makes sense to consider cubic spline interpolation instead. Figure 2-12 shows that the reconstruction error using the fast KIS approach remains well below the measurement noise, even after averaging.

The actual nose-to-nose implementation uses cubic interpolation to allow time base distortion compensation before averaging. This was based on the consideration that within the actual time window, the time base distortion is fixed for all M_{ij} realizations. As such, it makes sense to compensate first for time base distortion, before estimating the time base drift. As described earlier, the latter is necessary to allow regular averaging. It was found however that compensating for time base distortion first had no noticeable effect. As such, one may choose to apply averaging first to increase the SNR and to decrease the equivalent time jitter such that one can use cubic spline interpolation instead of cubic interpolation. Due to time constraints, this alternative approach is not implemented or tested.

Frequency domain interpolation using the chirp-z transform

In general, an estimate of the phase distortion of the sampling oscilloscope is required at a frequency grid which does not correspond to the original 250 MHz frequency grid,

1. relative bandwidth being defined in [9] as the signal bandwidth divided by half the sampling frequency.
2. this systematic deviation from an equidistant time grid is referred to as 'jitter deviation' in [9]

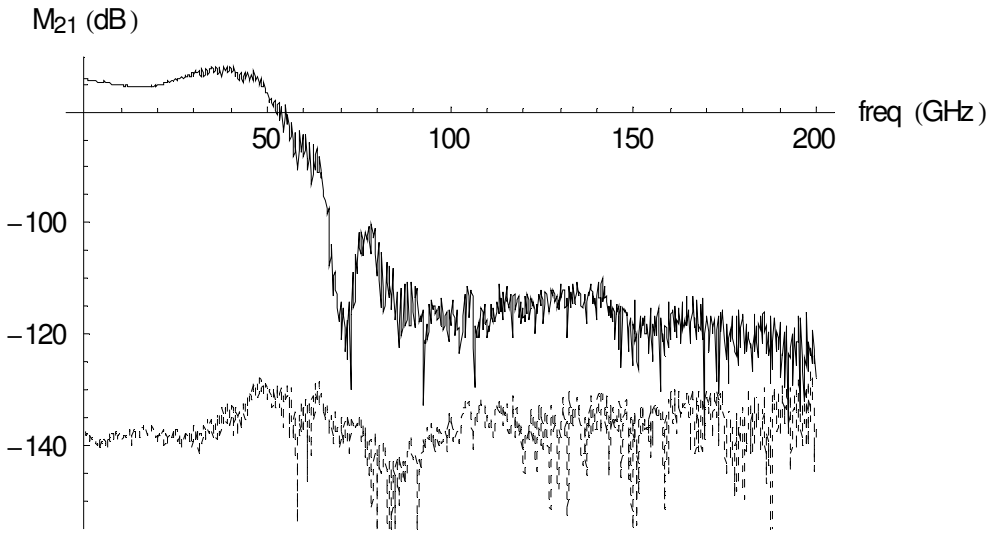

Figure 2-12. Example M_{ij} spectrum up to 200 GHz, based on 1000 averages and after TBDn compensation based on [1] (solid line). Dashed line: complex difference (in dB) between the TBDn-compensated M_{ij}, based on the time-consuming approach proposed by [1] and based on the fast KIS-alternative [9].

based on the initial 4 ns time window. For instance, for the calibration of the harmonic phase reference a frequency grid of 2 MHz is required. Because M_{ij} corresponds to an impulse response, one can use the fact that the signal is zero at both edges. As such it is possible to append zero values. The naive implementation is to do so in the time domain, but then one cannot obtain an arbitrary frequency grid. An arbitrary frequency grid can be obtained using a chirp-z transform [8].

Another option, which is not implemented or tested, is to use a linear time-invariant model.

Conclusions

The implementation of a streamlined process allowed to transfer the nose-to-nose calibration technique to the calibration lab of the Lightwave division within Agilent Technologies. It contributed significantly to the success of the Agilent 86030A 50 GHz *Lightwave Component Analyzer*.

Based on the demonstration and explanation of the process, the nose-to-nose calibration technique was evaluated and implemented successfully by people at NIST.

Over time, some of the original algorithms, described in [1], were replaced by new algorithms, which are superior either in speed or in quality.

References

[1] Jan Verspecht, “Calibration of a Measurement System for High Frequency Nonlinear Devices”, *PhD dissertation*, 1995.
[2] P. Hale, T. Clement, K. Coakley, C. Wand, D. DeGroot and A. Verdoni, “Estimating the Magnitude and Phase Response of a 50 GHz Sampling Oscilloscope Using the “Nose-To-Nose” Method”, *55th ARFTG Conf. Digest*, June 2000.
[3] D. DeGroot, P. Hale, M. Vanden Bossche, F. Verbeyst, and J. Verspecht, "Analysis of interconnection networks and mismatch in the nose-to-nose calibration," *55th ARFTG Conf. Digest*, pp. 116-121, June 2000.
[4] C. Wang, P. Hale and K. Coakley, “Least-Squares Estimation of Time-Base Distortion of Sampling Oscilloscopes,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 48, No. 6, December 1999
[5] C. Wang, P. Hale, K. Coakley and T. Clement, “Uncertainty of Oscilloscope Timebase Distortion Estimate,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 51, No. 1, February 2002
[6] K. Coakley and P. Hale, “Alignment of Noisy Signals,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No. 1, February 2001
[7] K. Coakley, C. Wang, P. Hale, T. Clement, “Adaptive Characterization of Jitter Noise in Sampled High-Speed Signals,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 52, No. 5, October 2003
[8] G. Vandersteen, Y. Rolain, J. Schoukens, “An Identification Technique for Data Acquisition Characterization in the Presence of Nonlinear Distortions and Time Base Distortions,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No. 5, October 2001
[9] Y. Rolain, J. Schoukens and G. Vandersteen, “Signal Reconstruction for Non-Equidistant Finite Length Sample Sets: a “KIS” approach,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 47, No. 5, October 1998, pp. 1046 - 1052
[10] A. Oppenheim and R. Shafer, “Digital Signal Processing,” *Prentice-Hall*, 1975

CHAPTER 3

Comparison of the nose-to-nose and EOS-based calibration technique.

Abstract

Only recently, the electro-optic sampling (EOS) system [1]-[4] at NIST allows to calibrate the impulse response of an opto-electrical (O/E) converter up to 110 GHz. If one then measures this impulse response using a high-speed sampling oscilloscope, one effectively measures the convolution of the (known) impulse response of the O/E and that of the sampling oscilloscope. In fact, the reality is more complex because one still needs to estimate and compensate for the time base drift, time base distortion and time base jitter of the oscilloscope. One also has to take into account the mismatch of the O/E and the oscilloscope and the S-parameters of the adapter to obtain the amplitude and phase distortion of the oscilloscope. Having done so, it becomes possible to compare the amplitude and phase distortion of a 50 GHz sampling oscilloscope plug-in, based on the nose-to-nose calibration technique, to the one obtained using a calibrated O/E.

A discrepancy was reported with respect to the phase distortion obtained by both calibration methods [5]. The difference starts around 20 GHz and increases as a function of the frequency. Earlier, a discrepancy was reported between the nose-to-nose based amplitude distortion and the amplitude distortion that is obtained using a stepped sine measurement [6].

The goal of this chapter is to verify the reported discrepancies. Therefore, an Agilent 83484A 50 GHz electrical plug-in was shipped to NIST, to measure the impulse response of an O/E converter, which has been calibrated using the EOS system. Based on these measurements, it is possible to verify the presence and the magnitude of the phase discrepancy reported in [5]. At the same time, the presence of the amplitude discrepancy is verified using two methods: one method that is similar to the one described in [6] and another new method.

As a consequence of this verification, the jitter estimation was given additional attention. This was the motivation for additional research and original work aimed at the estimation of drift in the presence of both additive and jitter noise. This work is presented in separate chapters.

Amplitude comparison

Overview

The amplitude characteristic of one channel of an Agilent 83484A 50 GHz electrical plug-in, inserted in a 83480A sampling oscilloscope mainframe, is obtained using four different methods (see figure 3-1 and figure 3-2). The first two approaches are based on stepped sine measurements, where a sine wave of fixed amplitude is stepped over the frequency band of interest in a number of steps, and the magnitude is measured separately at each frequency. The other two methods obtain the amplitude characteristic of the plug-in starting from impulse response measurements.

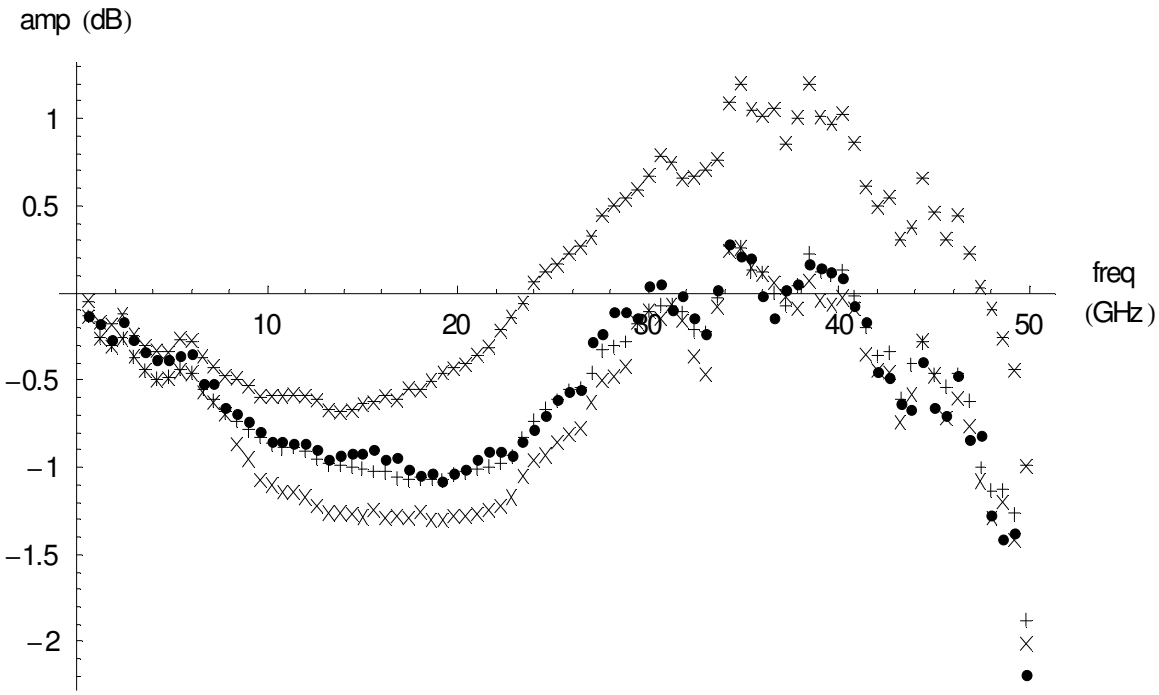

Figure 3-1. Estimated amplitude characteristic of a 50 GHz plug-in. (x: power meas. and histogram mode, +: power meas. and normal mode, -x-: nose-to-nose, dots: EOS).

The first (stepped sine) technique is similar to the one described in [6], except that the power of the incident voltage wave is measured using a one-port calibrated[1] *Large-Signal Network Analyzer* (LSNA) instead of using two power sensors. The sampling oscilloscope is used in freerun mode and a vertical histogram measurement is performed.

This technique has some disadvantages. In freerun mode, the one-step-ahead predictor used in the sampler of the scope is disabled [7], while it is enabled in

1. Short, Open, Load and power calibration (no phase calibration).

triggered mode. The power measured by the scope can be derived from the standard deviation of the histogram measurement. However, the latter assumes that in freerun mode, the time axis is randomly sampled using a uniform probability density function. Finally, the noise added by the sampling scope is measured without any signal being applied. Thus, it is assumed that this noise level is independent of the signal level. To get around these hypotheses, a second method that is similar to the first technique, uses the sampling oscilloscope in triggered mode. The disadvantage of this method is that one has to estimate and compensate all time base errors.

The third type of method uses the amplitude characteristic resulting from the nose-to-nose calibration technique. In the fourth method, the characteristic is obtained using the impulse response measurement of a transfer standard, an opto-electrical converter, which is calibrated using the EOS setup at NIST.

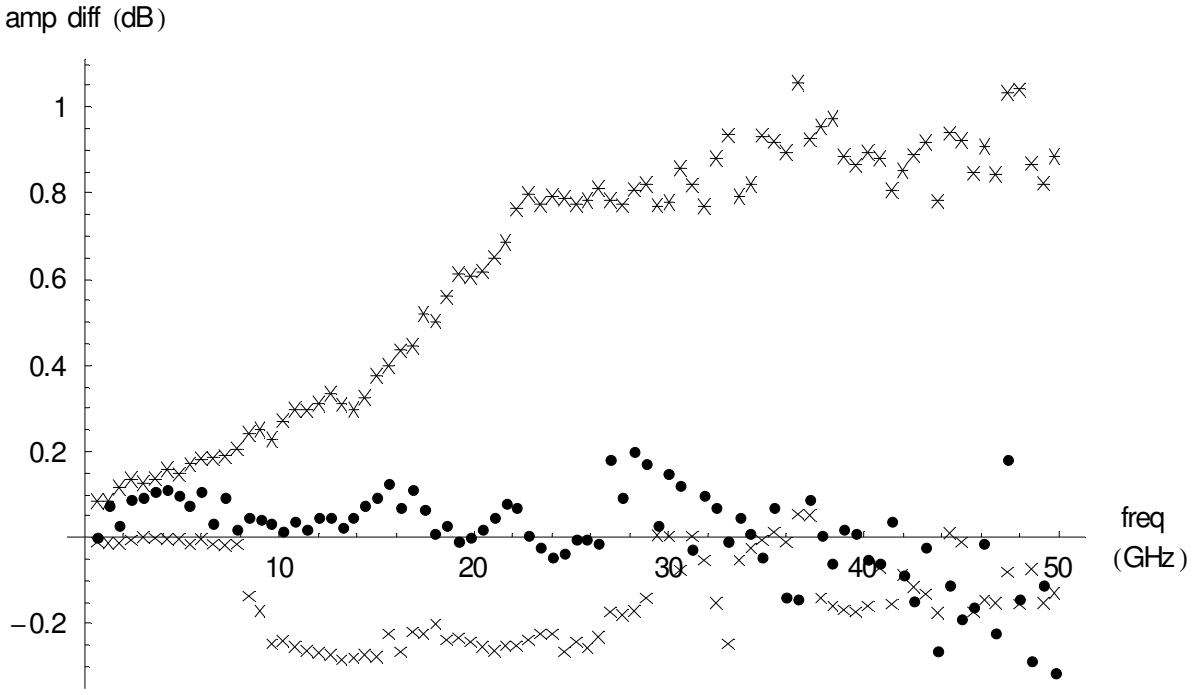

Figure 3-2. Difference of the estimated amplitude characteristic of a 50 GHz plug-in relative to method 2 (power meas. and normal mode). (x: power meas. and histogram mode, -x-: nose-to-nose, dots: EOS).

Method #1: LSNA power measurement and scope in histogram mode

Figure 3-3 shows a simplified version of the used measurement setup.

Because of its availability, the *Large-Signal Network Analyzer* is used to measure the incident voltage wave at the input connector of the sample scope. First a one-port short-open-load calibration and a power calibration are performed at the plane where the sample scope plug-in will be connected. No extra phase calibration is performed.

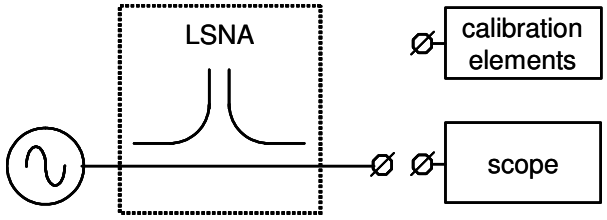

Figure 3-3. Measurement setup involving a *Large-Signal Network Analyzer* and a scope used in histogram mode.

Due to present limitations of the LSNA software, the measurement is performed on a 600 MHz grid from 600 MHz up to 50 GHz.

Due to the requirement for measurements up to 50 GHz, the cabling is kept as short as possible to minimize the losses. The source power is adapted to obtain an incident power of -10 dBm at the calibration plane. This corresponds to 100 mVp, and is sufficient in order to guarantee a linear operation of the sample scope. It is also verified that the harmonics that are generated by the source are sufficiently small (more than 40 dB down).

At each frequency, the amplitude of the incident voltage wave is measured by the LSNA. The scope is used in freerun triggering mode. The vertical histogram measurement reveals the probability density function (PDF) of a sine wave corrupted by additive Gaussian noise. In order to obtain the additive noise generated by the sampling oscilloscope as a result of the measurement, an additional vertical histogram measurement is performed without any external signal applied and with the input of the scope terminated in a 50 Ω load. This allows to correct the RMS value of the measured voltage for the noise of the scope using $\sqrt{\sigma_m{}^2 - \sigma_n^2}$, where σ_m corresponds to the measured standard deviation of the sine wave. A value $\sigma_n \cong 0.57$ mV is measured for the standard deviation when no signal is applied to the scope.

As a sanity check, the measured fraction of samples which lie within $\pm\sigma$ of the mean value is compared to the expected value of 50%, which is the theoretical value for a sine wave that is randomly sampled based on a uniform probability density function. σ corresponds to the measured standard deviation.

Method #2: LSNA power measurement and scope in normal mode

The second method is similar to the first method, except that the scope is used in triggered mode instead of freerun mode. As such, all time base corrections must be applied to the sample scope measurement.

First, a time base distortion measurement is performed. Subtle effects like the thermal tail[1] must be properly dealt with. Therefore, both the trigger rate and the time base resolution are selected to be identical during the measurements that are used to estimate the time base distortion and during the actual measurements. In order to avoid leakage after time base distortion correction, the width of the selected acquisition window must equal an integer number of periods. All applied frequencies are an integer multiple of 600 MHz and therefore an acquisition window width of 5 ns is selected.

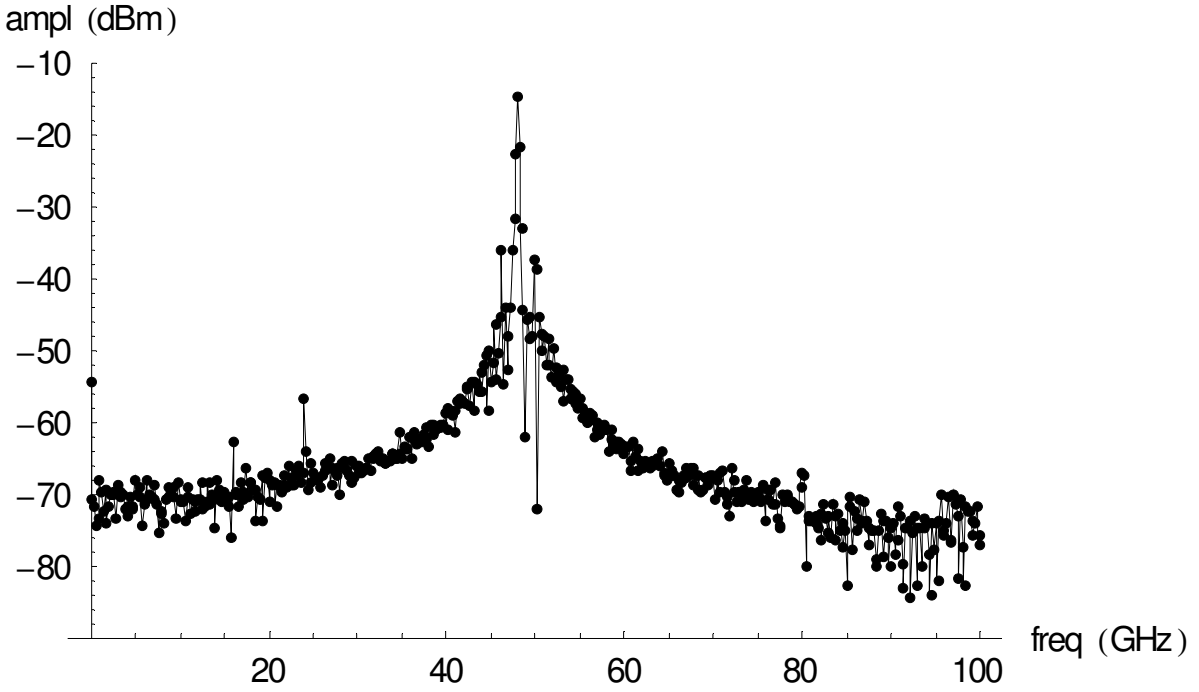

Figure 3-4. Measured averaged spectrum before time base distortion correction (48 GHz).

The time base drift estimation and the time base distortion correction are based on the methods described in the “Enhancements to the nose-to-nose calibration technique” chapter. Figure 3-4 shows the spectrum before time base distortion correction, based on the sample mean of 500 measured records of 4000 points each. The mean is calculated after time base drift estimation and compensation. The amplitude at 48 GHz is found to be -14.57 dBm. Figure 3-5 shows the spectrum after time base distortion correction. The amplitude at 48 GHz is now found to be -13.05 dBm. This demonstrates that a bias of 1.5 dB is introduced when one neglects the time base distortion. Figure 3-5 reveals a strange residual shaping of the spectrum around 48 GHz. Comparing the time signal to the pure sine wave based on the amplitude and phase at 48 GHz (figure 3-6) clearly shows some small residual “thermal tail” effect, which is assumed to be caused by small temperature variations in between the time base distortion measurements and the measurements at 48 GHz.

1. described in the “Enhancements to the nose-to-nose calibration technique” chapter.

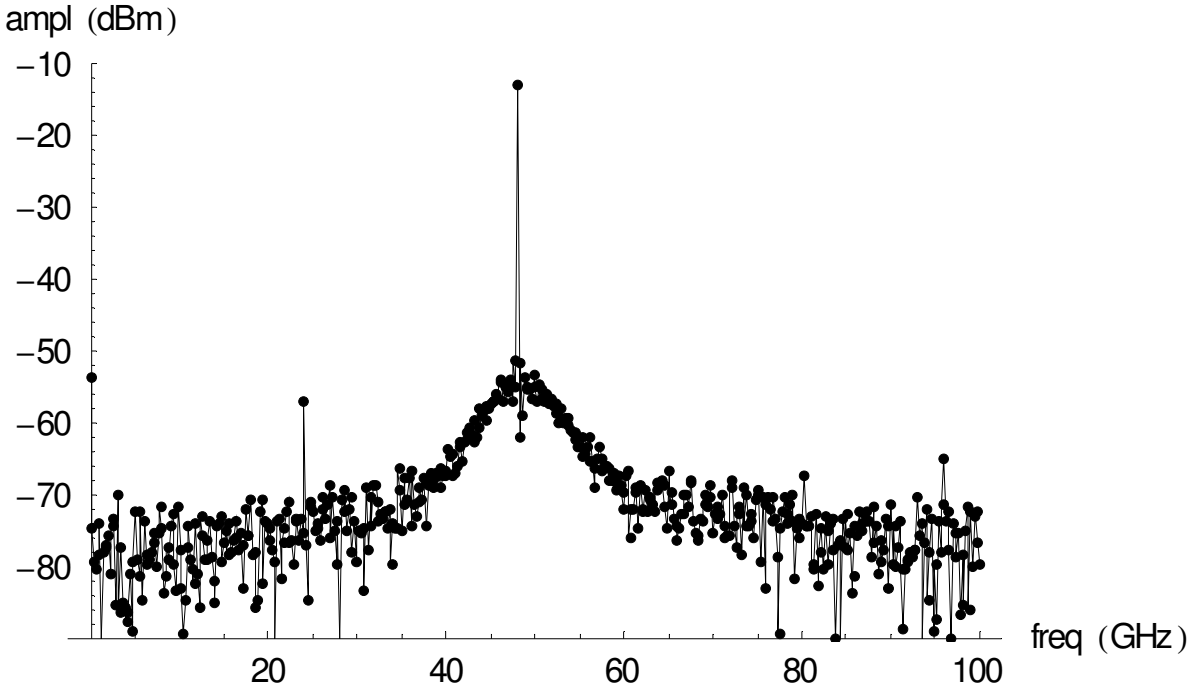

Figure 3-5. Measured averaged spectrum after time base distortion correction (48 GHz).

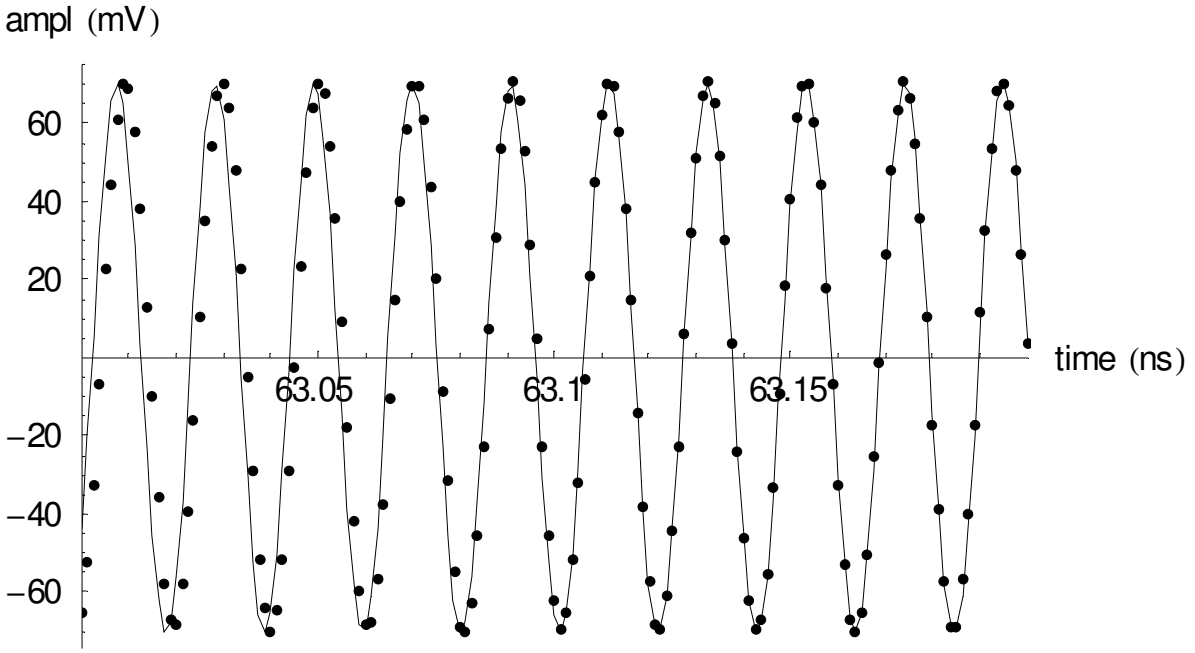

Figure 3-6. Residual “thermal tail” effect at the start of the acquisition window (48 GHz). Solid line: spectral component at 48 GHz only. Dots: all spectral components.

The removal of the minimal number of points (500) at the start of the acquisition window which allows to maintain an integer number of periods, results in figure 3-7. The amplitude at 48 GHz is now -13.03 dBm. Thus, the impact on the measured amplitude is 0.02 dB and therefore negligible. The remaining shaping of the spectrum around 48 GHz is believed to be caused by the phase noise of the 50 GHz Agilent 83650 source. Also, both the subharmonic at $f/2$ and the second harmonic are visible.

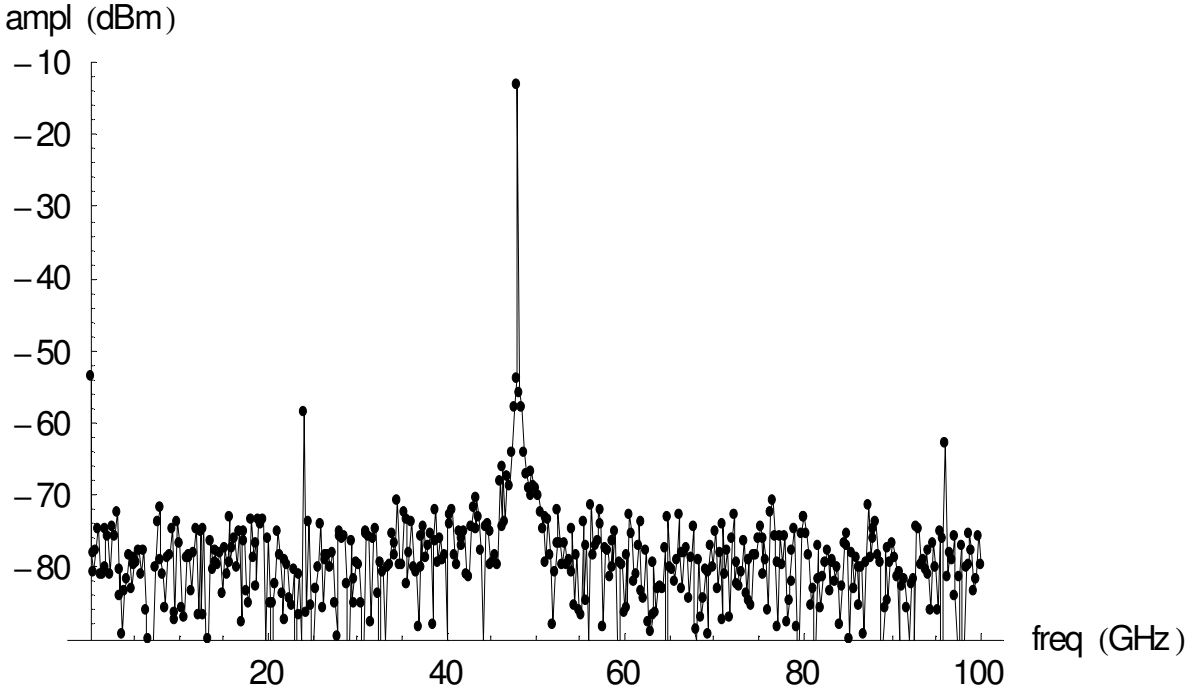

Figure 3-7. Measured averaged spectrum after time base distortion correction and after the removal of the samples impacted by the residual thermal tail (48 GHz).

The initial time base jitter estimation is performed using a first order approximation (Eq. 3-1) of the sample variance in the presence of both additive and jitter noise, as is described in more detail in the “System identification approach applied to jitter estimation“ chapter.

$$\tilde{\sigma}^2_{y_1}(t_i) = \sigma^2_{n_y} + \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i} \cdot \sigma^2_{n_t}. \qquad \text{Eq. 3-1}$$

A simple linear regression technique is used to estimate both the variance of the additive noise $\sigma^2_{n_y}$ and the variance of the jitter noise $\sigma^2_{n_t}$, using the measured sample variance and the squared derivative[1] of the measured sample mean. The resulting jitter standard deviation is 0.95 ps, while the standard deviation of the additive noise is found to be 7.8 mV. The latter is much larger than expected.

Assuming a normal probability density function, a jitter standard deviation of 0.95 ps at 48 GHz corresponds in an attenuation of the amplitude of 0.36 dB. The final amplitude at 48 GHz after all compensations becomes -12.67 dBm.

An additional study, which is described in detail in the “System identification approach applied to jitter estimation“ chapter, shows that the use of a straightforward time base drift compensation incorrectly shapes the sample variance. The study also proposes higher-order models for the sample variance in the presence of both additive and jitter noise. Correct time base drift compensation in combination with a higher-order model gives very good results in the case of an impulse response measurement for realistic values of the standard deviation of both jitter and additive noise and in the case there are no model errors. Applying this technique to sine wave measurements, yields estimates of $\sigma_{n_y}^2$, which strongly vary as a function of the selected model order and even can become negative.

Fortunately, one can derive the exact expression for both the expected value and the variance of a pure sine wave in the presence of normally distributed jitter noise and additive noise. Consider

$$y(t_i) = A\sin\{\omega(t_i + n_t(t_i)) + \phi\} + n_y(t_i) \qquad \text{Eq. 3-2}$$

$y(t_i)$ represents the observation of the pure sine wave, that is contaminated by both additive noise and jitter. The noise sources are considered to be part of the observation. Both $n_y(t_i)$ and $n_t(t_i)$ are assumed to be zero mean, normally distributed, independent and stationary with respect to t_i. ω and t_i are assumed to be known exactly.

Using the characteristic function of a normal distribution, it can be shown [8] that

$$\mu = E\{y(t_i)\} = A \cdot e^{-\frac{\omega^2 \cdot \sigma_{n_t}^2}{2}} \cdot \sin(\omega t_i + \phi) \qquad \text{Eq. 3-3}$$

Now one can also calculate the variance

$$\sigma^2 = E\left\{[y(t_i) - E\{y(t_i)\}]^2\right\} = E\left\{[y(t_i)]^2\right\} - [E\{y(t_i)\}]^2 \qquad \text{Eq. 3-4}$$

1. using the standard function implemented in *VEEtest™* from Agilent Technologies and based on a sliding fourth-order (five-point) polynomial.

The first term in this expression is found to be:

$$E\left\{\left[A\sin\{\omega(t_i+n_t(t_i))+\phi\}+n_y(t_i)\right]^2\right\}$$

$$= A^2E\left\{\left[\sin\{\omega(t_i+n_t(t_i))+\phi\}\right]^2\right\}+\sigma_{n_y}^2$$

$$= \frac{A^2}{2}-\frac{A^2}{2}\cdot E\{\cos[2\omega(t_i+n_t(t_i))+2\phi]\}+\sigma_{n_y}^2$$

$$= \frac{A^2}{2}-\frac{A^2}{2}\cdot e^{-\frac{(2\omega)^2\cdot\sigma_{n_t}^2}{2}}\cdot\cos[2(\omega t_i+\phi)]+\sigma_{n_y}^2 \qquad \text{Eq. 3-5}$$

Combining equations Eq. 3-3 up to Eq. 3-5, one obtains the value of the variance:

$$\sigma^2 = \sigma_{n_y}^2+\frac{A^2}{2}\left\{1-e^{-\omega^2\cdot\sigma_{n_t}^2}\right\}\left\{1+e^{-\omega^2\cdot\sigma_{n_t}^2}\cdot\cos[2(\omega t_i+\phi)]\right\} \qquad \text{Eq. 3-6}$$

In the absence of time base distortion, one can easily retrieve the values of A, ϕ, $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$ using the Fourier transform of Eq. 3-3 and Eq. 3-6.

When time base distortion is present, as is the case during our measurements, one can obtain an estimate for A, ϕ, $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$ by minimizing the following cost function with respect to these unknowns:

$$V_{(W)LS} = \sum_{i=1}^{N}\frac{1}{W_i^2}\left[\sigma_i^2-\sigma_{n_y}^2-\frac{A^2}{2}\left\{1-e^{-\omega^2\cdot\sigma_{n_t}^2}\right\}\left\{1+e^{-\omega^2\cdot\sigma_{n_t}^2}\cdot\cos(2\omega t_i)\right\}\right]^2$$

$$+\sum_{i=1}^{N}\frac{1}{W_i'^2}\left[\mu_i-A\cdot e^{-\frac{\omega^2\cdot\sigma_{n_t}^2}{2}}\cdot\sin(\omega t_i+\phi)\right]^2 \qquad \text{Eq. 3-7}$$

σ_i^2 and μ_i respectively represent the measured sample variance and sample mean at time instant t_i. The optional factors W_i^2 and $W_i'^2$ allow for a weighting. W_i^2 can be based on the sample variance of σ_i^2, if available, or can be evaluated using the variance of the χ^2 - distribution of σ_i^2; σ_i^2 itself can be used for $W_i'^2$.

First, the correctness of the implementation of this estimator (Eq. 3-7) is verified using simulations.

Next, the estimator is applied to the sine wave measurement at 48 GHz. First, a time base drift compensation is applied, as is explained in the “System identification approach applied to jitter estimation“ chapter. This yields the sample variance and the sample mean data on a non-equidistant time grid t_i. The latter is estimated based on a time base distortion measurement, which was performed up front.

Figure 3-8 and figure 3-9 show the sample mean and sample variance of the first two periods of the sine wave measurement obtained at an excitation frequency at 48 GHz for the LS estimator. Although the estimator also provides estimates for ϕ, $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$, the main parameter of interest here is A. It corresponds to the amplitude of the sine wave “before” the low-pass effect of the jitter. The LS estimator yields an amplitude of 74.19 mV, while its WLS equivalent provides a value of 74.14 mV ± 0.02 mV (95% confidence interval).

Starting from an initial value of -14.57 dBm, the amplitude of the sine wave at 48 GHz after compensation for the time base drift, the time base distortion and the time base jitter increases up to -12.59 dBm. This is only 0.1 dB larger than the -12.67 dB based on the initial, less correct, approach (see page 3-9).

A similar verification is performed for the sine wave measurement at 43.2 GHz and yields an amplitude of -11.58 dBm, which is only 0.05 dB larger than the value obtained using the initial approach.

As such, it is concluded that the initial approach (used for all other frequencies) is sufficiently accurate to be used during the comparison. It is represented by the ‘+’ symbols in figure 3-1 on page 3-3.

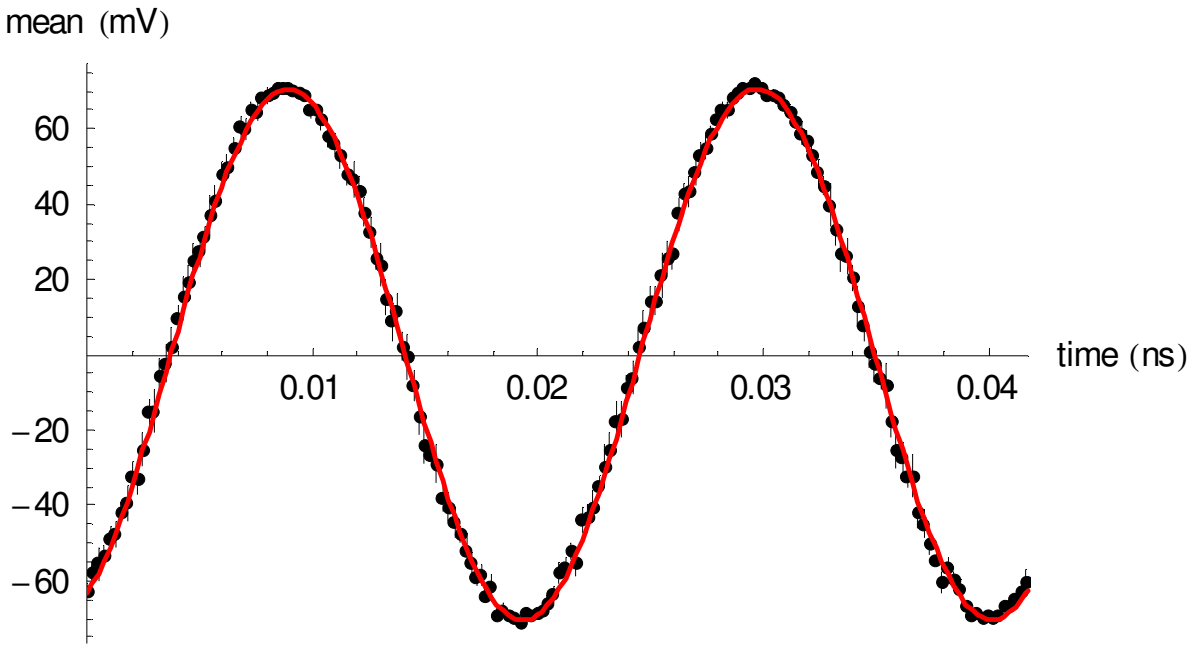

Figure 3-8. First two periods of the measured sample mean at 48 GHz (black dots), its 95% confidence interval (vertical black lines) and the estimated mean (red solid line).

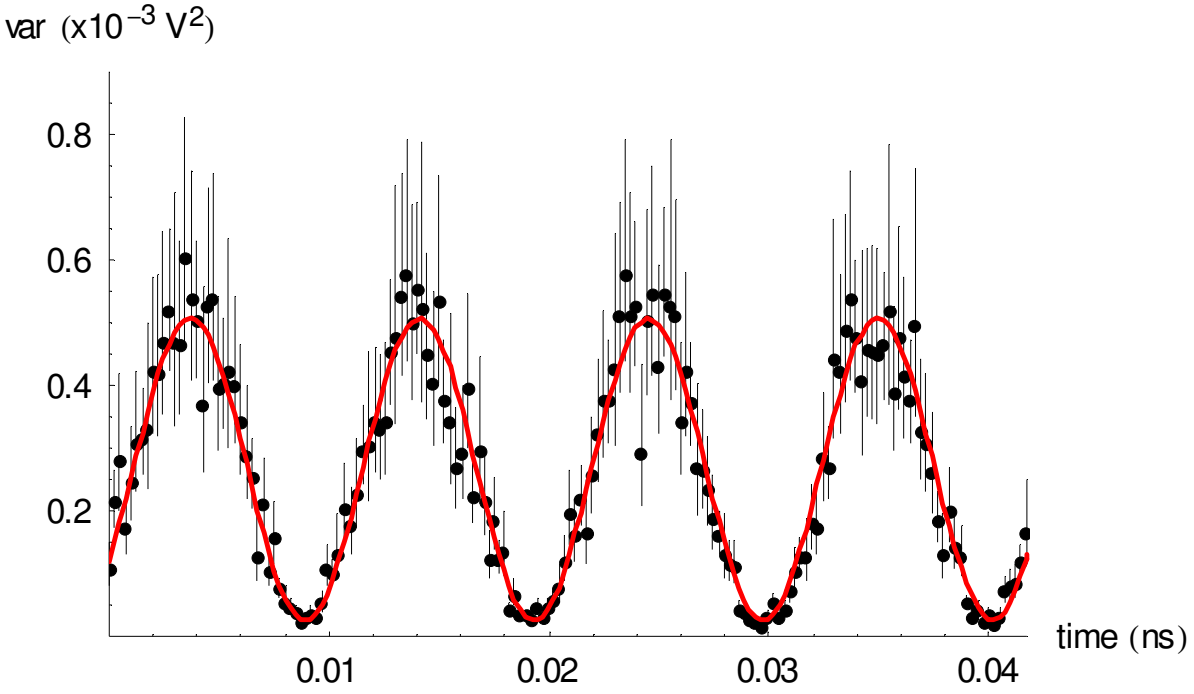

Figure 3-9. Corresponding measured sample variance at 48 GHz (black dots), its 95% confidence interval (vertical black lines) and the estimated variance (red solid line).

Method #3: amplitude characteristic based on nose-to-nose

The nose-to-nose calibration procedure used here is briefly explained in the "Enhancements to the nose-to-nose calibration technique" chapter.

The amplitude characteristic resulting from a nose-to-nose calibration is shown in figure 3-1 on page 3-3. The nose-to-nose calibration is used as the reference for the phase calibration of the *Large-Signal Network Analyzer*. Given a symmetrical jitter probability density function, it is shown [8] that jitter does not introduce any phase distortion. Therefore, although the jitter estimation and compensation was included as part of the amplitude correction, it was not given as much attention as the time base drift, the time base distortion and the mismatch compensation.

As part of the comparison of the amplitude characteristic obtained using different techniques, the initial jitter estimation and its compensation is, once more, compared to the enhanced method, which is explained in detail in the "System identification approach applied to jitter estimation" chapter.

The jitter is estimated for both the positive and negative kickout pulses. Given identical trigger conditions for both kickouts, the result is expected to be identical within the uncertainty on the estimate. Figure 3-10 and figure 3-11 show the measured and modelled variance, including the boundaries of their 95% confidence interval, corresponding to the main pulse and to the first reflection. Figure 3-11 also indicates

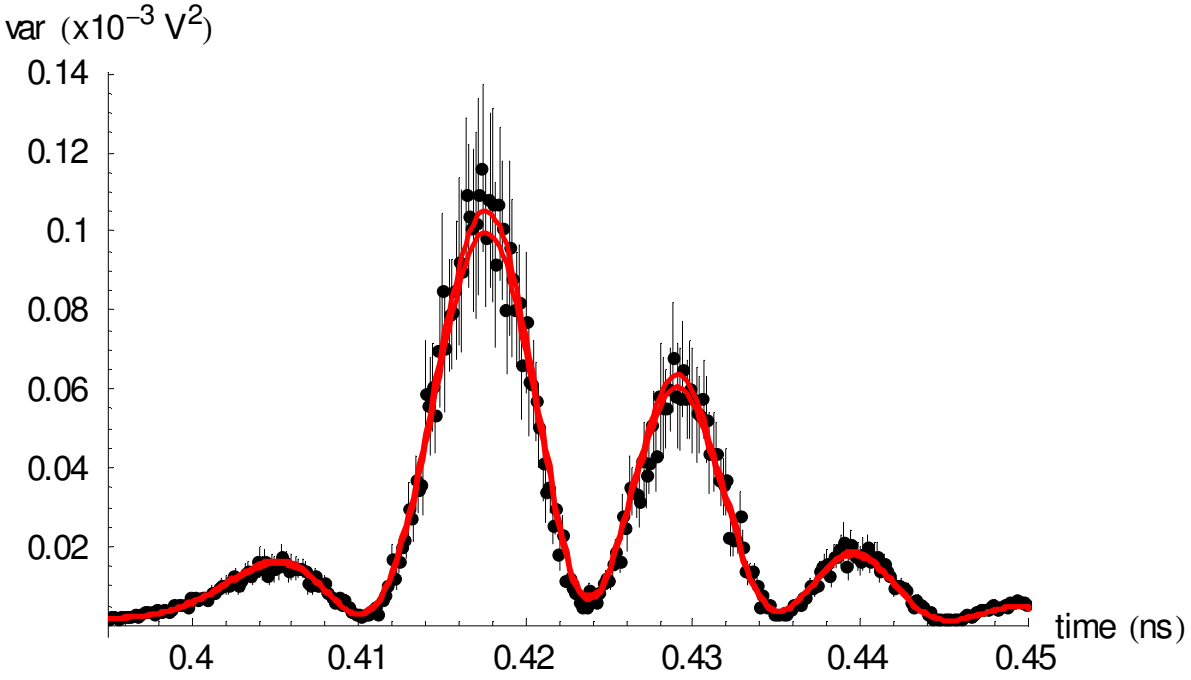

Figure 3-10. Measured (black dots, vertical black lines: 95% confidence interval) and modelled variance (solid red curves, boundaries of the 95% confidence interval) of the positive kickout pulse. Zooming into the variance corresponding to the main portion of the kickout pulse.

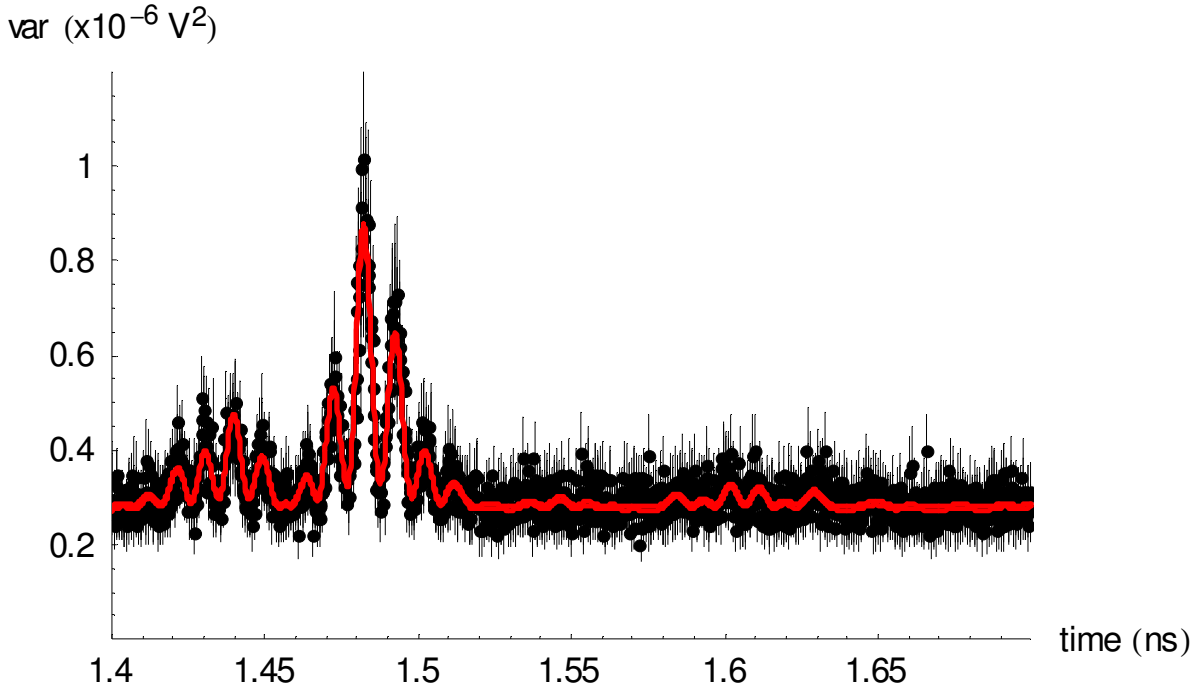

Figure 3-11. Measured (black dots, vertical black lines: 95% confidence interval) and modelled variance (solid red curves, boundaries of the 95% confidence interval) of the positive kickout pulse. Zooming into the variance corresponding to the first reflection of the kickout pulse.

that the model is also capable to model the variance of the additive noise. A least square estimator is used in combination with a third order model because the jitter standard deviation exceeds 1 ps. The rationale of this decision and the details of the jitter estimation are explained in the “System identification approach applied to jitter estimation“ chapter.

The estimate of the jitter standard deviation for the positive kickout pulse is found to be 1.383 ps ± 21 fs (95% confidence interval). The estimate of the jitter standard deviation for the negative kickout pulse is found to be 1.385 ps ± 19 fs and equals that of the positive kickout within the 95% confidence interval.

Figure 3-12 shows the relative difference (in dB) of the amplitude characteristic of the M_{12} measurement[1] after time base drift and time base jitter estimation and ditto compensation, based on the original and the new approach. The new approach gives a slightly larger amplitude at 50 GHz.

Given a difference of 80 mdB at 50 GHz, it can be concluded that the original approach is sufficiently accurate to be used during the actual comparison of the amplitude characteristic of the 50 GHz plug-in.

1. the M_{12} measurement uses plug-in 2 as kickout generator and plug-in 1 as kickout receiver.

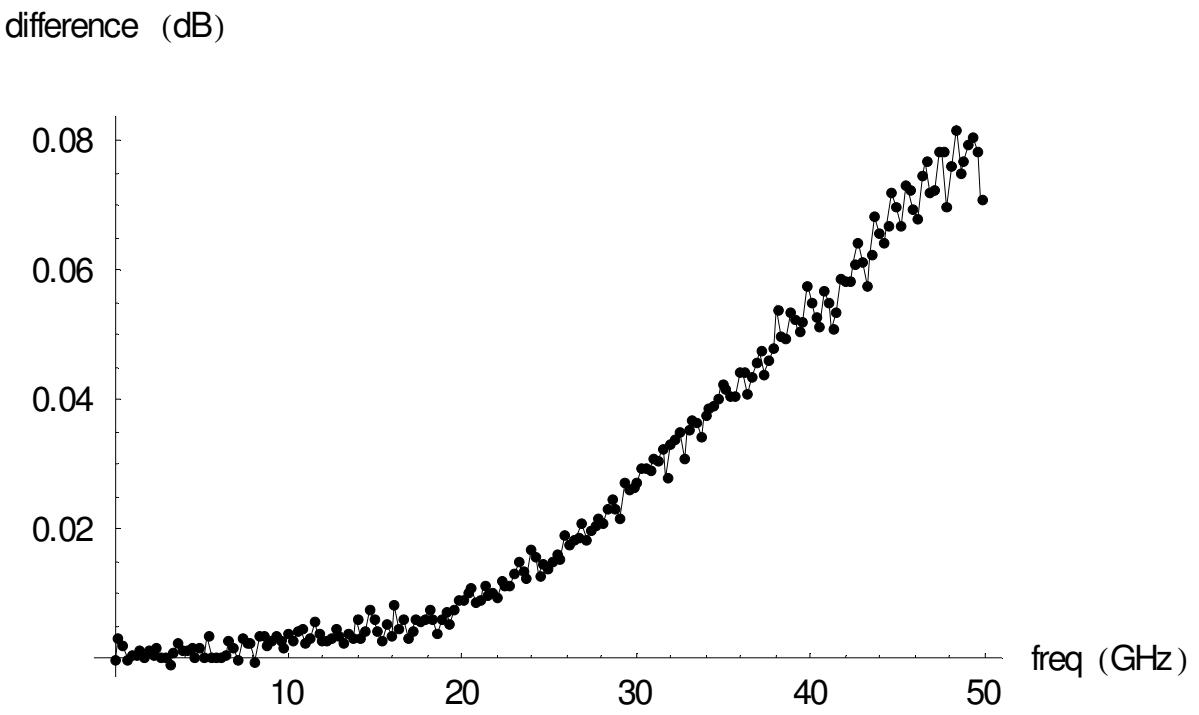

Figure 3-12. Relative difference between the estimated plug-in amplitude characteristic based on the new time base drift and jitter compensation and the original implementation.

Method #4: amplitude characteristic based on EOS

The last method uses an opto-electrical (O/E) converter which is calibrated up to 110 GHz using the electro-optical sampling system [1]-[4] at NIST as a reference element. This O/E converter is then used in the setup described in figure 3-13. The calibrated O/E converter is excited by an optical impulse. After compensating for all time base effects and mismatch effects, the impulse response measured by the sampling scope equals the convolution of the (known) impulse response of the O/E and the (unknown) impulse response of the sampling oscilloscope plug-in. As the impulse response of the O/E is known, the latter can be obtained. The second O/E in the trigger path is solely used to convert the optical pulse into an electrical pulse that can be used to trigger the sampling oscilloscope.

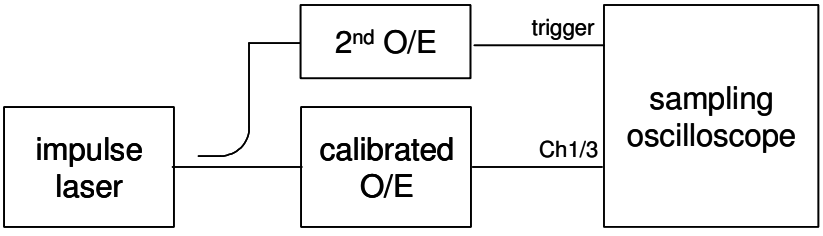

Figure 3-13. Block diagram of the setup used during the sampling oscilloscope calibration using a EOS-calibrated O/E converter.

Again, special attention is given to the time base jitter estimation and compensation.

Figure 3-14 compares the measured and modelled variance, focusing on the main portion of the impulse response. Some discrepancies between measurement and model are visible around 0.395 ns (relative to the start of the acquisition window). In the case of a least squares estimator and a third order model, the estimated jitter standard deviation turns out to be 1.601 ps ± 41 fs. This is significantly larger than the typical jitter standard deviation of about 1 ps which is obtained during other impulse response measurements, based on the same setup. Therefore, it may be possible that there was an issue during the measurement. One possible explanation is the selection of a less optimal setting of the trigger level of the sample scope.

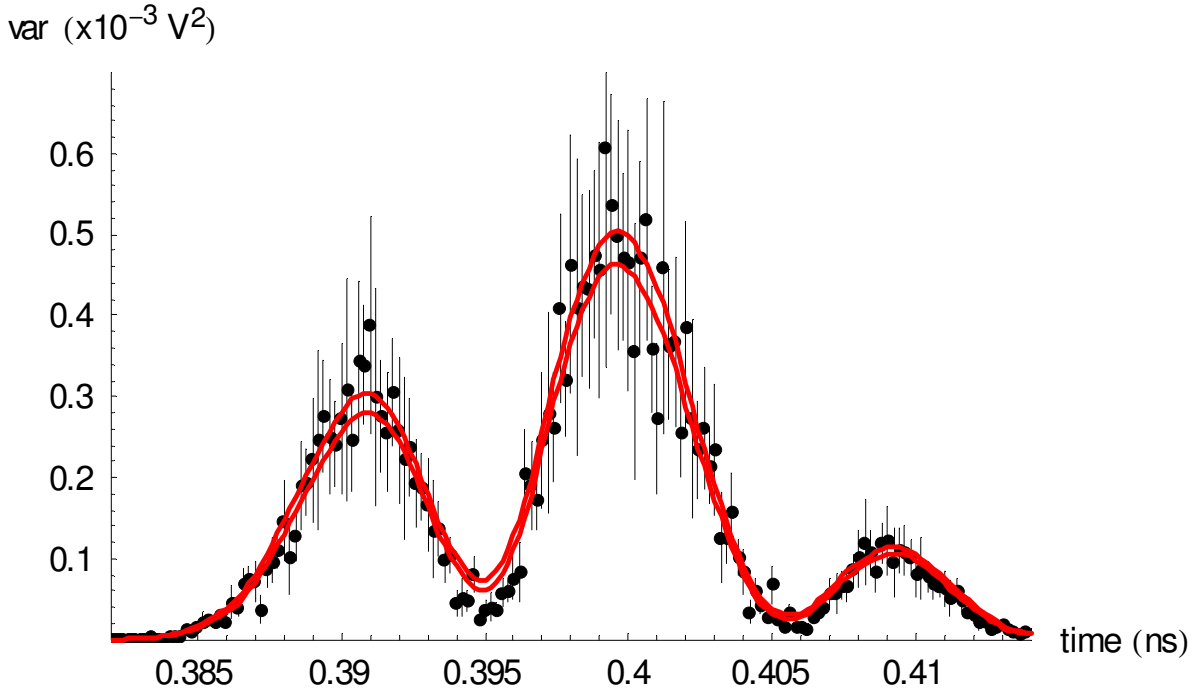

Figure 3-14. Measured (black dots, vertical black lines: 95% confidence interval) and modelled variance (LS estimator, 3rd order model) (solid red curves, boundaries of the 95% confidence interval) corresponding to the main portion of the impulse response of the O/E.

Figure 3-15 shows the modelled variance based on a WLS estimator and a third order model. The corresponding estimated jitter standard deviation is 1.477 ps ± 9 fs.

The LS estimate (figure 3-14) appears to do a better job for the larger values of the variance, while the WLS estimate (figure 3-15) performs better for smaller values of the variance. Taking into account the lower limit of the smallest estimate (1.468 ps) and the upper limit of the largest estimate (1.642 ps) of the jitter standard deviation yields a difference after jitter compensation of ± 0.12 dB at 50 GHz. The time base jitter compensation based on the initial[1] implementation falls within this uncertainty.

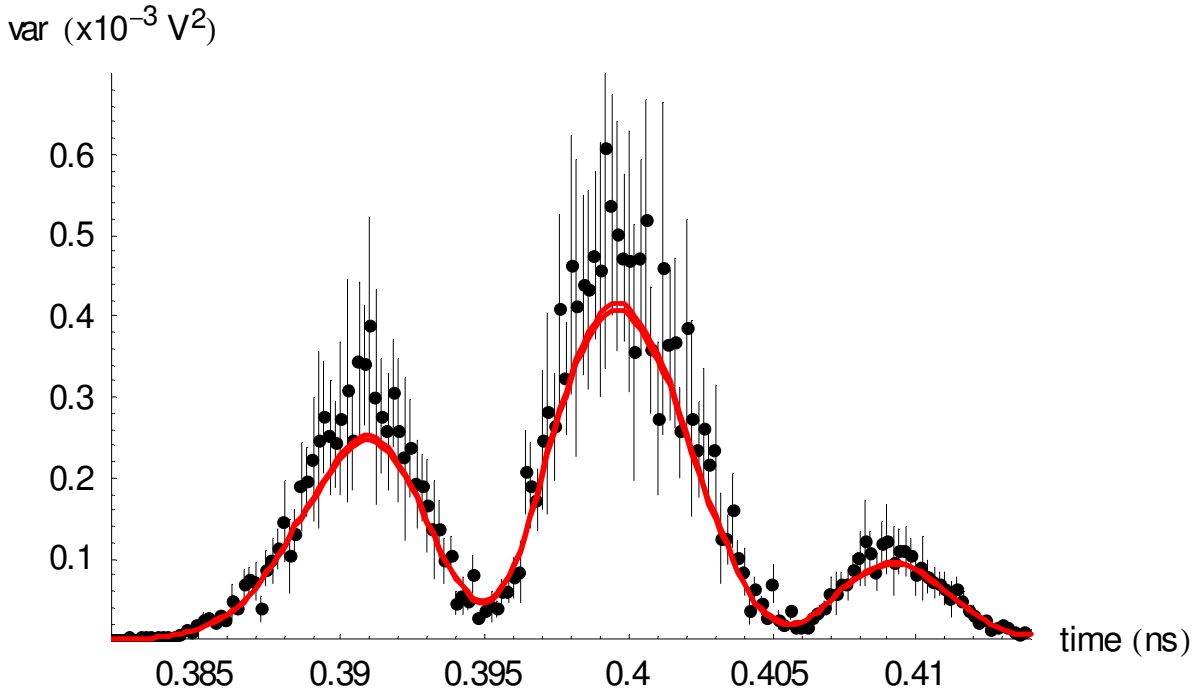

Figure 3-15. Measured (black dots, vertical black lines: 95% confidence interval) and modelled variance (WLS estimator, 3rd order model) (solid red curves, boundaries of the 95% confidence interval) corresponding to the main portion of the impulse response of the O/E.

Summary

Although the comparison does not include any confidence intervals, figure 3-1 on page 3-3 shows good correspondence between all methods, except for the nose-to-nose which seems too yield an amplitude characteristic which is consistently too large. A possible explanation can be found in [9]. The difference between the nose-to-nose and the other methods (see figure 3-2 on page 3-4) is very similar to the discrepancy reported by NIST [6], both in shape and in order of magnitude. The difference is larger than can be contributed to differences in time base drift and time base jitter correction.

1. less correct with respect to time base jitter estimation as explained in the “System identification approach applied to jitter estimation“ chapter.

Phase comparison

Presently, the phase characteristic of the sampling oscilloscope plug-in can only be estimated either using the nose-to-nose calibration technique or by measuring the impulse response of a photodiode (O/E), which itself was calibrated using the EOS system at NIST.

The nose-to-nose calibration procedure is briefly explained in the “Enhancements to the nose-to-nose calibration technique” chapter. The setup and postprocessing required for the second technique is briefly described on page 3-15.

Figure 3-16 shows the discrepancy between the phase characteristics up to 50 GHz, obtained by both techniques for the Agilent 83484A 50 GHz electrical plug-in which was shipped to NIST. A delay is applied such that the phase difference from DC up to 20 GHz falls within the 95% confidence interval[1] provided by NIST [10] for the phase response of the photodiode itself.

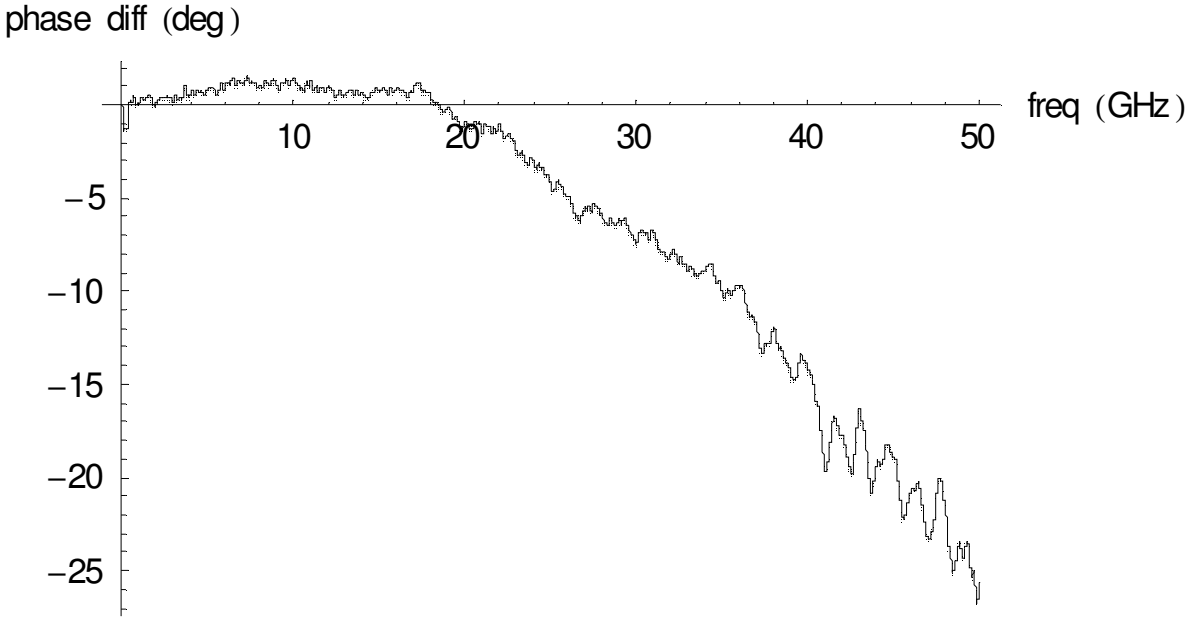

Figure 3-16. Difference (in degrees) between the estimated phase characteristic of the Agilent 83484A 50 GHz electrical plug-in based on the nose-to-nose calibration technique and using the photodiode which was calibrated using the EOS system at NIST.

Although the comparison does not include any confidence intervals, it confirms the discrepancy, which was recently reported [5].

1. ± 1.5 degrees

Conclusions

The performed comparisons confirm the discrepancies which were reported, both with respect to the amplitude characteristic and the phase characteristic of a 50 GHz electrical plug-in, obtained after a nose-to-nose calibration.

Finding an explanation for this discrepancy is left as topic for future research. Fitting a parametric model on the amplitude and phase discrepancy may provide some insight. Also, it makes sense to mention that [11] reports a much better correspondence for the amplitude characteristic of a sampler up to 120 GHz, based on a nose-to-nose calibration and a power measurement.

Eq. 3-6 gives the exact expression of the variance for a sine wave which is disturbed by both additive and jitter noise, in the case the latter has a normal probability density function. It is a rather small effort to provide the exact expression for other distributions, based on [8].

References

[1] D. Williams, P. Hale, T. Clement, and J. Morgan, "Mismatch corrections for electro-optic sampling systems," 56th ARFTG Conference Digest, pp. 141-145, Nov. 30-Dec. 1, 2000
[2] D. Williams, P. Hale, T. Clement, and J. Morgan, "Calibrating electro-optic sampling systems," Int. Microwave Symposium Digest, Phoenix, AZ, pp. 1527-1530, May 20-25, 2001
[3] T. Clement, D. Williams, P. Hale, and J. Morgan, "Calibrating photoreceiver response to 110 GHz," Proc. 15th Annual Meeting, IEEE Lasers and Electro-optics Soc., Glasgow, Scotland, 2002
[4] D. Williams, P. Hale, T. Clement, C-M. Wang, "Uncertainty of the NIST Electro-optic Sampling System," NIST Technical Note 1535, 2004
[5] D. Williams, P. Hale, T. Clement, “Electrical-phase Traceability to NIST’s EOS System,” research update presented at the 4th ARFTG NVNA User’s Forum, June 2004 (http://www.arftg.org/LSNA/4th/UsersForum_June2004_Minutes2.pdf)
[6] P. Hale, T. Clement, K. Coakley, C. Wand, D. DeGroot and A. Verdoni, “Estimating the Magnitude and Phase Response of a 50 GHz Sampling Oscilloscope Using the “Nose-To-Nose” Method”, 55th ARFTG Conf. Digest, June 2000.
[7] Marc Vanden Bossche, private communication 2005.
[8] T. Souders, D. Flach, C. Hagwood and G. Yang, “The Effects of Timing Jitter in Sampling Systems,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 39, No. 1, February 1990
[9] K. Remley, “The Impact of Internal Sampling Circuitry on the Phase Error of the Nose-to-Nose Oscilloscope Calibration”, NIST Technical Note 1528.
[10] T. Clement, P. Hale and D. Williams, “Report of special test (42161S)”, NIST ID number 814627.
[11] J. Scott, "Rapid Millimetre-wave Sampler Response Characterization to Well Beyond 120 GHz Using an Improved Nose-to-nose Method," Int. Microwave Symposium Digest, Philadelphia, pp. 1511-1514, June 2003

CHAPTER 4

System identification approach applied to jitter estimation.

Abstract

Given a symmetrical probability density function, jitter does not introduce phase distortion [1]. However, it has a low-pass effect on the amplitude characteristic. Because the nose-to-nose calibration procedure was mainly used to provide phase information, initially jitter estimation was of less importance.

However, the crossverification of the amplitude distortion of a 50 GHz sampling oscilloscope based on the nose-to-nose calibration technique and the electro-optic sampling system of NIST, justifies additional research with respect to jitter estimation.

A system identification[1] approach is applied to estimate the jitter introduced by a high-frequency sampling oscilloscope.

First, an extended model is proposed to describe the sample variance of a set of repeated (impulse response) measurements in the presence of additive and jitter noise. It is important to remember that the primary goal in this work is to estimate the jitter and as such not the deterministic part of the system, i.e. the impulse response.

Next, the (weighted) least-squares and maximum likelihood estimator are introduced.

Results are shown based on simulations. First, the simulated variance is based on a known model, involving both jitter and additive noise. This allows to test both the correctness of the implementations and to verify the ability to detect model errors. Next, more realistic simulations are performed using “real” jitter. The simulations also allow to study the effect of uncertainties on the input signal. More specifically, first the exact derivatives of the exact signal are used, while in a next step, these derivatives are calculated from the sample mean of the signal.

Finally, the jitter and additive noise standard deviation are estimated on real measurements by performing impulse response measurements using an Agilent 83480A sampling oscilloscope in combination with 83484A 50 GHz electrical plug-ins. Additional challenges, such as the conjugated effect of time base drift and time base distortion, are described and correctly taken care of, demonstrating the real power of a solid stochastical framework.

1. “The aim of identification theory is to provide a systematic approach to fit a mathematical model, as well as possible, to the deterministic part of the system, eliminating the noise distortions as much as possible.” (extracted from “An Introduction to System Identification”, Prof. J. Schoukens, published by the Vrije Universiteit Brussel)

Modeling variance in the presence of additive and jitter noise

High-frequency sampling oscilloscopes often use an equivalent-time sampling principle and suffer from both additive measurement noise $n_y(t_i)$ and timing jitter noise $n_t(t_i)$ at the sampling time instance t_i.

$$y(t_i) = y_0(t_i + n_t(t_i)) + n_y(t_i) \quad \text{Eq. 4-1}$$

$y(t_i)$ represents the measurement of the exact signal $y_0(t_i)$ when both additive noise and jitter are added as part of the measurement. Both $n_y(t_i)$ and $n_t(t_i)$ are assumed to be zero mean, normally distributed[1], independent and stationary with respect to t.

In general it is assumed that the time jitter is small compared to the characteristic time constant[2] of the exact signal $y_0(t_i)$. In that case $y_0(t_i + n_t(t_i))$ is approximated by its first order Taylor series approximation:

$$\tilde{y_1}(t_i) = y_0(t_i) + \left.\frac{dy_0}{dt}\right|_{t = t_i} \cdot n_t(t_i) + n_y(t_i) \quad \text{Eq. 4-2}$$

Given zero mean additive and jitter noise, the expected value of $\tilde{y_1}(t_i)$ equals $y_0(t_i)$ and it makes sense to note that this first order approximation cannot explain the low-pass effect introduced by jitter.

Furthermore, the variance of $\tilde{y_1}(t_i)$ equals

1. Extract of a private communication with Bernie Hovden, Technical Support Engineer Digital Signal Analysis of Agilent Technologies in Santa Rosa to support the assumption of a normal distribution of the jitter in the case of the Agilent 83480A sampling oscilloscope from a hardware point of view: “The inherent trigger jitter in the 83480A basically comes from the translation of amplitude noise (which is typically Gaussian) on the finite rise time trigger source to time in the trigger and time base circuits from the recognition of a trigger event to the firing of the sampler. The decision circuit typically operates over a linear part of the transition so the converted noise should remain Gaussian. Both the trigger and time base circuits have multiple independent stages where the threshold detection takes place sequentially, with no synchronization between the stages. An educated guess is that there are at least 14 separate translations of amplitude to time. The central limit theorem says that a large number of independent events with uncertainties will tend to a Gaussian PDF.”
2. defined as $1/\omega_{-3dB}$ in case of an impulse response or $1/\omega_{max}$ in case of a multi-tone

$$\tilde{\sigma}^2_{y_1}(t_i) = \sigma^2_{n_y} + \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i} \cdot \sigma^2_{n_t}. \qquad \text{Eq. 4-3}$$

Recent work with respect to jitter estimation [2],[3] is based on this first order approximation. Here, it is worthwhile to notice that according to Eq. 4-3, $\sigma^2_y(t_i)$ must be equal to $\sigma^2_{n_y}$ whenever $\left.\frac{dy_0}{dt}\right|_{t=t_i} = 0$. Based on the observation that the latter is not true neither for nose-to-nose nor for other high-frequency impulse response measurements, it is decided to extend the Taylor series approximation to include also the second and third order contributions. As such

$$\tilde{y}_3(t_i) = y_0(t_i) + \sum_{k=1}^{3} \frac{1}{k!} \cdot \left.\frac{d^k y_0}{dt^k}\right|_{t=t_i} \cdot n_t^k(t_i) + n_y(t_i) \qquad \text{Eq. 4-4}$$

Calculating the expected value of $\tilde{y}_3(t_i)$, one finds

$$E\{\tilde{y}_3(t_i)\} = y_0(t_i) + \frac{1}{2} \cdot \left.\frac{d^2 y_0}{dt^2}\right|_{t=t_i} \cdot \sigma^2_{n_t} \qquad \text{Eq. 4-5}$$

A bias now becomes apparent and approximates the low-pass effect introduced by jitter. Let $y_0(t) = A \cdot \sin\omega t$, then $E\{\tilde{y}_3(t_i)\} = A \cdot \left(1 - \frac{\omega^2 \cdot \sigma^2_{n_t}}{2}\right) \cdot \sin\omega t$.

Calculating the variance $\tilde{\sigma}^2_{y_3}(t_i)$ using $E\left\{(\tilde{y}_3(t_i) - E\{\tilde{y}_3(t_i)\})^2\right\}$ and based on the fact that all odd order moments of a normal distribution equal zero, while the fourth order moment equals $3\sigma^4_{n_t}$ and the sixth order moment equals $15\sigma^6_{n_t}$ [4] gives

$$\tilde{\sigma}^2_{y_3}(t_i) = \sigma^2_{n_y} + \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i} \cdot \sigma^2_{n_t} + \left\{\frac{1}{2}\cdot\left(\frac{d^2y_0}{dt^2}\right)^2 + \frac{dy_0}{dt}\cdot\frac{d^3y_0}{dt^3}\right\}\Bigg|_{t=t_i} \cdot \sigma^4_{n_t} + \frac{5}{12}\cdot\left(\frac{d^3y_0}{dt^3}\right)^2\Bigg|_{t=t_i} \cdot \sigma^6_{n_t}$$

Eq. 4-6

Based on Eq. 4-6, $\tilde{\sigma}^2_{y_3}(t_i)$ is now larger than $\sigma^2_{n_y}$ when $\frac{dy_0}{dt}\Big|_{t=t_i}$ equals zero, unless the second and third order derivatives are also both zero.

Remark

The study that follows, refers to different models based on the order of approximation of the Taylor series (Eq. 4-4) instead of the order of contributions of $\sigma^2_{n_t}$ to Eq. 4-6. It should be noticed that for a second order approximation of the Taylor series, the second term in the $\sigma^4_{n_t}$ contribution of Eq. 4-6 is not present.

Estimators

Linear and nonlinear least squares

Starting from N independent and identically distributed (i.i.d.) measurements, one minimizes the following cost with respect to the unknown variances $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$:

$$V_{LS} = \sum_{i=1}^{N} e^2(t_i) \quad , \qquad \text{Eq. 4-7}$$

$$e(t_i) = \frac{\sigma_{t_i}^2 - \tilde{\sigma}_y^2\left(t_i, \sigma_{n_y}^2, \sigma_{n_t}^2, \frac{d^k y_0}{dt^k}\right)}{W_i} . \qquad \text{Eq. 4-8}$$

$\sigma_{t_i}^2$ corresponds to the measured sample variance for $t = t_i$.

$\tilde{\sigma}_y^2\left(t_i, \sigma_{n_y}^2, \sigma_{n_t}^2, \frac{d^k y_0}{dt^k}\right)$ represents the model of the noise variance (see Eq. 4-3 and Eq. 4-6) and W_i is the optional weighting. For the unweighted least squares (LS), W_i is set to 1, while the square root of the sample variance of the sample variance at time instance t_i is used for the weighted least squares (WLS).

Using the first order model (Eq. 4-3) for the noise variance, the error $e(t_i)$ is linear in the unknowns $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$. However, if the model (Eq. 4-6) is expanded towards a second or third order Taylor approximation (Eq. 4-4), the problem is no longer linear in $\sigma_{n_t}^2$.

Maximum Likelihood (ML) estimator

In [2] it is shown that the ML estimator produces statistically more efficient estimates for $\sigma^2_{n_y}(t_i)$ and $\sigma^2_{n_t}(t_i)$ than a linear least squares estimator, assuming that the model equals

$$\tilde{\sigma}^2_{y_1}(t_i) = \sigma^2_{n_y} + \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i} \cdot \sigma^2_{n_t} \qquad \text{Eq. 4-9}$$

and corresponds to the exact model.

This estimator uses the knowledge that if each stochastical variable X_i has a $N(0, 1)$ normal distribution with unity variance and zero mean, then $\sum_{i=1}^{r} X_i^2$ has a χ^2_r chi-squared distribution with r degrees of freedom. If the mean value of X_i is unknown, the (sample) mean has to be calculated and the number of degrees of freedom has to be decreased by 1.

The model described by Eq. 4-6 is now used to extend the ML estimator described in [2].

First, we derive the ML estimator in more detail based on a first order model,

$$\tilde{y}_1(t_i) = y_0(t_i) + \frac{dy_0}{dt}\Bigg|_{t=t_i} \cdot n_t(t_i) + n_y(t_i) \qquad \text{Eq. 4-10}$$

Then the distribution of the modelled output $\tilde{y}_1(t_i)$ is derived, based on the observation that for $n_t(t_i) \sim N(0, \sigma_{n_t})$ and $n_y(t_i) \sim N(0, \sigma_{n_y})$,

$$\tilde{y}_1(t_i) \sim N\left(y_0(t_i), \sqrt{\sigma^2_{n_y} + \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i} \cdot \sigma^2_{n_t}}\right). \qquad \text{Eq. 4-11}$$

The distribution can then be normalized as follows:

$$y_N(t_i) = \frac{\tilde{y_1}(t_i) - y_0(t_i)}{\sqrt{\sigma_{n_y}^2 + \left(\frac{dy_0}{dt}\right)^2\Big|_{t=t_i} \cdot \sigma_{n_t}^2}} \sim N(0, 1) \qquad \text{Eq. 4-12}$$

Acquiring K realizations of y at $t = t_i$, k representing the realization index, it follows that

$$\sum_{k=1}^{K} y_N^2(t_i, k) = \sum_{k=1}^{K} \frac{\{y(t_i, k) - y_0(t_i)\}^2}{\sigma_{n_y}^2 + \left(\frac{dy_0}{dt}\right)^2\Big|_{t=t_i} \cdot \sigma_{n_t}^2} \sim \chi_K^2 . \qquad \text{Eq. 4-13}$$

Eq. 4-13 assumes that $y_0(t_i)$ is known.

Using the sample mean and sample variance $\sigma_{t_i}^2$ of $y(t_i)$

$$\sigma_{t_i}^2 = \frac{1}{K-1} \cdot \sum_{k=1}^{K} \left\{ y(t_i, k) - \frac{1}{K} \cdot \sum_{l=1}^{K} y(t_i, l) \right\}^2 , \qquad \text{Eq. 4-14}$$

Eq. 4-13 becomes

$$\sum_{k=1}^{K} y_N^2(t_i, k) = \frac{(K-1) \cdot \sigma_{t_i}^2}{\sigma_{n_y}^2 + \left(\frac{dy_0}{dt}\right)^2\Big|_{t=t_i} \cdot \sigma_{n_t}^2} \sim \chi_{(K-1)}^2 . \qquad \text{Eq. 4-15}$$

Let $n = K - 1$ to simplify the notation.

In order to derive the log likelihood function, one starts from the chi-squared probability distribution function with n degrees of freedom

$$F(\chi^2) = \int_0^{\chi^2} \frac{1}{2^{n/2} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot x^{\left(\frac{n}{2}-1\right)} \cdot e^{\frac{-x}{2}} dx. \qquad \text{Eq. 4-16}$$

Here $x = \dfrac{n \cdot \sigma_{t_i}^2}{\sigma_{n_y}^2 + \left(\dfrac{dy_0}{dt}\right)^2\Big|_{t=t_i} \cdot \sigma_{n_t}^2}$.

First the integrand of Eq. 4-16 is evaluated.

Substitution of $x' = \sigma_{t_i}^2$ and $\alpha = \sigma_{n_y}^2 + \left(\dfrac{dy_0}{dt}\right)^2\Big|_{t=t_i} \cdot \sigma_{n_t}^2$, α representing the model, leads to

$$\frac{1}{2^{n/2} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot x^{\left(\frac{n}{2}-1\right)} \cdot e^{\frac{-x}{2}} dx = \frac{1}{2^{n/2} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot \left(\frac{n \cdot x'}{\alpha}\right)^{\left(\frac{n}{2}-1\right)} \cdot e^{\frac{-n \cdot x'}{2\alpha}} \cdot \frac{n}{\alpha} \cdot dx'.$$

When maximizing the log likelihood function with respect to the model parameters, the constant terms can be omitted. As such, the log likelihood function to be maximized equals

$$\ln f = -\left(\frac{n}{2}-1\right) \cdot \ln\alpha - \frac{n \cdot x'}{2\alpha} - \ln\alpha = -\frac{n}{2} \cdot \left(\ln\alpha + \frac{x'}{\alpha}\right). \qquad \text{Eq. 4-17}$$

Based on N i.i.d. measurements, the following cost needs to be minimized with respect to $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$:

$$V_{ML} = \frac{n}{2} \cdot \sum_{i=1}^{N} \left\{ \ln\left(\sigma_{n_y}^2 + \left(\frac{dy_0}{dt} \right)^2 \Big|_{t=t_i} \cdot \sigma_{n_t}^2 \right) + \frac{\sigma_{t_i}^2}{\sigma_{n_y}^2 + \left(\frac{dy_0}{dt} \right)^2 \Big|_{t=t_i} \cdot \sigma_{n_t}^2} \right\} .$$ Eq. 4-18

As such, the MLE cost in case of a third order model becomes $\frac{n}{2} \cdot \sum_{i=1}^{N} \left\{ \ln\beta + \frac{\sigma_{t_i}^2}{\beta} \right\}$ where β must be substituted using Eq. 4-6.

Also, it is straightforward to extend Eq. 4-18 to deal with situations where the number of degrees of freedom varies with i. The relevance of this extension will become clear in "*Step 4: influence of time base drift"* on page 4-44. Eq. 4-18 becomes

$$V_{ML} = \sum_{i=1}^{N} \left\{ \frac{n_i}{2} \cdot \left[\ln\left(\sigma_{n_y}^2 + \left(\frac{dy_0}{dt} \right)^2 \Big|_{t=t_i} \cdot \sigma_{n_t}^2 \right) + \frac{\sigma_{t_i}^2}{\sigma_{n_y}^2 + \left(\frac{dy_0}{dt} \right)^2 \Big|_{t=t_i} \cdot \sigma_{n_t}^2} \right] \right\} ,$$ Eq. 4-19

where n_i corresponds to the number of degrees of freedom for $t = t_i$.

Curiosity

During simulations it was found that the second term contributing to the MLE cost in case of a first order model (Eq. 4-18) always equals N. At first, this was believed to be a programming error. However, it is proven here that in the solution, this contribution always equals N.

Let $x_i = \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i}$, while a and b are estimates of $\sigma_{n_y}^2$ and $\sigma_{n_t}^2$ by minimizing the cost (Eq. 4-18).

Then Eq. 4-18 becomes $\sum_{i=1}^{N}\left\{\ln(a+b\cdot x_i)+\frac{\sigma_{t_i}^2}{a+b\cdot x_i}\right\}$ to be minimized with respect to a and b.

$$\frac{\partial}{\partial a}=0\Leftrightarrow\sum_{i=1}^{N}\left\{\frac{1}{a+b\cdot x_i}-\frac{\sigma_{t_i}^2}{(a+b\cdot x_i)^2}\right\}=0 \qquad \text{Eq. 4-20}$$

$$\frac{\partial}{\partial b}=0\Leftrightarrow\sum_{i=1}^{N}\left\{\frac{x_i}{a+b\cdot x_i}-\frac{\sigma_{t_i}^2\cdot x_i}{(a+b\cdot x_i)^2}\right\}=0 \qquad \text{Eq. 4-21}$$

Multiplying Eq. 4-20 by a and Eq. 4-21 by b and adding both equations, the sum is zero in the solution.

$\sum_{i=1}^{N}\left\{1-\frac{\sigma_{t_i}^2}{a+b\cdot x_i}\right\}=0$, which proves that $\sum_{i=1}^{N}\frac{\sigma_{t_i}^2}{\sigma_{n_y}^2+\left(\frac{dy_0}{dt}\right)^2\Big|_{t=t_i}\cdot\sigma_{n_t}^2}=N.$

This is exactly the second term in the cost function (Eq. 4-18).

Generation of simulation data

First the different models and estimators are tested using simulation data.

This data should be kept as realistic as possible. Therefore, the combined impulse response of a real-world opto-electrical converter (O/E) and a 50 GHz sampling oscilloscope is used as a starting point. Figure 4-1 shows the block diagram of the required setup. The second O/E in the trigger path is solely used to convert the optical pulse into an electrical pulse that can be used to trigger the sampling oscilloscope.

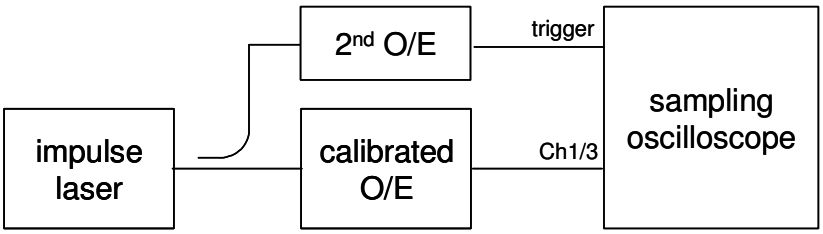

Figure 4-1. Block diagram of the setup used during the impulse measurement.

For the measurement of the impulse response, a time record of 5 ns is used starting at 143 ns and 500 records of 4096 points are acquired. The data are corrected for time base drift and time base distortion. The resulting averaged impulse response is shown in figure 4-2. Figure 4-3 shows the same information on a logarithmic scale.

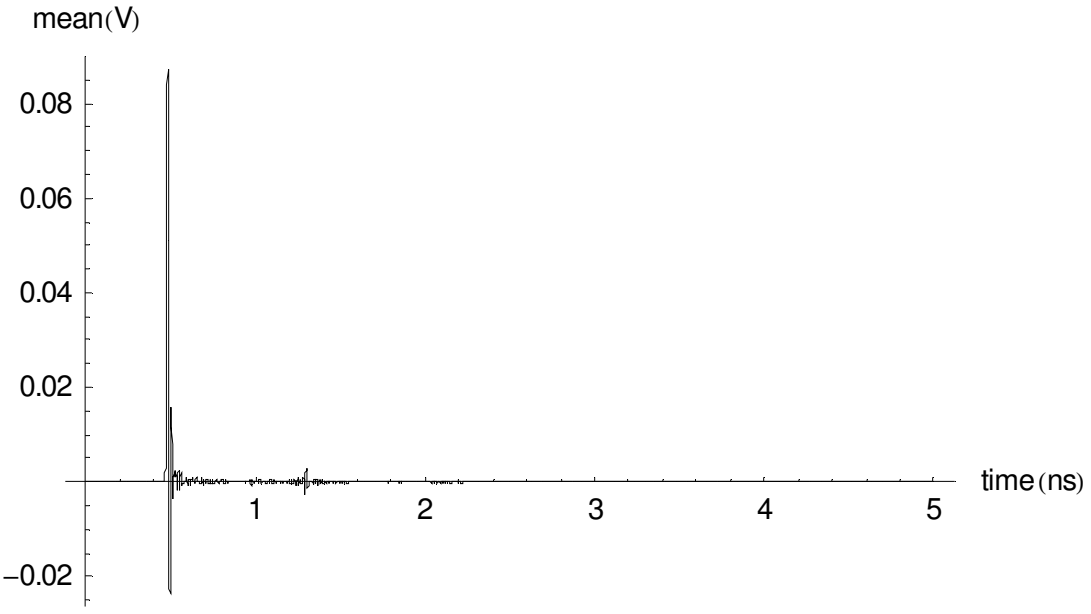

Figure 4-2. Averaged impulse response.

Figure 4-4 zooms in to the main portion of the averaged pulse and its corresponding sample variance. It is clearly shown that at the time instants where the averaged pulse

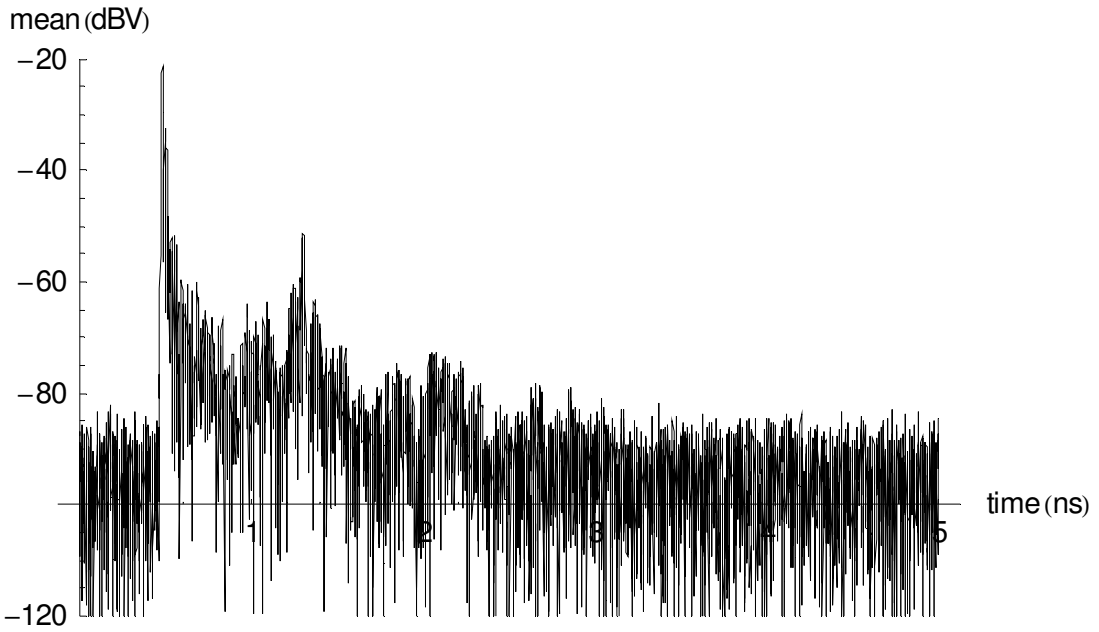

Figure 4-3. Averaged impulse response (logarithmic scale).

has a zero slope, the variance is larger than the constant level at both sides of the pulse, which corresponds to the variance of the additive noise. This means that Eq. 4-9 does not correspond to the exact model for the measured jittered signal.

To further increase the S/N ratio of the test signal, a rectangular window is first applied in the frequency domain. All frequency components up to the first transmission zero (located at about 70 GHz) are kept. The other lines are set to zero. In order to remove the small ringing at the edges in the time domain, an ad hoc window is applied there too.

The corresponding analytical expression for the time signal is then given by its Fourier series

$$x(t) = re\left\{ \sum_{m=0}^{M} X(m) \cdot e^{j2\pi m \cdot \Delta f \cdot t} \right\} \qquad \text{Eq. 4-22}$$

and allows to calculate the exact derivatives. $re(x)$ represents the real part of the complex value x. If the number of relevant spectral lines M becomes too large, the calculation of $x(t)$ using Eq. 4-22 becomes very time consuming and a fast implementation of the inverse Fourier transform is used instead. However, when simulating jitter noise, the time samples are no longer on an equidistant grid. In order to avoid the calculation using Eq. 4-22, a two step approach is used. First, the equidistant $x(t)$ is evaluated on a sufficiently oversampled time grid and then cubic

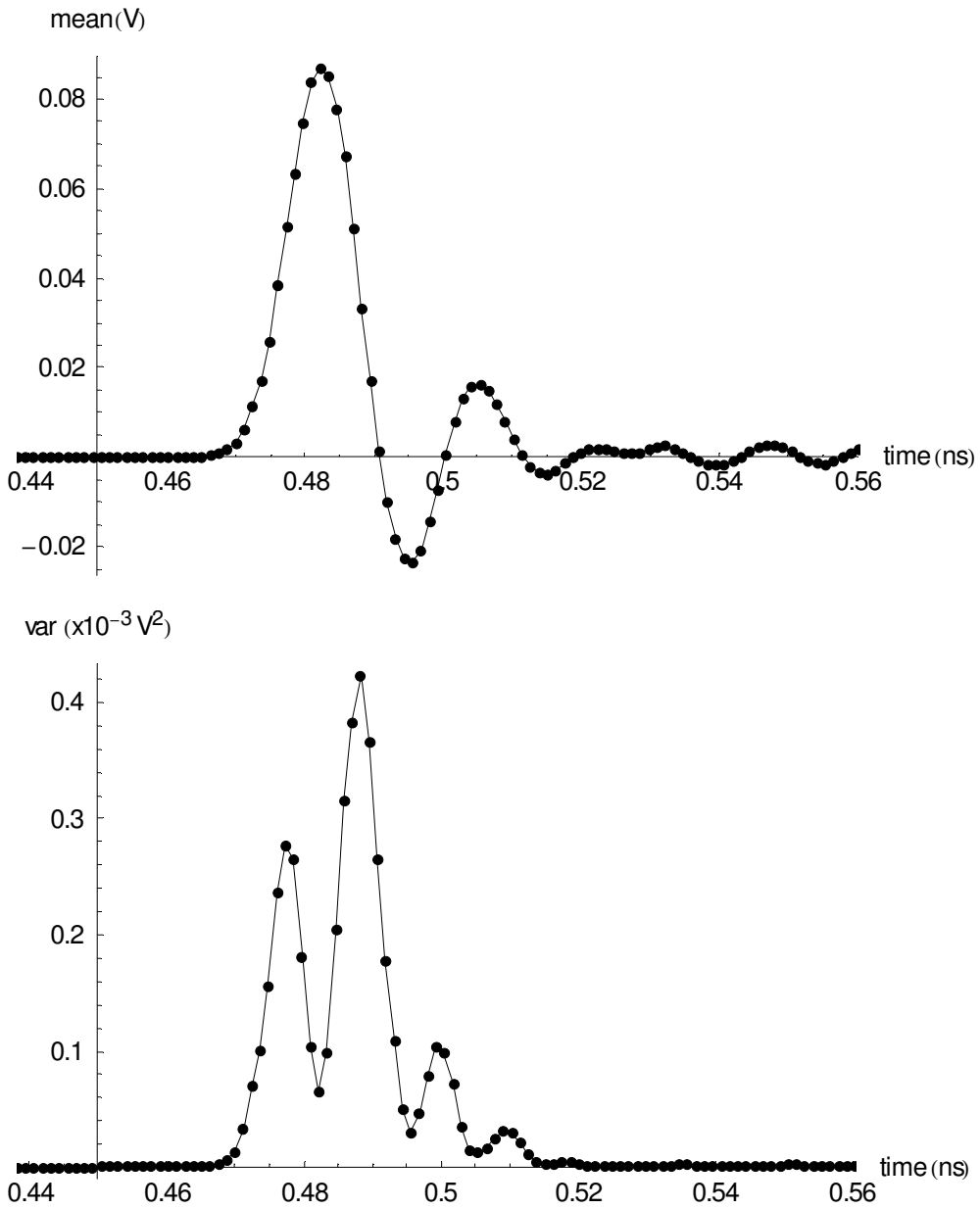

Figure 4-4. Zooming into the main portion of the averaged impulse response and its variance.

interpolation is used to obtain the value at $x(t + n_t(t))$. Given the above impulse response, it was found that oversampling by a factor of 128 in combination with cubic interpolation leads to an RMS value for the difference between the exact and interpolated signal that is about 200 dB down with respect to the RMS value of the signal.

Random number generator

Another critical element during simulation is the random number generator that is used to generate the noisy simulations. It was found that it makes sense to verify the fitness of the generator before questioning the correctness of the implementation of the model extraction. The latter is implemented in C in order to keep the simulation time reasonable. For performance reasons, the Vector Statistical Library [5] included in the Intel® Math Kernel Library is used.

Initially, the sample variance of the variance of the additive noise, estimated using both the MLE and WLS for different values of the variance of the jitter noise, consistently turned out to be approximately 50% smaller than the variance predicted by the corresponding element of the parameter covariance matrix.

It was found that this artifact was caused by a poor selection of the underlying basic random number generator. Replacing the 31-bit multiplicative congruential generator[1] (MCG31m1) by a combined multiple recursive generator with two components of order 3 (MRG32k3a) [5], this problem was solved. The main problem with the original basic random number generator is its relatively small period length with respect to the relatively large sampling.

The fitness of the random number generator is tested by calculating the sample mean and sample variance of N_e realizations of the following cost

$$V = \sum_{j=1}^{N_t} \frac{(\overline{z_{N_w,j}} - N_s)^2}{S^2_{N_w,j}} \qquad \text{Eq. 4-23}$$

where $\overline{z_{N_w,j}}$ and $S^2_{N_w,j}$ respectively represent the sample mean and sample variance of N_w realizations of $z_j = \sum_{i=1}^{N_s} x_{ij}^2$ with $x_{ij} \sim N(0,1)$, $i = 1 \,..\, N_s$ and $j = 1 \,..\, N_t$. As such $z_j \sim \chi^2_{N_s}$ and its expected value $E\{z_j\} = N_s$.

1. According to Intel MKL support, "MCG31 has a rather short period (~2^{32}) and is not recommended for applications demanding a large volume of random numbers."
The number of samples used during step 1 of the simulations (see page 4-21) is 100 x 4096 x 200 x 250 ~= 2x 10^{10} while 2^{32} ~= 4x 10^9.
The MRG32k3a pseudo-random generator has a period of ~2^{191} (~3x 10^{57}).

If N_s is chosen large enough, then $z_j \sim N(N_s, \sqrt{2N_s})$, its sample mean and sample variance are known to be independent and [6] can be used to verify the properties of the sample mean and sample variance based on Eq. 4-23.

$$\mu_V = E\{V\} = \frac{N_w - 1}{N_w - 3} \cdot N_t \qquad \text{Eq. 4-24}$$

$$\sigma_V^2 = E\{(V - E\{V\})^2\} = \frac{(N_w - 1)^3}{(N_w - 3)^2 \cdot (N_w - 5)} \cdot 2N_t \qquad \text{Eq. 4-25}$$

Eq. 4-24 and Eq. 4-25 represent the increase of the expected value and variance of the cost when replacing the exact variance of z_j by its sample variance based on N_w real (non-complex) data sets.

If N_t is chosen sufficiently large, then

$$V \sim N\left(\frac{N_w - 1}{N_w - 3} \cdot N_t, \sqrt{\frac{(N_w - 1)^3}{(N_w - 3)^2 \cdot (N_w - 5)} \cdot 2N_t}\right) \qquad \text{Eq. 4-26}$$

and one can calculate the 95% confidence intervals of its sample mean and sample variance using the Student-t and $\chi^2_{(N_e - 1)}$ distribution for N_e realizations [7].

It was found that using the Box-Müller transformation [5] in combination with the 31-bit multiplicative congruential generator (MCG31m1) of the Intel® Math Kernel Library did not pass the 95% confidence interval test for $N_s = 100$, $N_w = 200$, $N_t = 400$ and $N_e = 500$, while the combined multiple recursive generator with two components of order 3 (MRG32k3a) did. As such the latter is used during simulation.

Step 1: third order approximation of variance, known derivatives

In order to test the correctness of the model parameter extraction software, simulation data is generated using Eq. 4-6. The required derivatives are based on Eq. 4-22 and are assumed to be known exactly. In step 2, the variance is based on additive noise and "real" jitter noise, starting from Eq. 4-1. In step 3, the derivatives are no longer assumed to be known, but will be estimated based on the sample mean of the pulse and the estimated jitter standard deviation.

Based on measurements using the Agilent 83480A Digital Communications Analyzer, the standard deviation of the additive noise during the simulations was set to 0.6 mV, while the jitter standard deviation was stepped from 0 to 2 ps in 0.2 ps steps.

Model parameters are extracted for a maximum likelihood estimator, a least squares and a weighted least squares estimator. The model is based on portions of Eq. 4-6 corresponding to a first, second and third order approximation of Eq. 4-4.

Simulation data for the sample variance is obtained from a χ^2 - distributed random variable with a number of degrees of freedom of 100. Each trace contains 4096 time points. This process is repeated 200 times to estimate the sample variance of the sample variance. The set of 200 repeated simulations is used to estimate the variance of the additive and jitter noise simultaneously. In turn, this estimation is repeated 250 times.

Estimated jitter standard deviation

Figure 4-5 up to figure 4-7 show the sample mean of the absolute error of the estimated jitter standard deviation. The absolute error e_{n_t} is defined as the estimated jitter standard deviation minus the exact jitter standard deviation and can be both positive and negative.

$$e_{n_t} = \hat{\sigma}^2_{n_t} - \sigma^2_{n_{t_0}} \qquad \text{Eq. 4-27}$$

Based on the fact that the simulation data are generated using Eq. 4-6, corresponding to the third order Taylor approximation of Eq. 4-4, within the uncertainty of the parameters, one expects to find the exact values using a third order approximation. Lower order approximations are expected to perform well for small jitter values and to show deviations for larger jitter values.

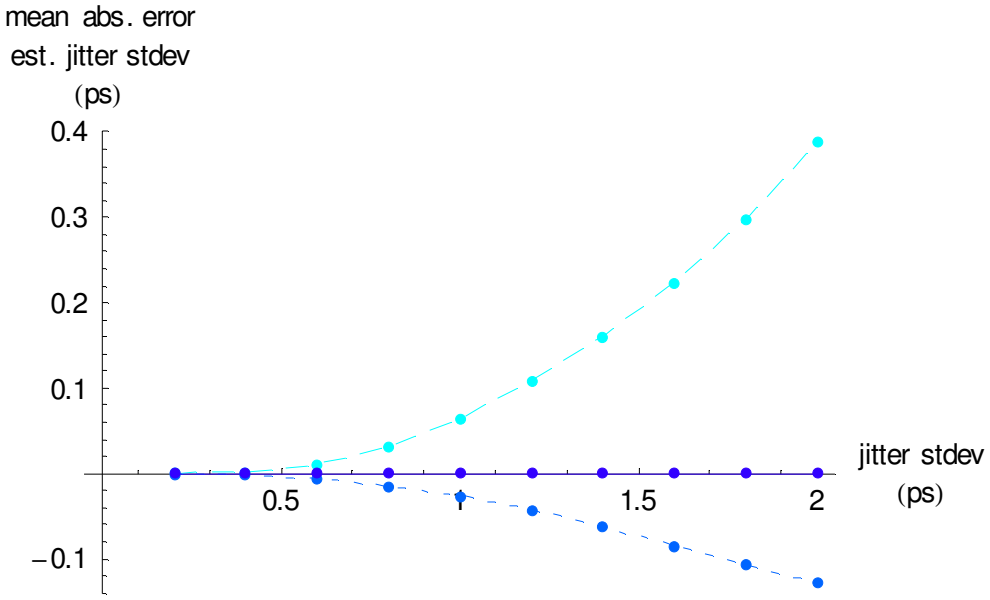

Figure 4-5. Mean absolute error of the estimated standard deviation of the jitter noise using the ML estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

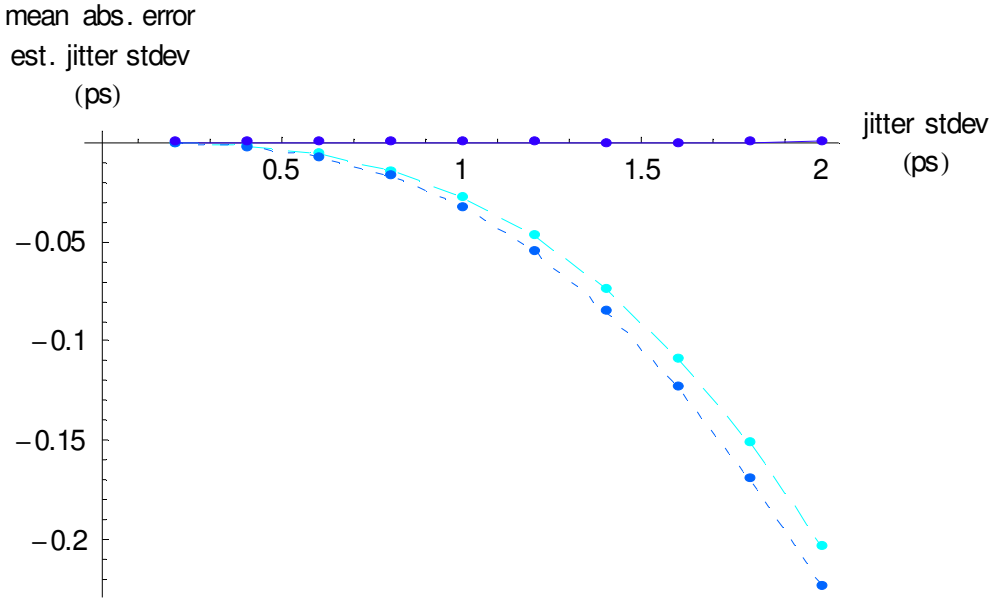

Figure 4-6. Mean absolute error of the estimated standard deviation of the jitter noise using the LS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

Estimated additive noise standard deviation

Figure 4-8 on page 4-19 up to figure 4-10 show the estimated standard deviation of the additive noise. The exact value is 0.6 mV. It makes sense to notice the small bias

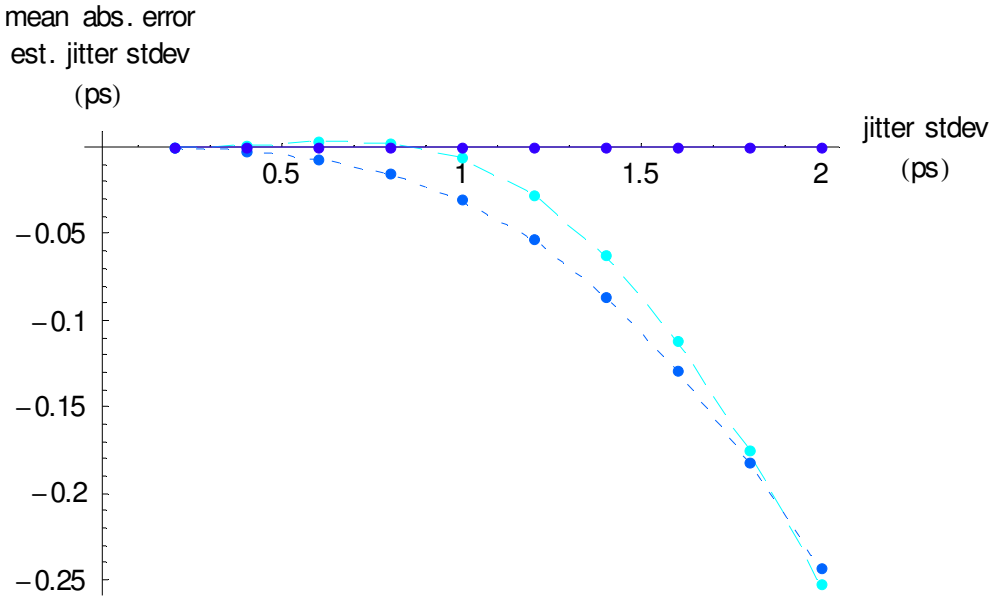

Figure 4-7. Mean absolute error of the estimated standard deviation of the jitter noise using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

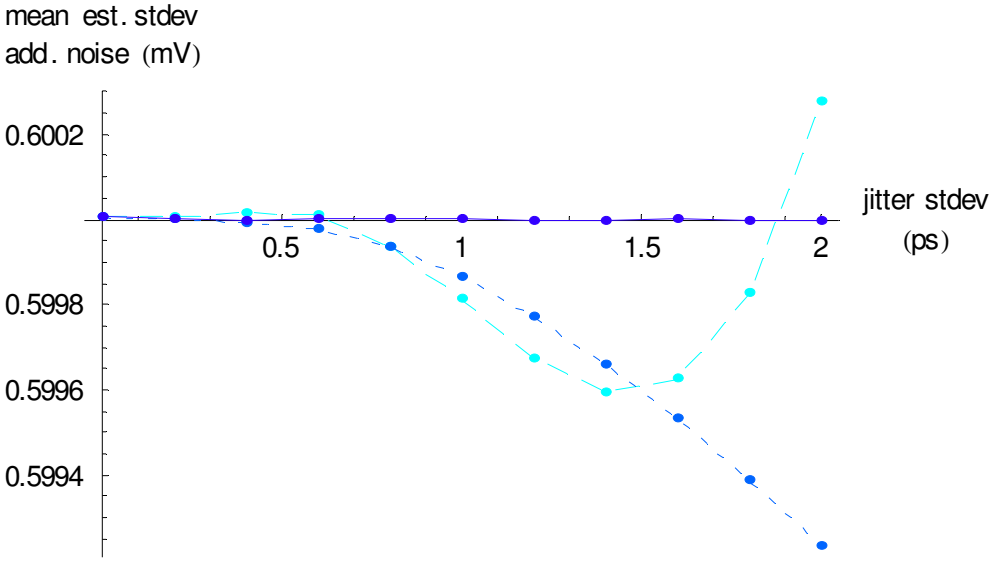

Figure 4-8. Mean estimated standard deviation of the additive noise using the ML estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

present in the WLS estimate. This bias turns out to be independent of the jitter value. The relative error of the estimated variance is empirically found to equal $\frac{4}{N_s \cdot N_w}$, where N_s equals the number of degrees of freedom of the sampled χ^2 - distribution,

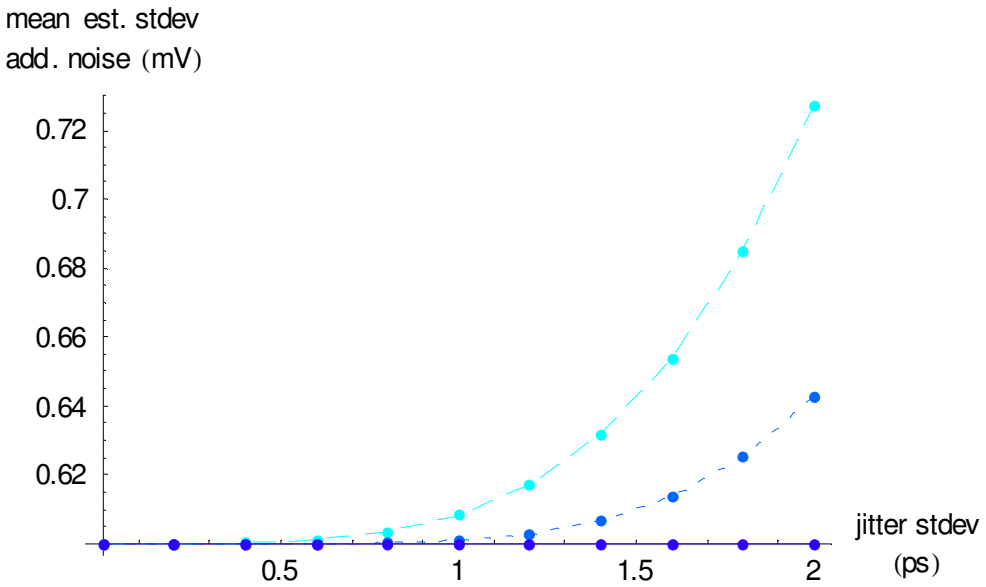

Figure 4-9. Mean estimated standard deviation of the additive noise using the LS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

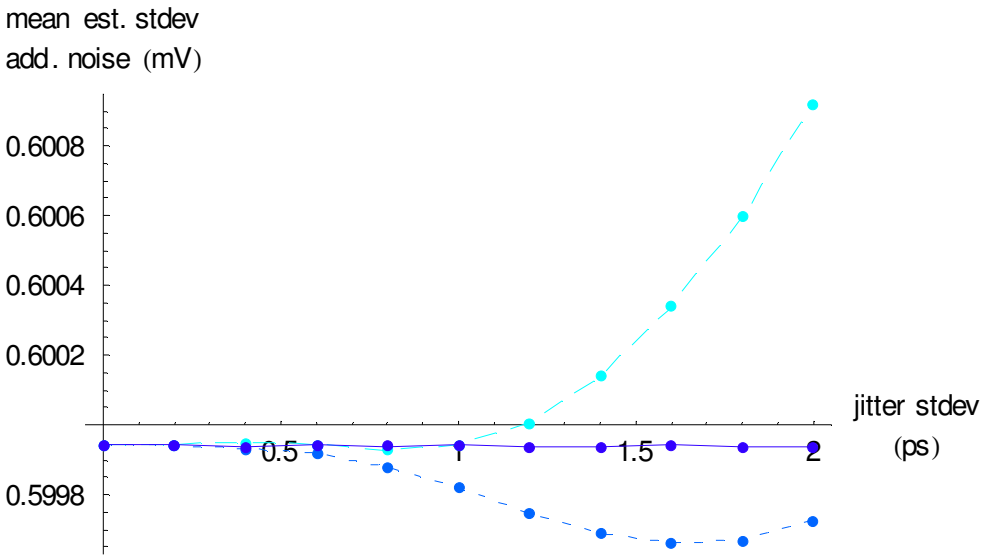

Figure 4-10. Mean estimated standard deviation of the additive noise using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

and N_w is the number of repeated realizations. As such $N_s \cdot N_w$ equals the overall number of degrees of freedom of the averaged data used to estimate both parameters. During this simulation $N_s = 100$ and $N_w = 200$. The corresponding relative error is $2 \cdot 10^{-4}$, the estimated standard deviation is

$0.6 \cdot \sqrt{1 - 2 \cdot 10^{-4}} = 0.59994$ mV. This value corresponds to the small offset visible for the third order model in figure 4-10.

Value of the cost function

The value of the cost function in case of a weighted least-squares deserves special attention, because one can calculate the expected value of the cost and its 95% confidence interval. Figure 4-11 shows the cost for the first, second and third order model. It clearly demonstrates the sensitivity of the cost with respect to model errors. Figure 4-12 and figure 4-13 on page 4-23 zoom into the sample mean and sample variance of the cost when there are no model errors.

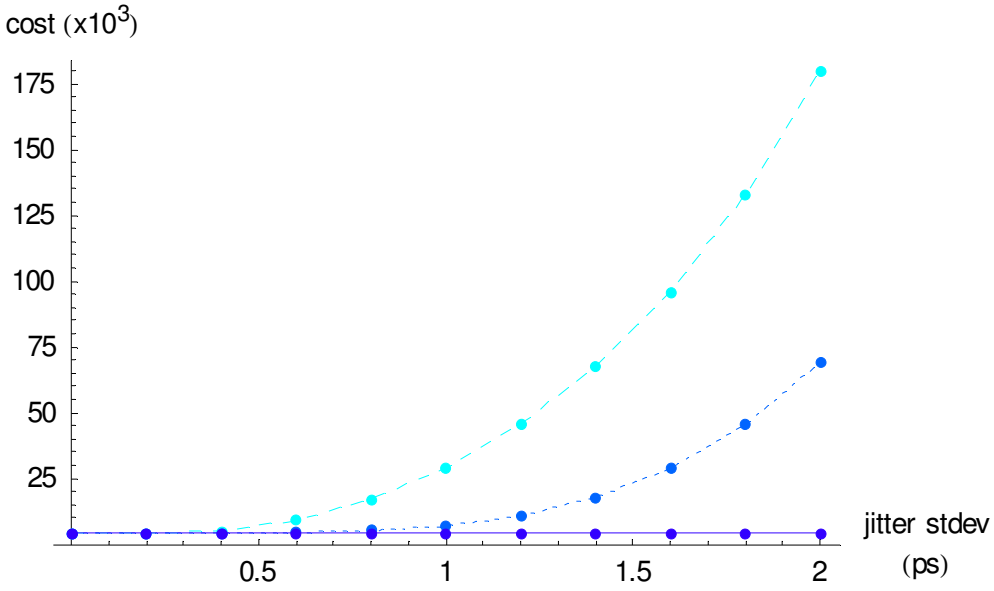

Figure 4-11. Mean value of the cost using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

In order to obtain the expected value and the variance of the cost, one can reuse the explanation found under "*Random number generator*" on page 4-15, taking into account that the number of degrees of freedom is decreased by the number of parameters ($p = 2$). During the simulation, the following values were used: $N_s = 100$, $N_t = 4096$, $N_w = 200$ and $N_e = 250$. Replacing N_t by $N_t - p$ in Eq. 4-26, the expected value of the cost μ_V is found to be 4136. The variance of the cost σ_V^2 equals 8526. Based on N_e realizations of the cost, one can calculate the 95% confidence intervals of its sample mean $\overline{V_{N_e}}$ and sample variance $S_{N_e}^2$ using the

Student-t and $\chi^2_{(N_e-1)}$ - distribution:

$$P\left(\frac{(N_e-1)\,S^2_{N_e}}{\chi^2_{N_e-1,\,1-\frac{\alpha}{2}}} \le \sigma^2_V \le \frac{(N_e-1)\,S^2_{N_e}}{\chi^2_{N_e-1,\,\frac{\alpha}{2}}}\right) = 1-\alpha \qquad \text{Eq. 4-28}$$

$$P\left(\overline{V_{N_e}} - \frac{S_{N_e}}{\sqrt{N_e}}\,t_{N_e-1,\,1-\frac{\alpha}{2}} \le \mu_V \le \overline{V_{N_e}} + \frac{S_{N_e}}{\sqrt{N_e}}\,t_{N_e-1,\,1-\frac{\alpha}{2}}\right) = 1-\alpha \qquad \text{Eq. 4-29}$$

For the actual simulation parameters, the 95% confidence intervals become:

$$0.845\,S^2_{N_e} \le \sigma^2_V \le 1.202\,S^2_{N_e} \qquad \text{Eq. 4-30}$$

$$\overline{V_{N_e}} - 1.970\sqrt{\frac{S^2_{N_e}}{N_e}} \le \mu_V \le \overline{V_{N_e}} + 1.970\sqrt{\frac{S^2_{N_e}}{N_e}} \qquad \text{Eq. 4-31}$$

The 95% confidence intervals are shown in figure 4-12 and figure 4-13.

Based on the simulation results it can be concluded that the estimated parameters converge to the exact parameters when there are no model errors. A small bias becomes apparent for the estimated standard deviation of the additive noise in case of the weighted least-squares estimator. This bias decreases as function of an increasing number of averages[1] used while estimating the standard deviation of the additive noise.

In case of a WLS estimator, it is demonstrated that both the sample mean and sample variance of the cost match their expected value within their 95% confidence intervals when there are no model errors, while there is a significant difference when there are model errors. This clearly shows the capability of a WLS estimator with respect to model selection.

1. Corresponding to a χ^2 - distribution with an increasing number of degrees of freedom, as such converging to a normal distribution.

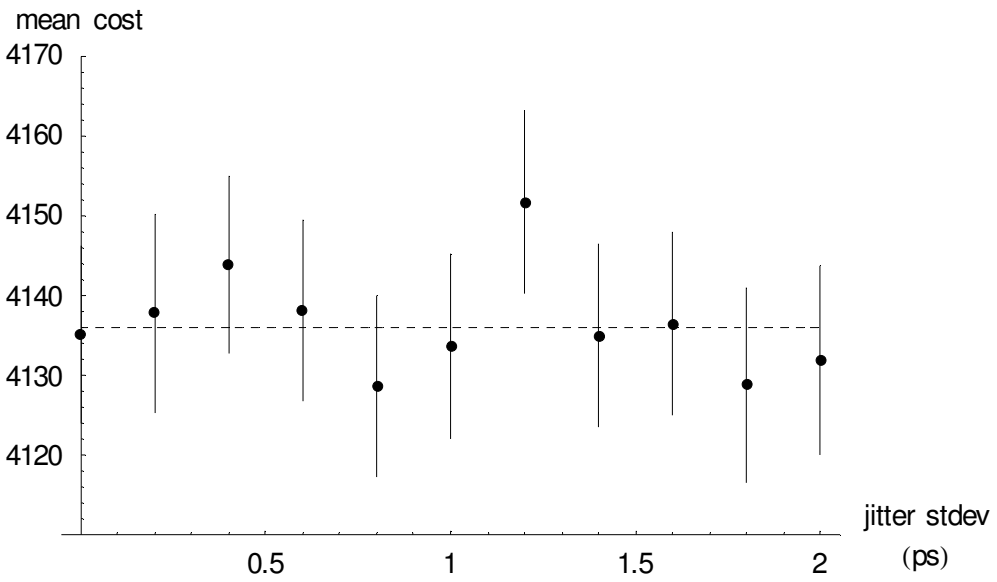

Figure 4-12. Expected value (dashed line) of the cost of the 3rd order model using the WLS estimator, compared to its sample mean value (250 realizations) and its 95% confidence interval.

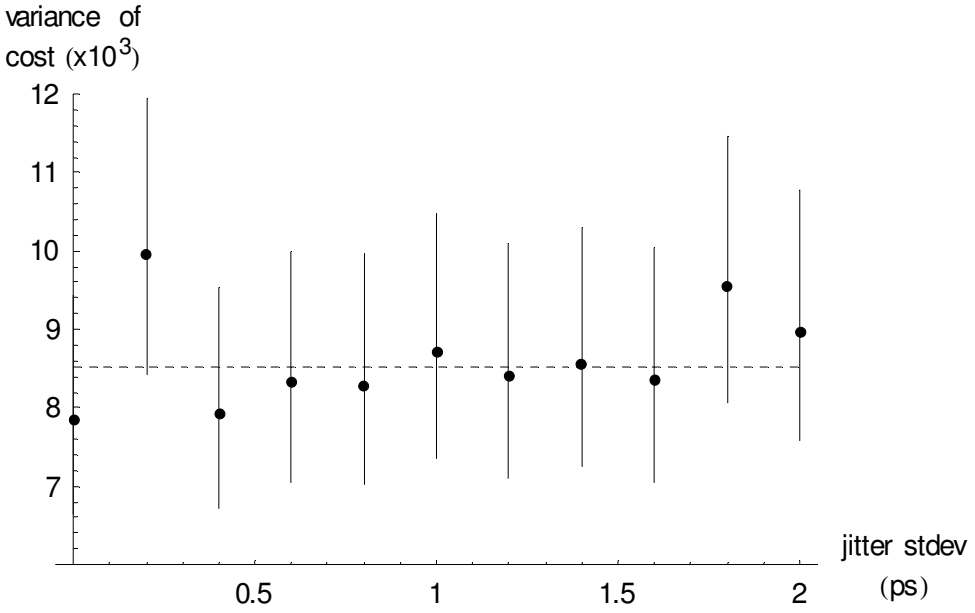

Figure 4-13. Expected variance (dashed line) of the cost of the 3rd order model using the WLS estimator, compared to its sample variance value (250 realizations) and its 95% confidence interval.

Using the covariance matrix of the parameters

Finally it is verified that the uncertainty on the parameters, indicated by the parameter covariance matrix, corresponds to the sample variance of the parameters based on N_e repeated estimations.

In order not to clutter the figures, the 95% confidence interval of the sample variance of the estimated variances σ is not shown.

For $N_e = 250$ estimations, $0.845\, S_{N_e}^2 \leq \sigma^2 \leq 1.202\, S_{N_e}^2$.

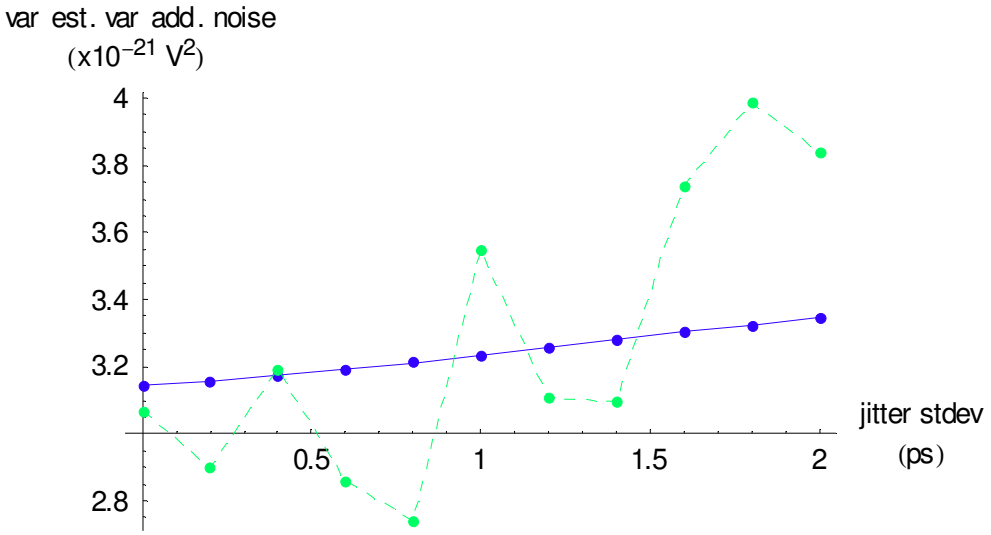

Figure 4-14. Variance of the estimated variance of the additive noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 250 estimations), 3rd order model using the WLS estimator.

Figure 4-14 shows the comparison of the estimated variance of the additive noise for a third order model in combination with the WLS estimator. Figure 4-15 repeats this comparison for the estimated variance of the jitter noise for the same model and estimator. Finally, figure 4-16 shows that the uncertainty on the estimated jitter variance is indeed smaller using a WLS estimator instead of a (unweighted) LS estimator.

It can be concluded that the parameter covariance matrix can be used to obtain an estimate of the uncertainty on the estimated parameters. As such, one does not have to perform repeated estimations in order to get an idea of the uncertainty on the estimated parameters.

An overview of the simulation results can be found in Table 4-1 and Table 4-2.

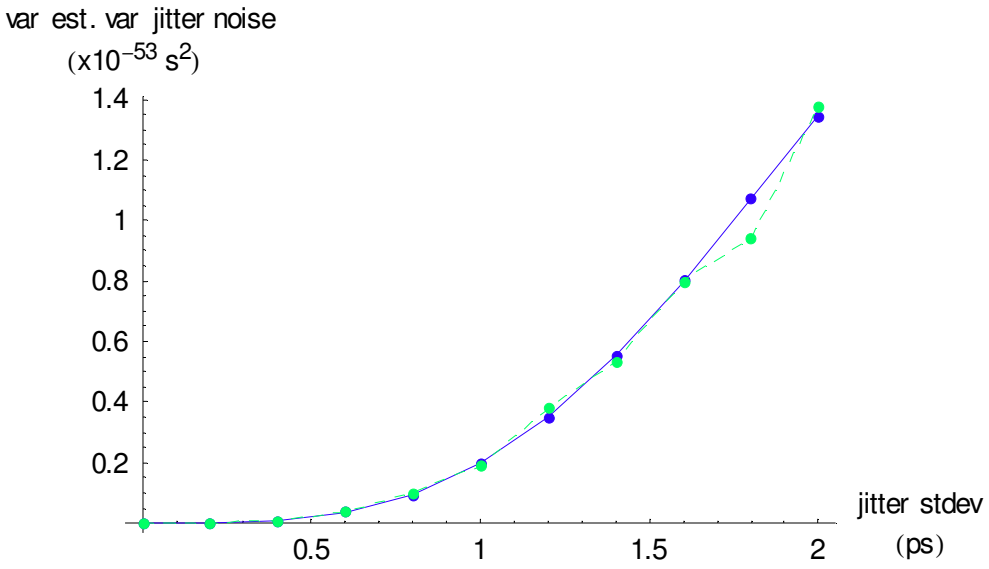

Figure 4-15. Variance of the estimated variance of the jitter noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 250 estimations), 3rd order model using the WLS estimator.

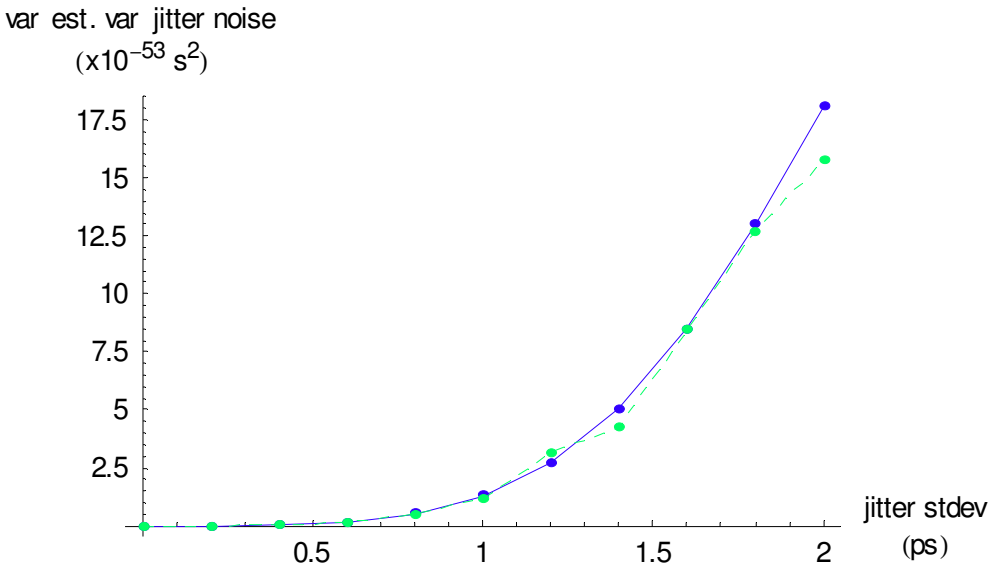

Figure 4-16. Variance of the estimated variance of the jitter noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 250 estimations), 3rd order model using the LS estimator.

estimator	order of model	mean abs. err. of est. σ_{n_t} (ps)	one sigma unc. on est. σ_{n_t} (fs)	mean abs. err. of est. σ_{n_y} (mV)	one sigma unc. on est. σ_{n_y} (μV)
LS	1	-0.027	1.645	0.008	0.486
	2	-0.032	1.589	0.001	0.504
	3	0.000	1.801	0.000	0.501
WLS	1	-0.007	0.765	0.000	0.047
	2	-0.031	0.636	0.000	0.047
	3	0.000	0.699	0.000	0.047
MLE	1	0.065	0.843	0.000	0.048
	2	-0.026	0.633	0.000	0.048
	3	0.000	0.703	0.000	0.048

Table 4-1. Summary of the simulation results in case of a 3rd order approximation of the variance and in case of known derivatives. Case of known standard deviation of additive noise σ_{n_y} of 0.6 mV and known standard deviation of jitter noise σ_{n_t} of 1 ps.

estimator	order of model	mean abs. err. of est. σ_{n_t} (ps)	one sigma unc. on est. σ_{n_t} (fs)	mean abs. err. of est. σ_{n_y} (mV)	one sigma unc. on est. σ_{n_y} (μV)
LS	1	-0.203	2.727	0.127	1.673
	2	-0.223	2.400	0.043	1.748
	3	0.000	3.364	0.000	1.509
WLS	1	-0.253	0.985	0.001	0.048
	2	-0.244	0.760	0.000	0.048
	3	0.000	0.917	0.000	0.048
MLE	1	0.388	1.414	0.000	0.049
	2	-0.128	0.739	-0.001	0.049
	3	0.000	0.922	0.000	0.048

Table 4-2. Summary of the simulation results in case of a 3rd order approximation of the variance and in case of known derivatives. Case of known standard deviation of additive noise σ_{n_y} of 0.6 mV and known standard deviation of jitter noise σ_{n_t} of 2 ps.

Step 2: realistic variance, known derivatives

Here, the simulated sample variance is obtained based on

$$y(t_i) = y_0(t_i + n_t(t_i)) + n_y(t_i). \qquad \text{Eq. 4-32}$$

The standard deviation of the additive noise $n_y(t_i)$ is fixed to 0.6 mV, while the standard deviation of the jitter noise $n_t(t_i)$ is stepped from 0 to 2 ps in 0.2 ps steps.

The required derivatives of the pulse are assumed to be known. In step 3, the derivatives will be calculated starting from the sample mean of the pulse and the estimated variance of the jitter noise. This allows to verify the impact of uncertainties on the required derivatives.

Simulation data for the sample mean and sample variance of the pulse is obtained using a normal distribution for both $n_y(t_i)$ and $n_t(t_i)$, $i = 1 .. 4096$ and based on 100 realizations of the pulse. The sample variance of the sample variance is based on 100 repeated realizations. Finally, the variance of the additive and jitter noise is estimated and this process is repeated 50 times.

Estimated jitter standard deviation

Figure 4-17 up to figure 4-19 show the sample mean of the absolute error of the estimated jitter standard deviation. It is clear that the third order model in combination with an unweighted LS estimator provides the best estimate for the jitter standard deviation. The mean relative error is 0.03% for a jitter standard deviation of 1 ps, while it is 0.3% using the WLS estimator.

Estimated additive noise standard deviation

Figure 4-20 on page 4-29 up to figure 4-22 show the estimated standard deviation of the additive noise. The exact value is 0.6 mV. Note that the mean relative error of a first order model in combination with an unweighted LS estimator is 1.25% for a jitter standard deviation of 1 ps. Using a third order model, this relative error is reduced to 0.11% and using a WLS estimator this error is further reduced to 0.02%.

Value of the cost function

Again, the value of the cost function in case of a weighted least-squares deserves special attention. Figure 4-23 shows the cost for the first, second and third order

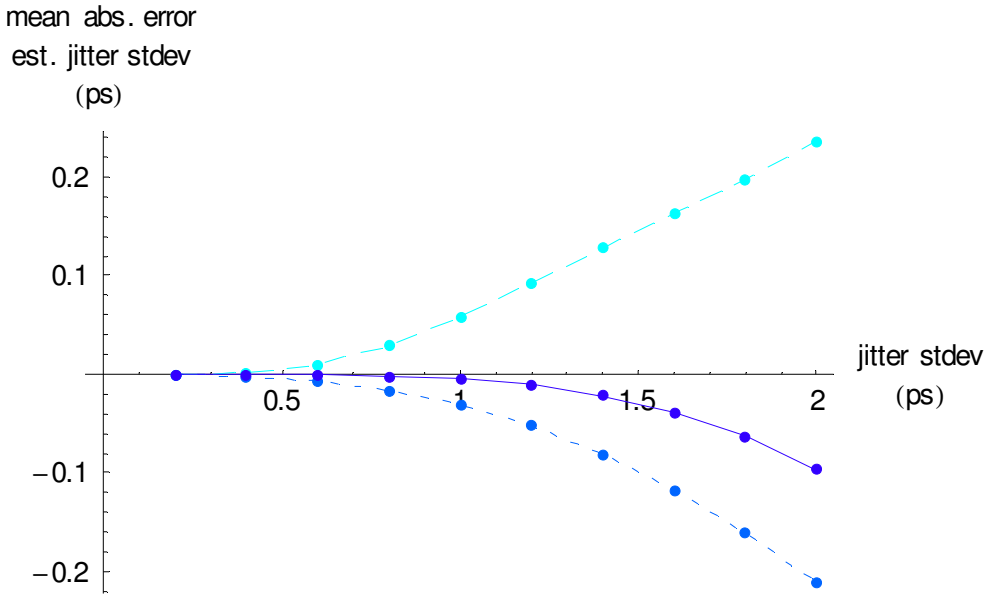

Figure 4-17. Mean absolute error of the estimated standard deviation of the jitter noise using the ML estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

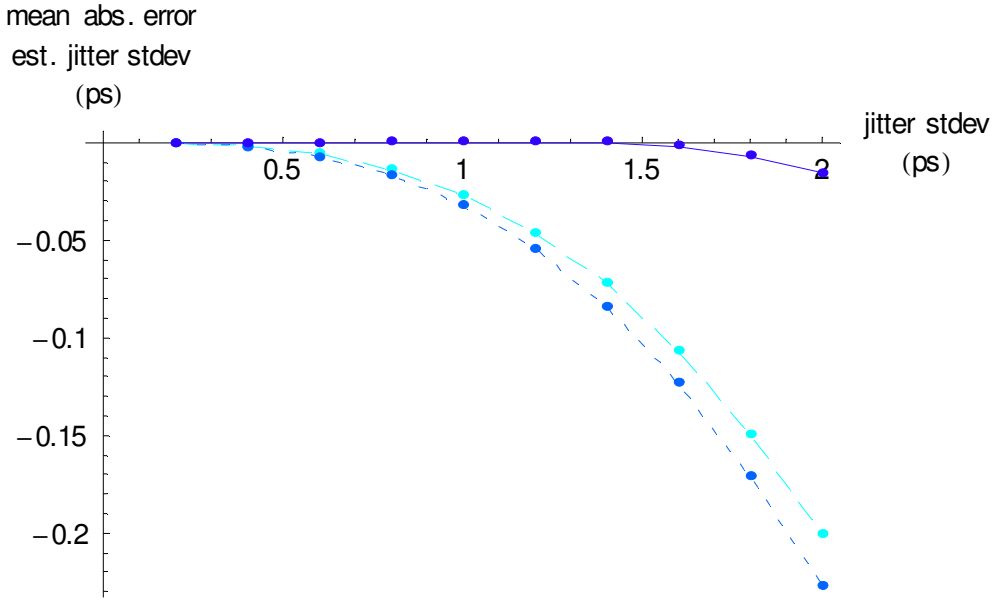

Figure 4-18. Mean absolute error of the estimated standard deviation of the jitter noise using the LS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

models and demonstrates the sensitivity of the cost with respect to model errors. Figure 4-24 and figure 4-25 on page 4-32 zoom into the sample mean and sample variance of the cost for the third order model.

In order to obtain the expected value and the variance of the cost, one can reuse the explanation found under "*Random number generator*" on page 4-15, taken into

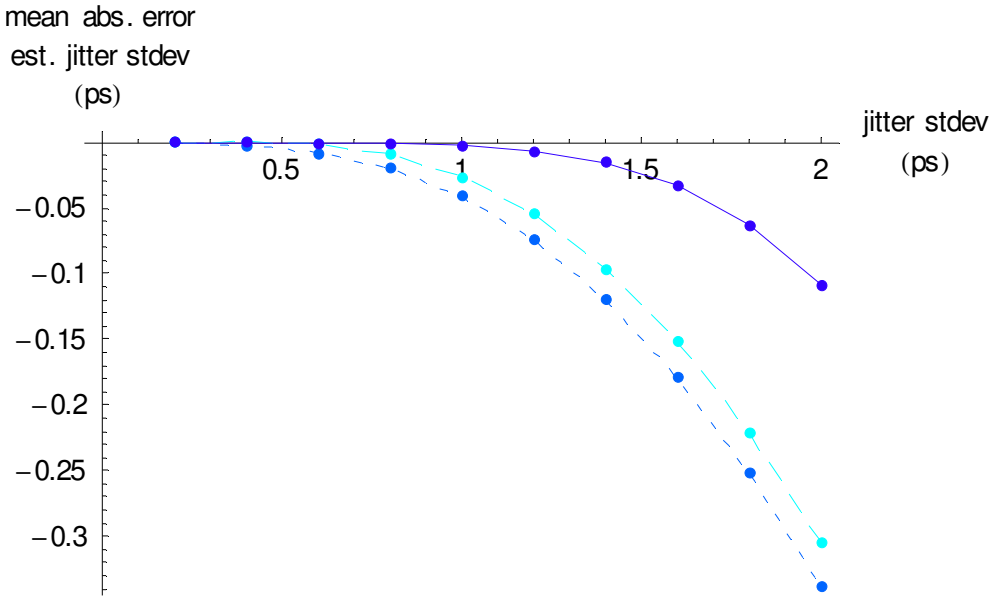

Figure 4-19. Mean absolute error of the estimated standard deviation of the jitter noise using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

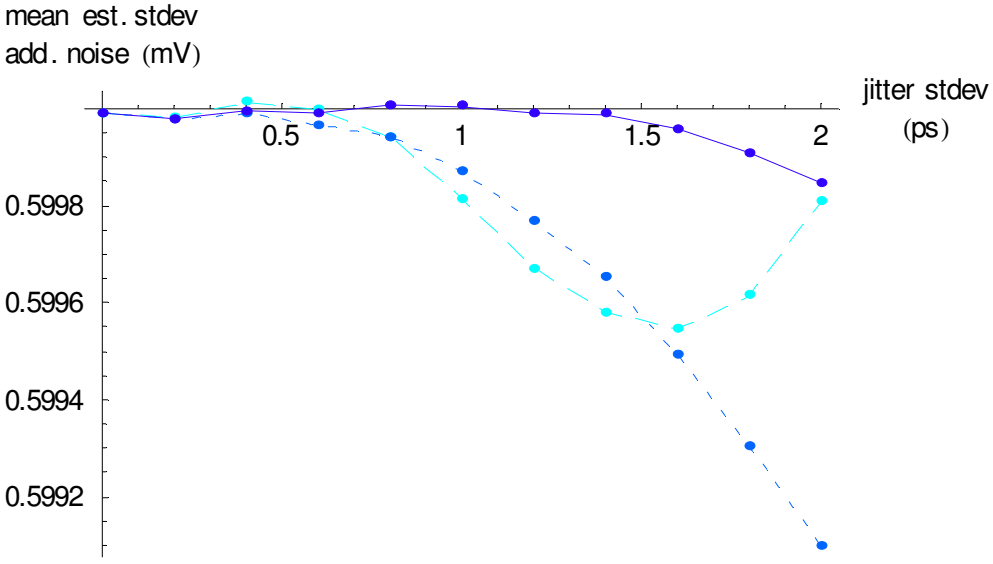

Figure 4-20. Mean estimated standard deviation of the additive noise using the ML estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

account that the number of degrees of freedom is decreased by the number of parameters ($p = 2$). During this simulation, the following values were used: $N_s = 100$, $N_t = 4096$, $N_w = 100$ and $N_e = 50$. Replacing N_t by $N_t - p$ in Eq. 4-26, the expected value of the cost μ_V is found to be 4178. The variance of the cost

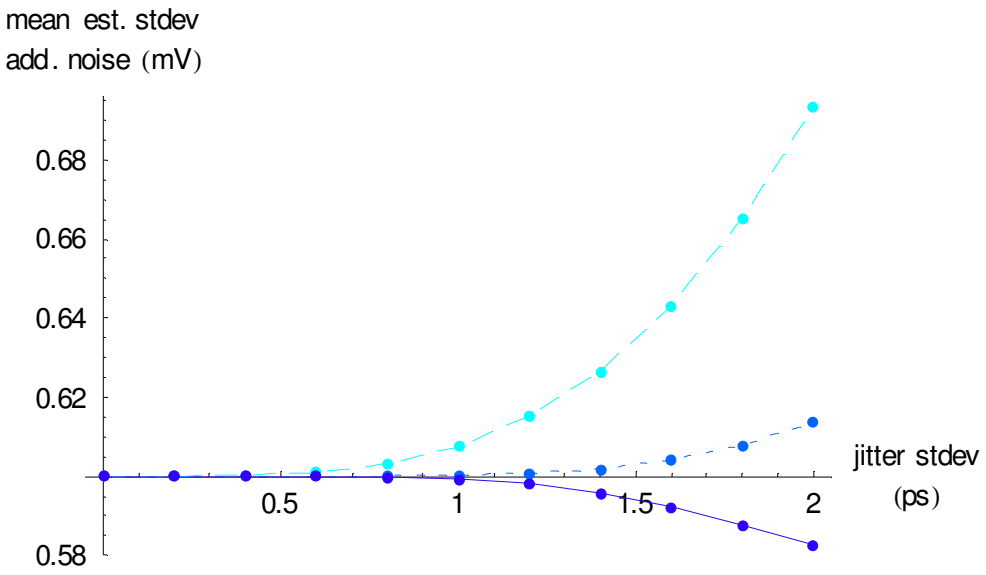

Figure 4-21. Mean estimated standard deviation of the additive noise using the LS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

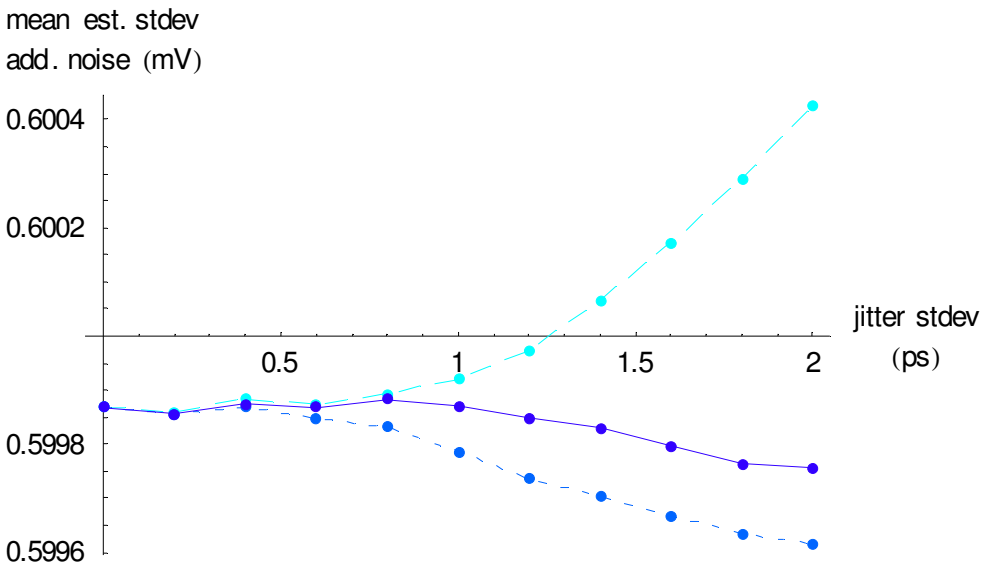

Figure 4-22. Mean estimated standard deviation of the additive noise using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

σ_V^2 equals 8888. Using Eq. 4-28 and Eq. 4-29, one can calculate the 95% confidence intervals for the expected value of the cost and its variance. These are shown in figure 4-24 and figure 4-25. For jitter standard deviations of more than 1 ps, the realized costs clearly deviate from the expected value of the cost. The asymmetry of the 95% confidence interval of the sample variance of the cost is due to the

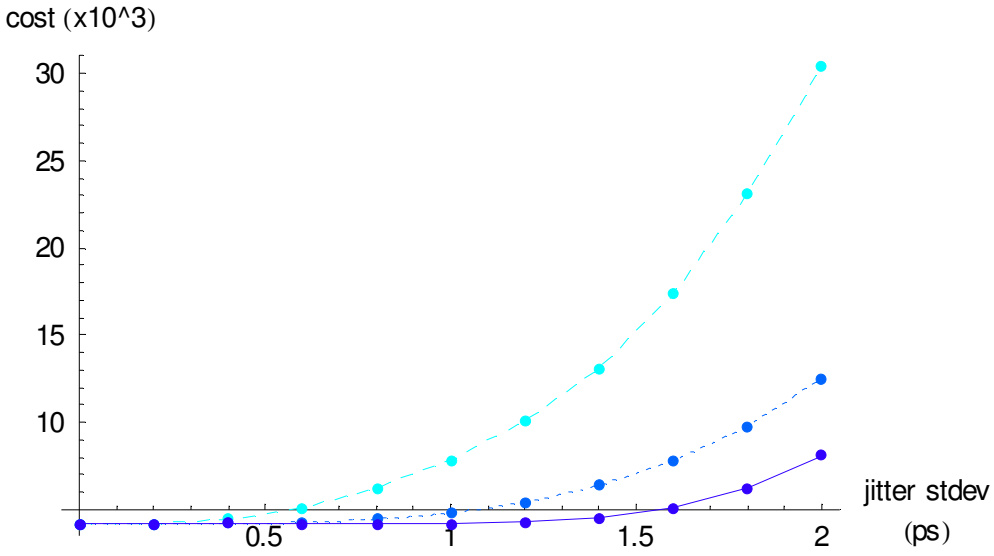

Figure 4-23. Mean value of the cost using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

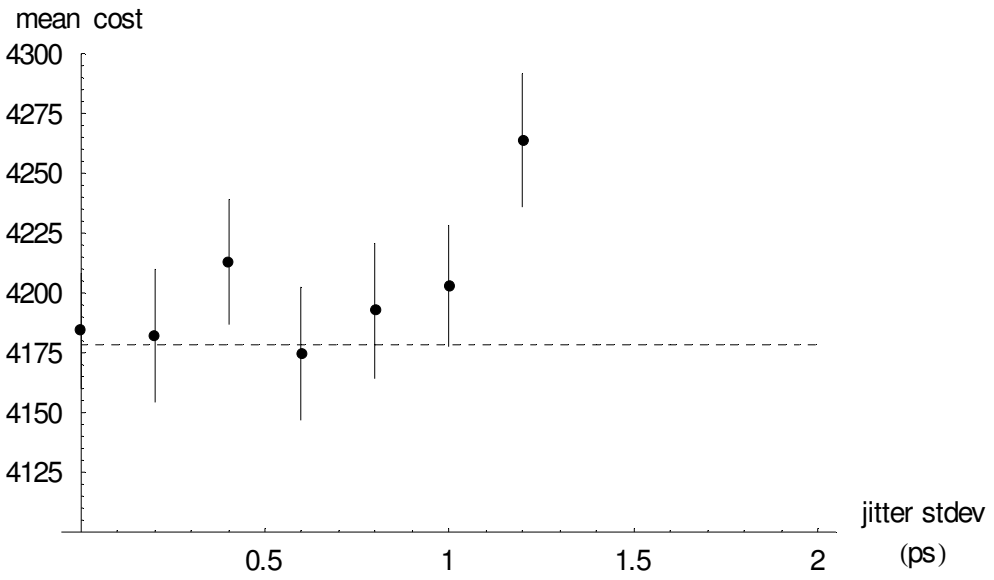

Figure 4-24. Expected value (dashed line) of the cost of the 3rd order model using the WLS estimator, compared to its sample mean value (50 realizations) and its 95% confidence interval. The mean cost for jitter standard deviations of more than 1.2 ps fall outside the selected vertical range.

asymmetric probability density function of a χ^2 - distribution for a relatively small number of degrees of freedom (49).

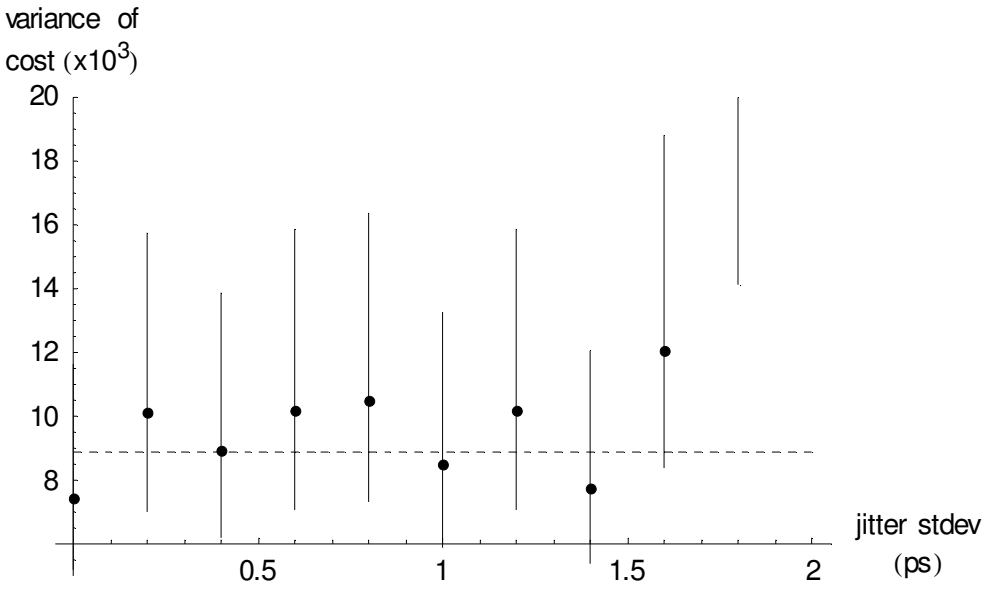

Figure 4-25. Expected variance (dashed line) of the cost of the 3rd order model using the WLS estimator, compared to its sample variance value (50 realizations) and its 95% confidence interval. The sample variance of the cost for a jitter standard deviation of 1.8 ps and 2 ps falls outside the selected vertical range.

It can be concluded that the unweighted LS estimator provides better estimates for the jitter standard deviation, while the WLS estimator outperforms the LS when estimating the standard deviation of the additive noise.

Using a third order model in combination with the WLS estimator, figure 4-24 clearly shows that the sample mean of the cost and its 95% confidence interval include the expected value of the cost for jitter standard deviation values up to 1 ps. For higher jitter values, the sample mean of the cost clearly starts to deviate from the expected values indicating the presence of model errors.

Using the covariance matrix of the parameters

Again it is verified that the uncertainty on the parameters, indicated by the parameter covariance matrix, corresponds to the sample variance of the parameters based on N_e repeated estimations.

Figure 4-26 shows the comparison for the estimated variance of the additive noise for a third order model in combination with the WLS estimator based on 50 estimations. Figure 4-27 repeats this comparison for the estimated variance of the jitter noise for the same model and estimator. Good correspondence is found, especially when taking the 95% confidence interval of the sample variance into account, based on a χ^2 - distribution with 49 degrees of freedom:

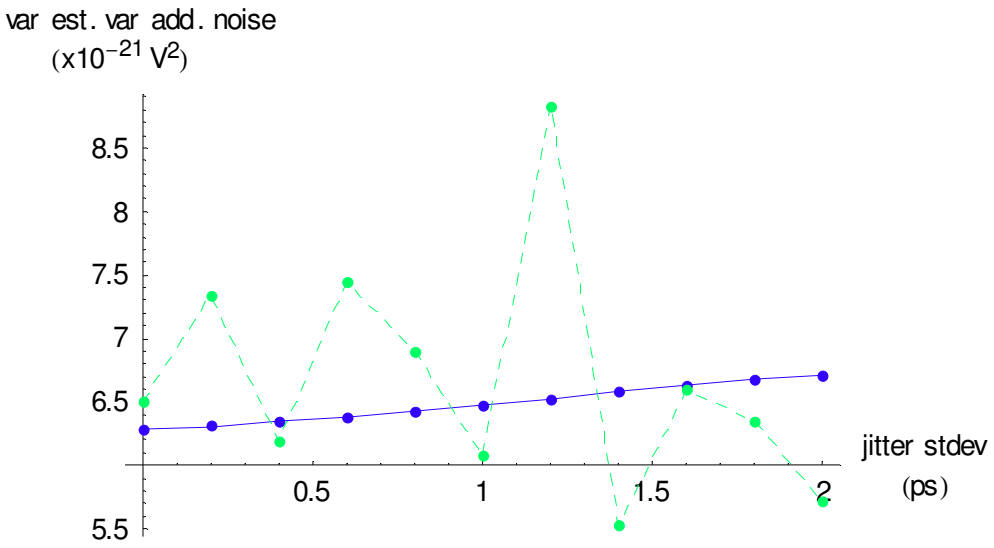

Figure 4-26. Variance of the estimated variance of the additive noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 50 estimations), 3rd order model using the WLS estimator.

$$0.698\, S_{N_e}^2 \leq \sigma_V^2 \leq 1.553\, S_{N_e}^2 \qquad \text{Eq. 4-33}$$

Finally, figure 4-28 shows that the uncertainty on the estimated jitter variance is indeed smaller when using a WLS estimator instead of a (unweighted) LS estimator. However, the uncertainty on the estimated jitter standard deviation is sufficiently small to prefer the LS estimator over the WLS. Indeed, starting from the sample variance of the sample variance, one can calculate the standard deviation on the standard deviation:

$$\sigma_{\sqrt{x}} \cong \sqrt{\frac{\sigma_x^2}{4x}} \qquad \text{Eq. 4-34}$$

As such, the variance of the estimated variance of 2.5 10^{-53} boils down to a 95% confidence interval on the estimated jitter standard deviation of ± 5 fs in case of a known jitter standard deviation of 1 ps.

It should be noticed that the number of pulse realizations used for one estimate of the parameters, is 10^4. Decreasing this number will increase the uncertainty on the parameters. Given a specified tolerance on the estimated parameters, using a smaller number of realizations, the larger uncertainty of the LS-based estimates may become an argument to prefer the WLS estimator over the LS estimator. Finally, one should recall that the WLS estimator has the additional advantage of providing a cost which

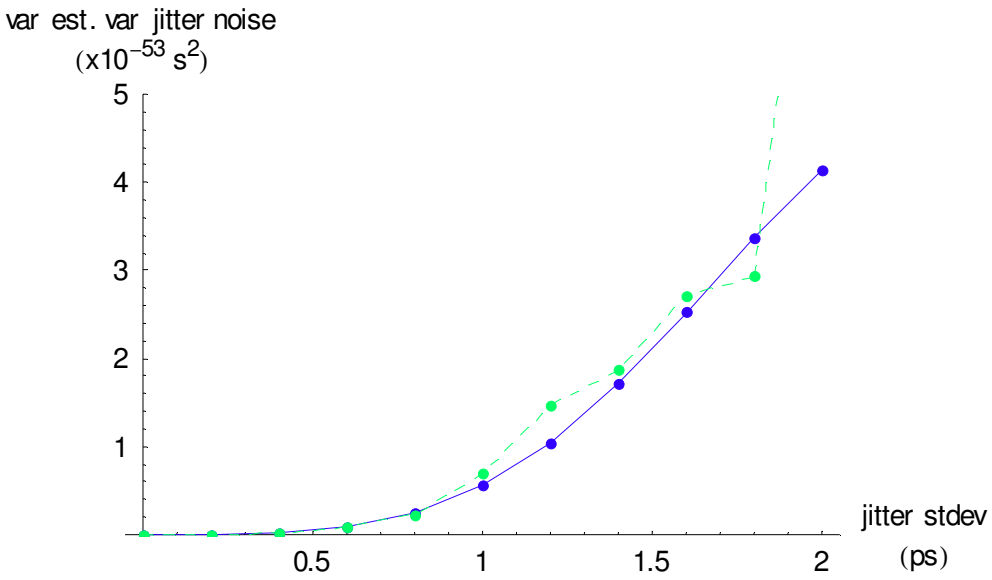

Figure 4-27. Variance of the estimated variance of the jitter noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 50 estimations), 3rd order model using the WLS estimator.

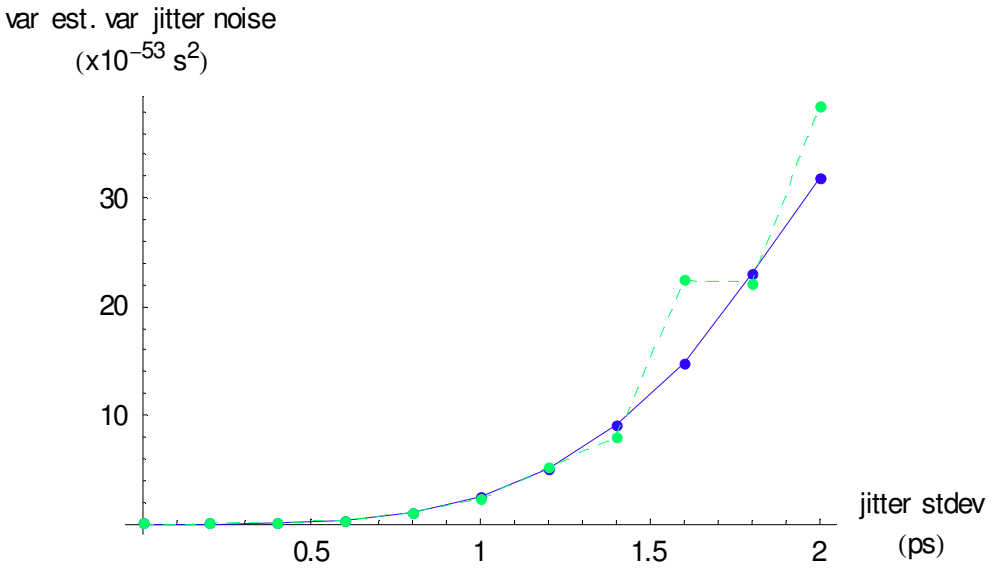

Figure 4-28. Variance of the estimated variance of the jitter noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 50 estimations), 3rd order model using the LS estimator.

can be compared to the expected value of the cost and allows to detect model errors or other anomalies.

An overview of the simulation results can be found in Table 4-3 and Table 4-4.

estimator	order of model	mean abs. err. of est. σ_{n_t} (ps)	one sigma unc. on est. σ_{n_t} (fs)	mean abs. err. of est. σ_{n_y} (mV)	one sigma unc. on est. σ_{n_y} (μV)
LS	1	-0.027	2.249	0.008	0.683
	2	-0.032	2.175	0.000	0.701
	3	0.000	2.468	-0.001	0.696
WLS	1	-0.026	1.128	0.000	0.067
	2	-0.041	1.028	0.000	0.067
	3	-0.003	1.183	0.000	0.067
MLE	1	0.058	1.180	0.000	0.067
	2	-0.031	0.890	0.000	0.067
	3	-0.005	0.987	0.000	0.067

Table 4-3. Summary of the simulation results in case of a realistic variance and in case of known derivatives. Case of known standard deviation of additive noise σ_{n_y} of 0.6 mV and known standard deviation of jitter noise σ_{n_t} of 1 ps.

estimator	order of model	mean abs. err. of est. σ_{n_t} (ps)	one sigma unc. on est. σ_{n_t} (fs)	mean abs. err. of est. σ_{n_y} (mV)	one sigma unc. on est. σ_{n_y} (μV)
LS	1	-0.200	3.417	0.093	2.361
	2	-0.227	3.064	0.014	2.286
	3	-0.016	4.457	-0.017	1.957
WLS	1	-0.305	1.305	0.000	0.068
	2	-0.337	1.095	0.000	0.068
	3	-0.108	1.607	0.000	0.068
MLE	1	0.236	1.811	0.000	0.068
	2	-0.211	0.987	-0.001	0.068
	3	-0.097	1.223	0.000	0.068

Table 4-4. Summary of the simulation results in case of a realistic variance and in case of known derivatives. Case of known standard deviation of additive noise σ_{n_y} of 0.6 mV and known standard deviation of jitter noise σ_{n_t} of 2 ps.

Step 3: realistic variance, derivatives based on sample mean

In this step, it is no longer assumed that the derivatives are known. Instead, the derivatives are calculated starting from the sample mean of the pulse. Due to the low-pass effect of the jitter, the sample mean of the pulse is a filtered version of the true pulse and must be corrected first, based on the estimated variance of the jitter noise. Therefore, during iteration i of the estimation process, the sample mean is corrected using the jitter variance estimated during iteration $i-1$. The derivatives of this corrected pulse are calculated via the frequency domain.

Using the same random seed, the simulation data is identical to the one used in step 2 for easy comparison. The results below don't show any negative consequences of the fact that the derivatives must be calculated.

Estimated jitter standard deviation

Figure 4-29 up to figure 4-31 show the sample mean of the absolute error of the estimated jitter standard deviation. Here too, it is clear that the third order model in combination with an unweighted LS estimator provides a better estimate for the jitter standard deviation. The mean relative error is 0.04% for a jitter standard deviation of 1 ps, while it is 0.3% using the WLS estimator. As such, for the estimation of the jitter

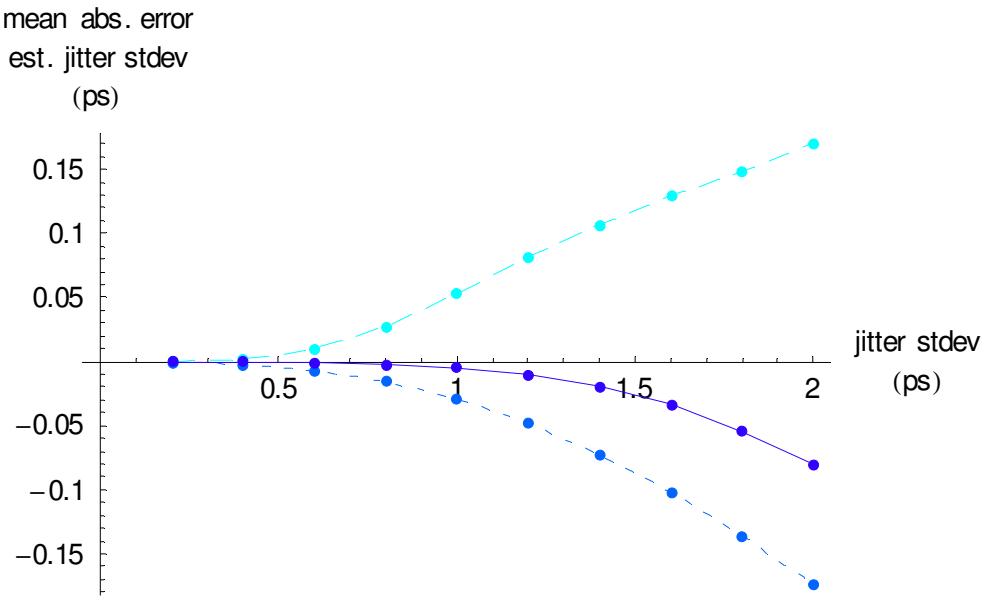

Figure 4-29. Mean absolute error of the estimated standard deviation of the jitter noise using the ML estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

standard deviation, the unweighted LS estimator does a better job than the WLS.

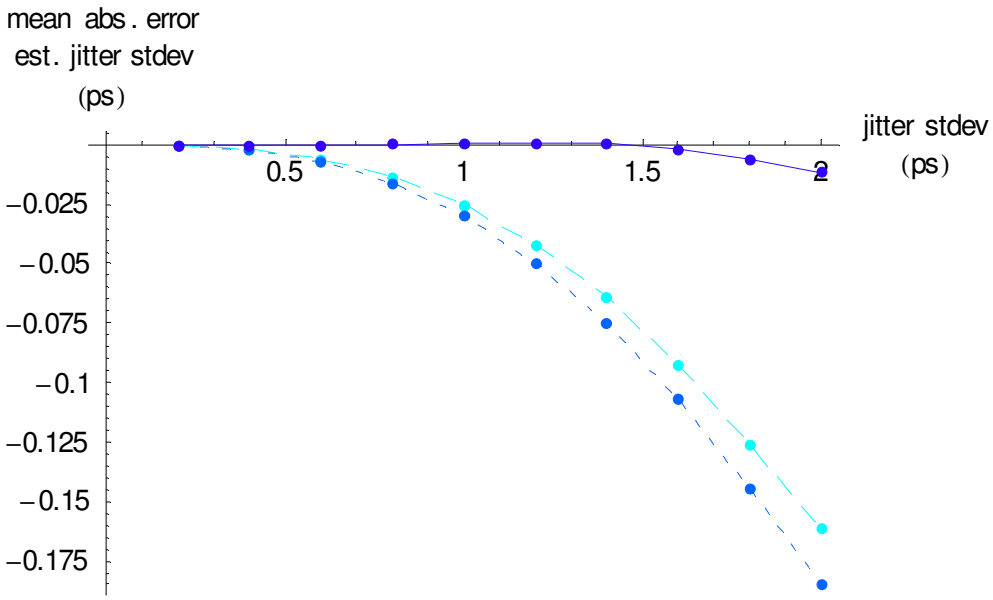

Figure 4-30. Mean absolute error of the estimated standard deviation of the jitter noise using the LS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

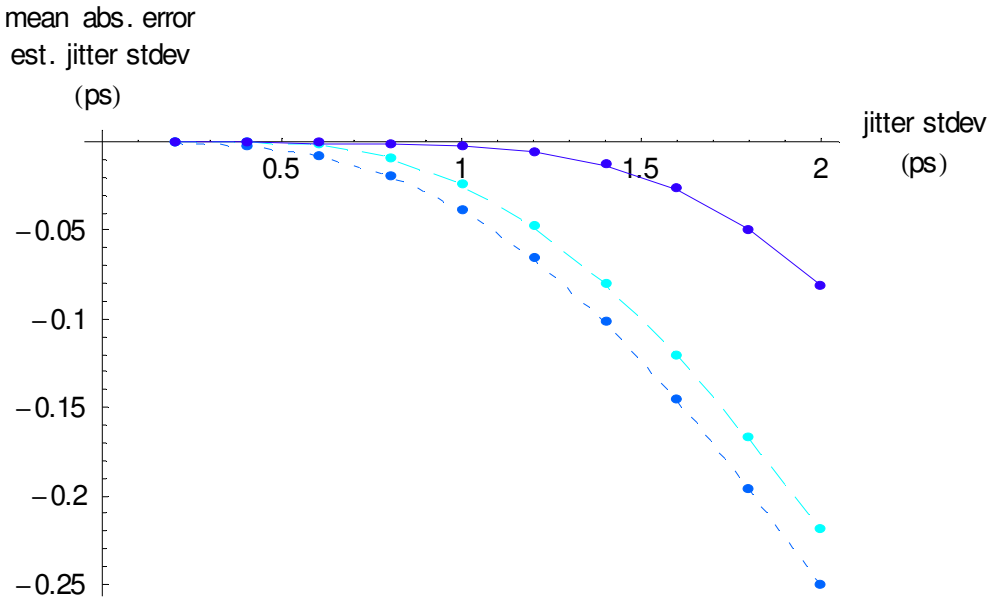

Figure 4-31. Mean absolute error of the estimated standard deviation of the jitter noise using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

Estimated additive noise standard deviation

Figure 4-32 up to figure 4-34 show the estimated standard deviation of the additive noise. The exact value is 0.6 mV. Note that the mean relative error of a first order

model in combination with an unweighted LS estimator is 1.24% for a jitter standard deviation of 1 ps. Using a third order model, this relative error is reduced to 0.12% and using a WLS estimator this error is further reduced to 0.02%. As such, for the

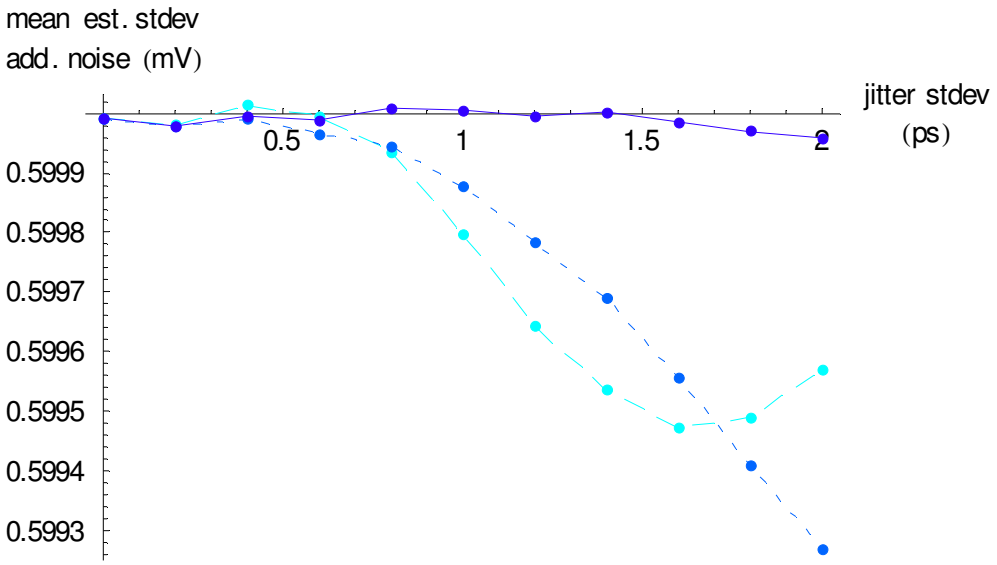

Figure 4-32. Mean estimated standard deviation of the additive noise using the ML estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

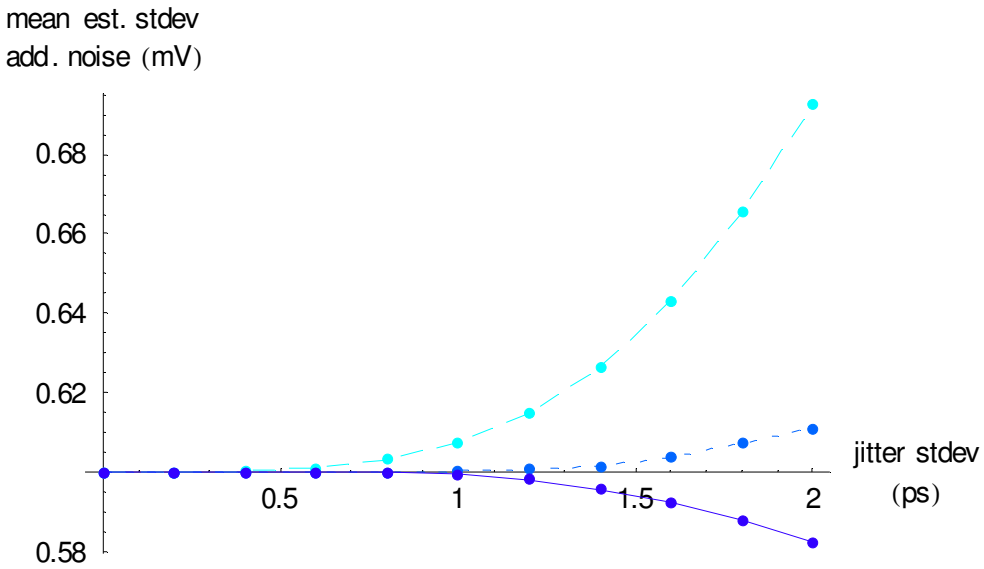

Figure 4-33. Mean estimated standard deviation of the additive noise using the LS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

estimation of the standard deviation of the additive noise, one should prefer the WLS estimator over the unweighted LS.

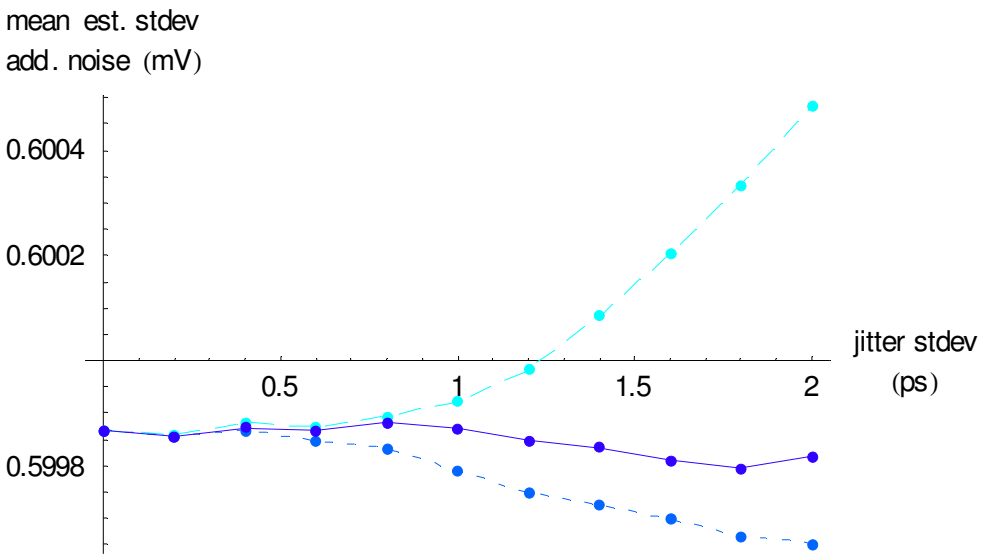

Figure 4-34. Mean estimated standard deviation of the additive noise using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

Value of the cost function

Figure 4-35 shows the mean cost for the first, second and third order model in combination with a WLS estimator. Figure 4-36 and figure 4-37 zoom into the sample mean and sample variance of the cost for the third order model.

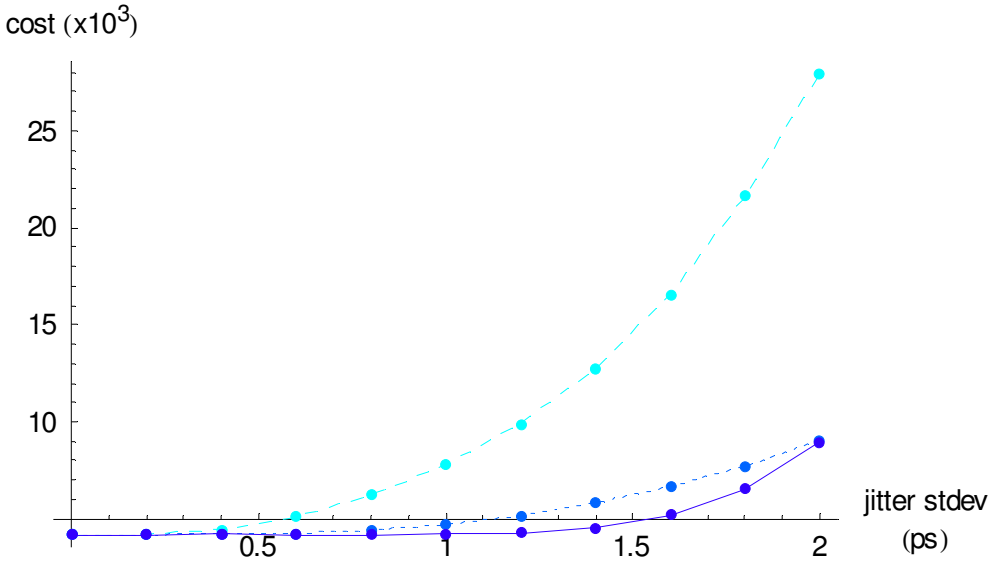

Figure 4-35. Mean value of the cost using the WLS estimator. (1st order: long dashed line, 2nd order: short dashed line, 3rd order: solid line)

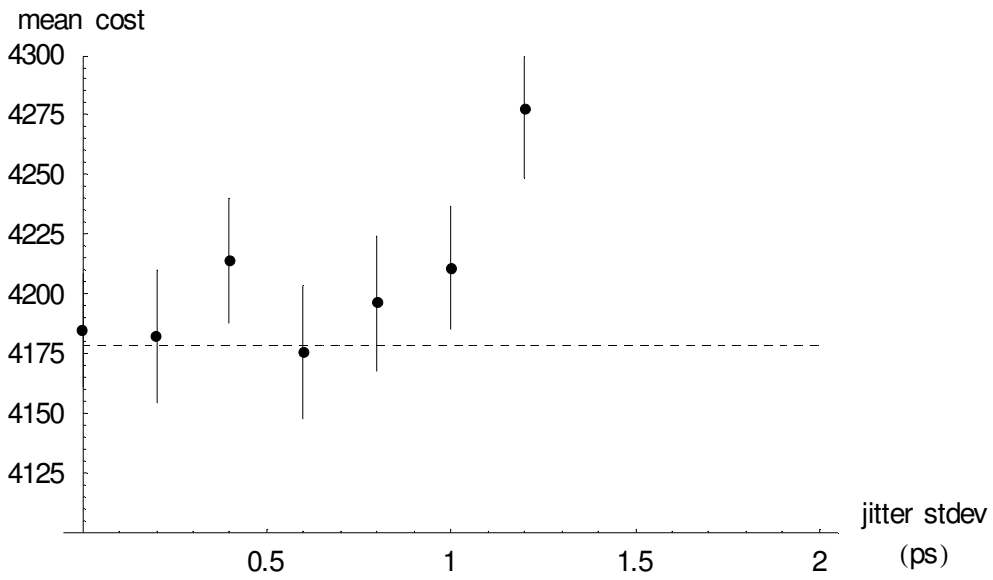

Figure 4-36. Expected value (dashed line) of the cost of the 3rd order model using the WLS estimator, compared to its sample mean value (50 realizations) and its 95% confidence interval. The mean cost for jitter standard deviations of more than 1.2 ps fall outside the selected vertical range.

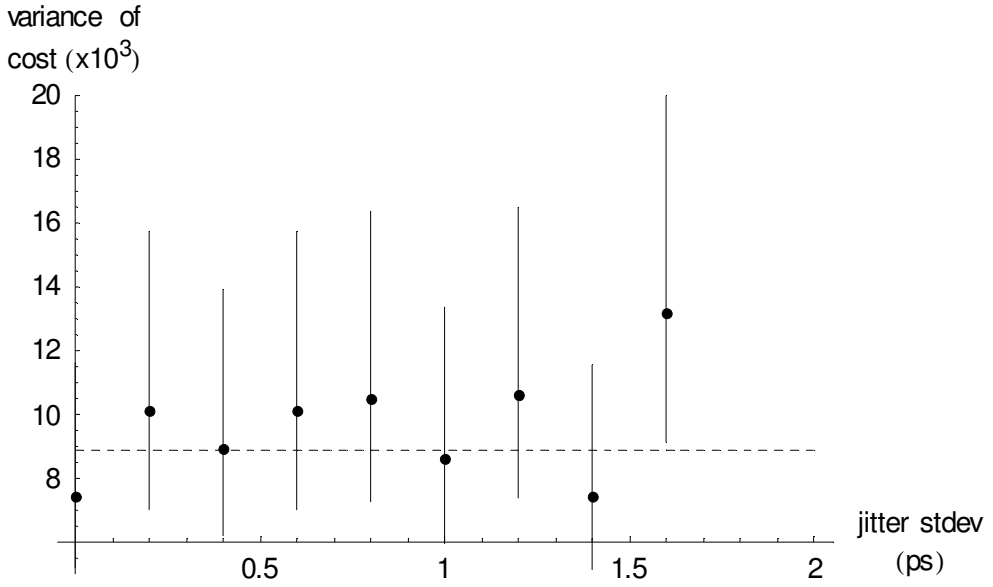

Figure 4-37. Expected variance (dashed line) of the cost of the 3rd order model using the WLS estimator, compared to its sample variance value (50 realizations) and its 95% confidence interval. The sample variance of the cost for a jitter standard deviation of 1.8 ps and 2 ps falls outside the selected vertical range.

The expected value of the cost, μ_V, its variance σ_V^2 and their 95% confidence interval equal those of step 2. These are shown in figure 4-36 and figure 4-37.

Comparing figure 4-24 to figure 4-36 and figure 4-25 to figure 4-37, it is clear that the effect of not knowing the exact derivatives is negligible. As was the case in step 2, the cost starts deviating from the expected cost for jitter values larger than 1 ps, implying the presence of model errors.

Using the covariance matrix of the parameters

Again, the uncertainty on the parameters obtained from theory and simulation are compared. It is verified if the uncertainty on the parameters, indicated by the parameter covariance matrix, corresponds to the sample variance of the parameters based on N_e repeated estimations.

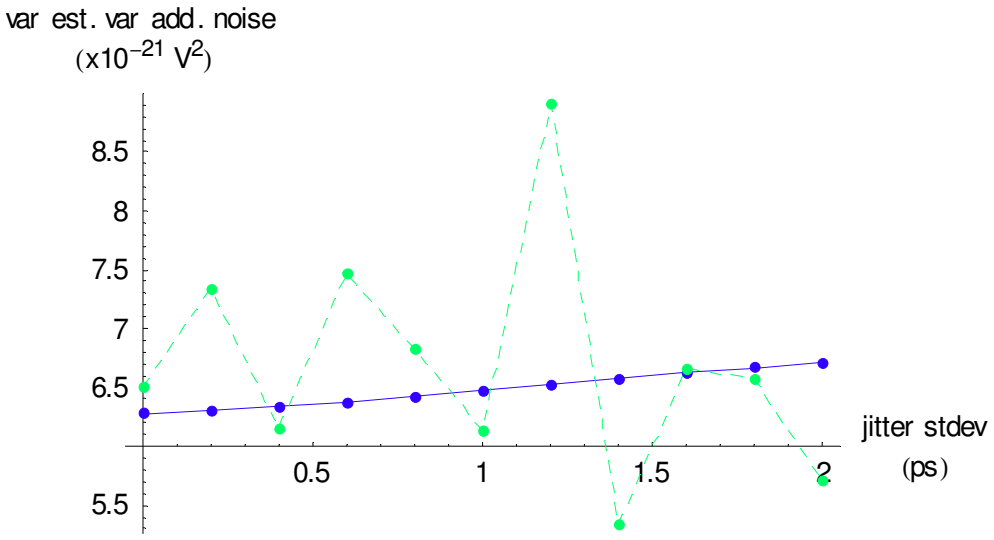

Figure 4-38. Variance of the estimated variance of the additive noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 50 estimations), 3rd order model using the WLS estimator.

Figure 4-38 shows the comparison for the estimated variance of the additive noise for a third order model in combination with the WLS estimator based on 50 estimations. Again, good correspondence is found after taking the 95% confidence interval of the sample variance into account, based on a χ^2 - distribution with 49 degrees of freedom (Eq. 4-33). Figure 4-39 repeats this comparison for the estimated variance of the jitter noise for the same model and estimator. Finally, figure 4-40 shows that once more the uncertainty on the estimated jitter variance is smaller using a WLS estimator instead of a (unweighted) LS estimator. Again, the uncertainty on the estimated jitter standard deviation is similar to that of step 2 and as such sufficiently small to prefer the LS estimator over the WLS for the estimation of the jitter standard deviation, while the WLS is preferred over the LS estimator when estimating the standard deviation of the additive noise.

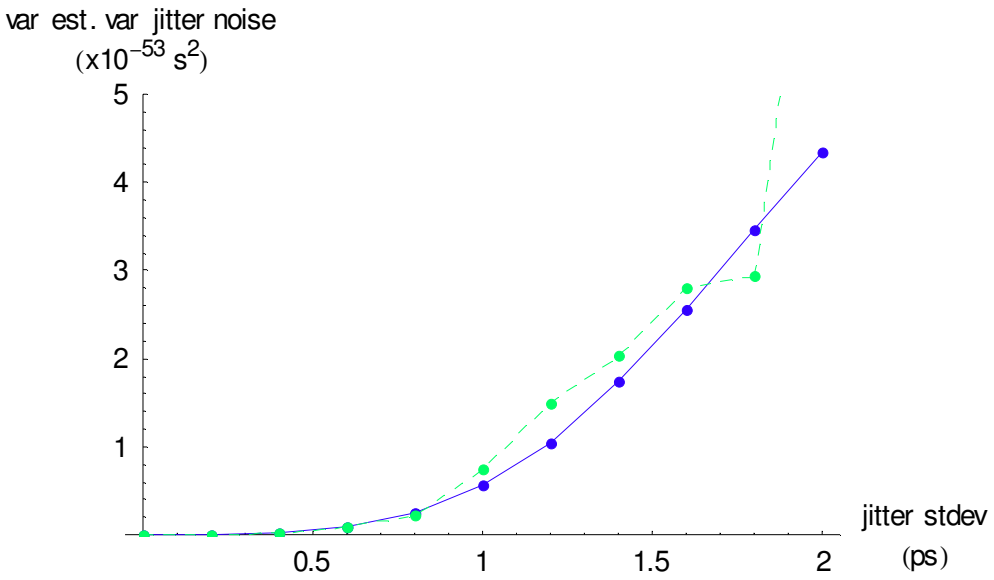

Figure 4-39. Variance of the estimated variance of the jitter noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 50 estimations), 3rd order model using the WLS estimator.

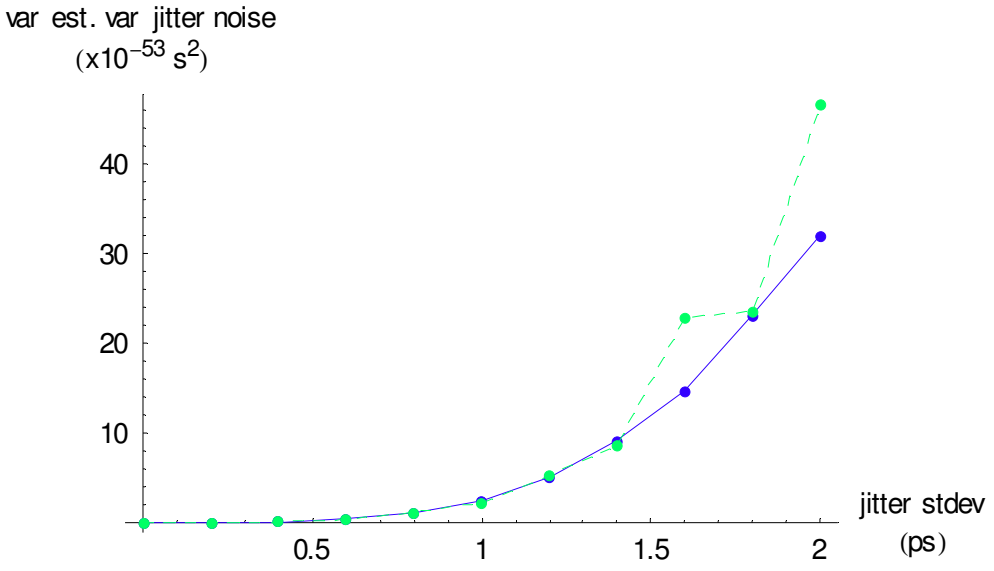

Figure 4-40. Variance of the estimated variance of the jitter noise (solid line: based on the parameter covariance matrix, dashed line: sample variance based on 50 estimations), 3rd order model using the LS estimator.

An overview of the simulation results can be found in Table 4-5 and Table 4-6.

estimator	order of model	mean abs. err. of est. σ_{n_t} (ps)	one sigma unc. on est. σ_{n_t} (fs)	mean abs. err. of est. σ_{n_y} (mV)	one sigma unc. on est. σ_{n_y} (μV)
LS	1	-0.025	2.257	0.007	0.682
	2	-0.030	2.184	0.000	0.701
	3	0.000	2.468	-0.001	0.696
WLS	1	-0.024	1.133	0.000	0.067
	2	-0.038	1.034	0.000	0.067
	3	-0.003	1.184	0.000	0.067
MLE	1	0.053	1.173	0.000	0.068
	2	-0.029	0.898	0.000	0.068
	3	-0.005	0.993	0.000	0.068

Table 4-5. Summary of the simulation results in case of a realistic variance and in case of derivatives based on the sample mean. Case of known standard deviation of additive noise σ_{n_y} of 0.6 mV and known standard deviation of jitter noise σ_{n_t} of 1 ps.

estimator	order of model	mean abs. err. of est. σ_{n_t} (ps)	one sigma unc. on est. σ_{n_t} (fs)	mean abs. err. of est. σ_{n_y} (mV)	one sigma unc. on est. σ_{n_y} (μV)
LS	1	-0.161	3.564	0.093	2.351
	2	-0.185	3.199	0.011	2.270
	3	-0.012	4.474	-0.018	1.954
WLS	1	-0.219	1.435	0.000	0.068
	2	-0.250	1.195	0.000	0.068
	3	-0.082	1.646	0.000	0.068
MLE	1	0.170	1.700	0.000	0.069
	2	-0.173	1.047	-0.001	0.069
	3	-0.080	1.257	0.000	0.069

Table 4-6. Summary of the simulation results in case of a realistic variance and in case of derivatives based on the sample mean. Case of known standard deviation of additive noise σ_{n_y} of 0.6 mV and known standard deviation of jitter noise σ_{n_t} of 2 ps.

Step 4: influence of time base drift

The above simulations show that all estimators perform reasonably well for the used time signal (Eq. 4-22) when the jitter standard deviation is limited to 1 ps. Applying them to measured data, large discrepancies were found between the measured and estimated variance, even though the estimated jitter standard deviation also turns out to be approximately 1 ps. The sampling oscilloscope measurements add extra challenges due to time base drift and time base distortion that are present in the measurements and absent in the simulations. It is decided to study the effect of time base drift and its compensation on the estimated parameters.

Let $x(t)$ be a band-limited signal. $x_n = x(n \cdot \Delta t)$ represents the sampled version of $x(t)$.

It is possible to reconstruct $x(t)$:

$$x(t) = \sum_{k=-\infty}^{+\infty} x_k \cdot \mathrm{sinc}\left\{\pi\left(\frac{t}{\Delta t} - k\right)\right\}, \text{ where } \mathrm{sinc}(\theta) = \frac{\sin(\theta)}{\theta}. \qquad \text{Eq. 4-35}$$

Applying a delay τ to $x(t)$:

$$x'(t) = x(t-\tau) = \sum_{k=-\infty}^{+\infty} x_k \cdot \mathrm{sinc}\left\{\pi\left(\frac{t-\tau}{\Delta t} - k\right)\right\}. \qquad \text{Eq. 4-36}$$

The sampled version of this delayed signal then becomes

$$x'_n = x'(n \cdot \Delta t) = \sum_{k=-\infty}^{+\infty} x_k \cdot \mathrm{sinc}\left\{\pi\left(n - k - \frac{\tau}{\Delta t}\right)\right\} \qquad \text{Eq. 4-37}$$

Eq. 4-37 clearly shows that for an arbitrary delay τ, x'_n depends on all x_k. However, if the applied delay equals a multiple of the sampling period, i.e. $\tau = m \cdot \Delta t$, all contributions to Eq. 4-37 are zero, except for $k = n - m$. As such, $x'_n = x_{n-m}$.

Based on the above, it is clear that, starting from a time sequence x_n, $n = 1 .. N$, of independent stochastical variables with known variance σ_n^2, the variance of the sampled version of the arbitrarily delayed signal $x'_n = \sum_{k=1}^{N} a_k \cdot x_k$ equals $\sigma'^2_n = \sum_{k=1}^{N} a_k^2 \cdot \sigma_k^2$. Moreover, x'_n is no longer independent with respect to n.

Applying the same delay to the original variance σ_n^2, $\tilde{\sigma}_n^2 = \sum_{k=1}^{N} a_k \cdot \sigma_k^2$.

Obviously $\sigma'^2_n \neq \tilde{\sigma}_n^2$ for a general sequence a_k.

If the applied delay is an integer multiple of the sampling period, i.e. $\tau = m \cdot \Delta t$, then $x'_n = x_{n-m}$. In that case $\sigma'^2_n = \tilde{\sigma}_n^2$.

The analytical pulse (Eq. 4-22) is used to study the shaping of the variance as function of $\frac{\tau}{\Delta t}$. First a known delay $\tau = k \cdot \frac{\Delta t}{10}$, $k = 1 .. 5$, $\Delta t = \frac{5\ \text{ns}}{4096}$ is applied to the analytical expression. The sample variance of this delayed pulse is obtained based on 1000 realizations using a standard deviation of 0.6 mV of the additive noise and a standard deviation of 1 ps of the jitter noise. Next the inverse delay is applied to this sample variance and compared to the sample variance of the original pulse $(\Delta t = 0)$.

Figure 4-41 clearly indicates that the shaping of the sample variance is very limited when the delay is limited to $0.1\Delta t$.

However, when a delay is applied of $0.5\Delta t$, the effect is significant.

Especially the increased level of the minima of the sample variance, corresponding to a zero slope of the average pulse, incorrectly amplifies the observation that the first order approximation of the model equation (Eq. 4-3) no longer holds.

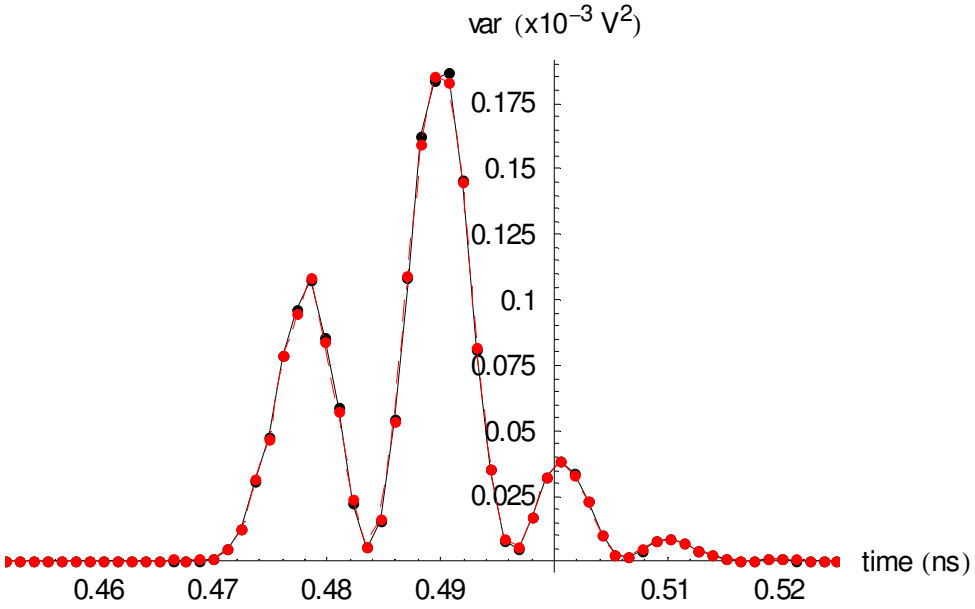

Figure 4-41. Minimal shaping of the sample variance (solid line: original variance, dashed line: variance after delay compensation of 0.1Δt)

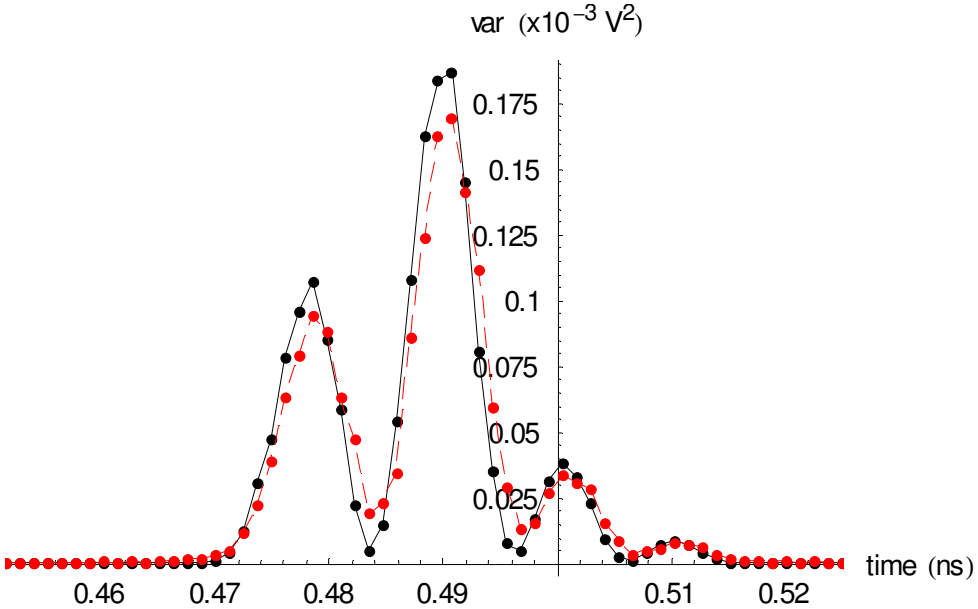

Figure 4-42. Clear shaping of the sample variance (solid line: original variance, dashed line: variance after delay compensation of 0.5Δt)

Starting from the information shown in figure 4-41 and figure 4-42, figure 4-43 shows the difference between the original variance and the variance after delay compensation of $0.1\Delta t$ and $0.5\Delta t$.

The above implies that one cannot apply an arbitrary delay to a signal before averaging without introducing errors in the sample variance, which is assumed to be based on stochastically independent variables, with respect to t.

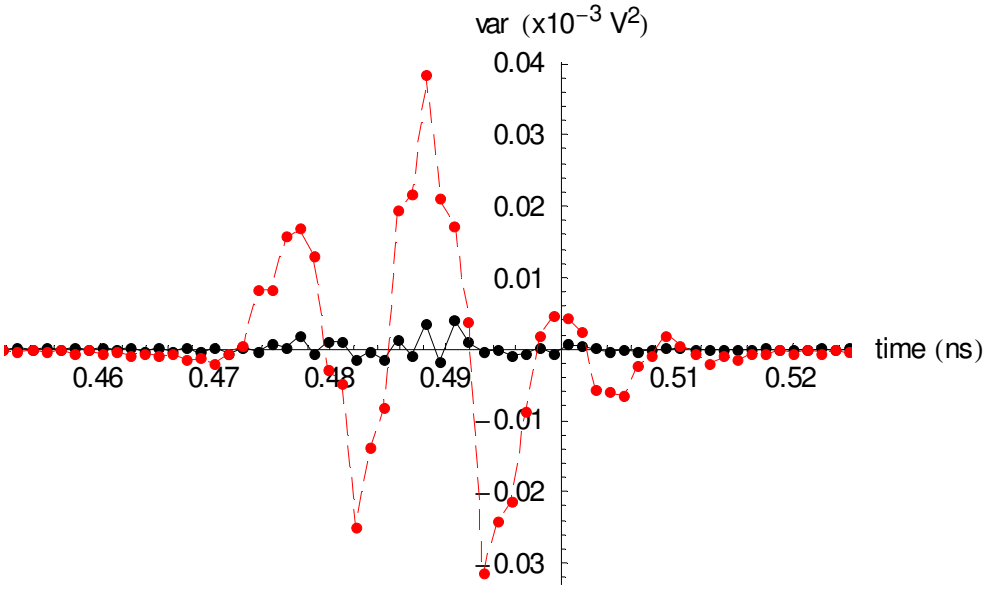

Figure 4-43. Difference between the original variance and the one after delay compensation (solid line: delay of 0.1Δt, dashed line: delay of 0.5Δt)

According to figure 4-4 on page 4-14, which shows the sample variance of impulse response measurements after time base drift compensation, the measurement indeed suffers from the above error. Using this sample variance, the estimated time base jitter is incorrect.

A similar problem becomes apparent when compensating for the time base distortion. Any interpolation combines two or more values of the non-equidistant time grid to obtain an estimate of the value at the equidistant time grid points. This again shapes the variance and introduces correlation between successive time points.

The error introduced by time base drift compensation before averaging can be limited based on the observations that

- applying a delay which is an integer multiple of the sampling period Δt introduces no error,
- for the impulse response of interest, the shaping of the variance and as such the error is very small[1] when the delay is limited to $0.1\Delta t$.

Therefore, the different realizations of the pulse are first delayed by integer multiples of Δt such that all realizations are aligned within $\pm 0.5\Delta t$.

Then all realizations are divided into time buckets which are $\pm 0.1\Delta t$ wide, resulting in a 5-times oversampled signal as compared to the original signal.

1. The RMS value of the difference is 1.6% of the RMS value of the original variance, while the maximum value of the difference is 2.1% of the maximum value of the original variance.

Depending on the shape of the time base drift, in general the buckets will not be equally filled. As such, the uncertainty on the sample variance will vary as function of the bucket index. This is not an issue when using a weighted least squares to extract the model parameters.

Time base jitter interpretable as time base drift

In case of measurements, the time base drift of each acquisition with respect to the first one is estimated by minimizing

$$V = \sum_{i=1}^{N} \left| X_{ref}(\omega_i) - e^{-j\omega_i \tau} \cdot X(\omega_i) \right|^2 \qquad \text{Eq. 4-38}$$

within the bandwidth of the signal $X(\omega_i)$.

It is impossible to fully separate time base drift and time base jitter, because the latter can be interpreted as time base drift as shown by the following simulation. The analytical pulse is distorted by additive noise (0.6 mV standard deviation) and jitter noise (1 ps standard deviation). 500 realizations of 4096 points each are generated.

First the drift - which is known to be zero - is estimated using the first realization as reference signal. Next, the drift of each realization with respect to any other one is estimated in order to obtain an enhanced drift estimate with respect to the first realization, as advised in [8]. Figure 4-44 shows no noticeable difference between both approaches.

The sample standard deviation of the estimated drift, in the case of a jitter standard deviation of 1 ps and based on the 500 realizations, equals 0.26 ps (0.25 ps for the enhanced estimate).

Given the above and based on reasonable time constants[1] corresponding to thermal effects, it makes sense to apply smoothing to the estimated time base drift. This will average out the effect of the fast jitter, while it will leave the slow drift mainly unaffected. Additional motivation for this smoothing can be found in the chapter on drift estimation in the presence of both jitter and additive noise.

1. Assumptions with respect to time constants are explained in step 5.

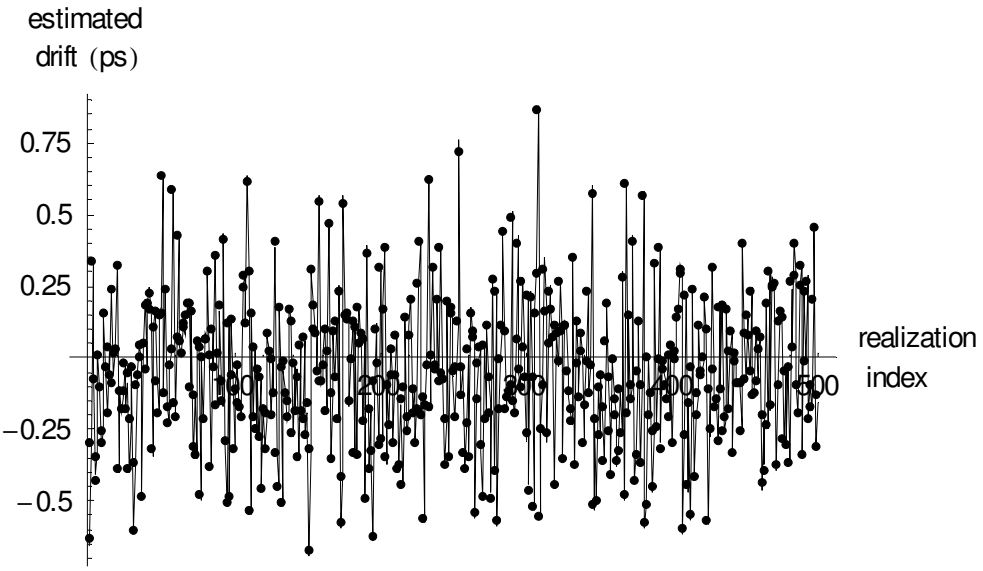

Figure 4-44. Time base jitter interpreted as time base drift (full line: “naive” estimate, dots: enhanced estimate).

Step 5: measurements

The measurement data correspond to the combined impulse response of an opto-electrical converter and a 50 GHz sampling oscilloscope (see figure 4-1 on page 4-12). The impulse response measurement and the required time base distortion measurements were performed by Tracy Clement at NIST. This contribution is gratefully acknowledged.

A time record of 5 ns was used starting at 143 ns delay and 5000 records of 4096 points were acquired.

Time base drift estimation

The drift is estimated, using the first measurement as reference. Based on the fact that simulations confirmed that time base jitter may incorrectly be interpreted as time base drift (see above) and based on reasonable time constants corresponding to thermal effects, the estimated drift is smoothed.

The thermal time constant of the sampling oscilloscope is assumed to be of the order of minutes. The collection and storage of 5000 records of 4096 points is found to take 32 minutes. As such, approximately 150 records are collected per minute. Also, practical experience shows that the shape of the time base drift strongly depends on the environment, both with respect to temperature variations and airflow.

The smoothing can either be done using a global model such as a low-order polynomial or by applying a local model such as a moving average window. The latter has the disadvantage that some realizations at both edges cannot be used, because no moving average exists for these realizations.

Figure 4-45 shows the estimated time base drift and its smoothed version using a 4^{th} order polynomial. The residue (figure 4-46) has a standard deviation of 0.30 ps and as such approximates the 0.26 ps found earlier[1] based on a jitter standard deviation of 1 ps.

Figure 4-47 shows the equivalent when using a moving average window, which is 101 realizations wide. The standard deviation of the residue (figure 4-48) is slightly smaller: 0.28 ps.

It is unclear which smoothing method is to be preferred. It looks like the measurements at NIST were performed under very good conditions with respect to temperature variations and airflow. Given the disadvantage of the moving average with respect to the loss of measurements, the smoothing based on the 4^{th} order

1. see end of step 4 on page 4-48.

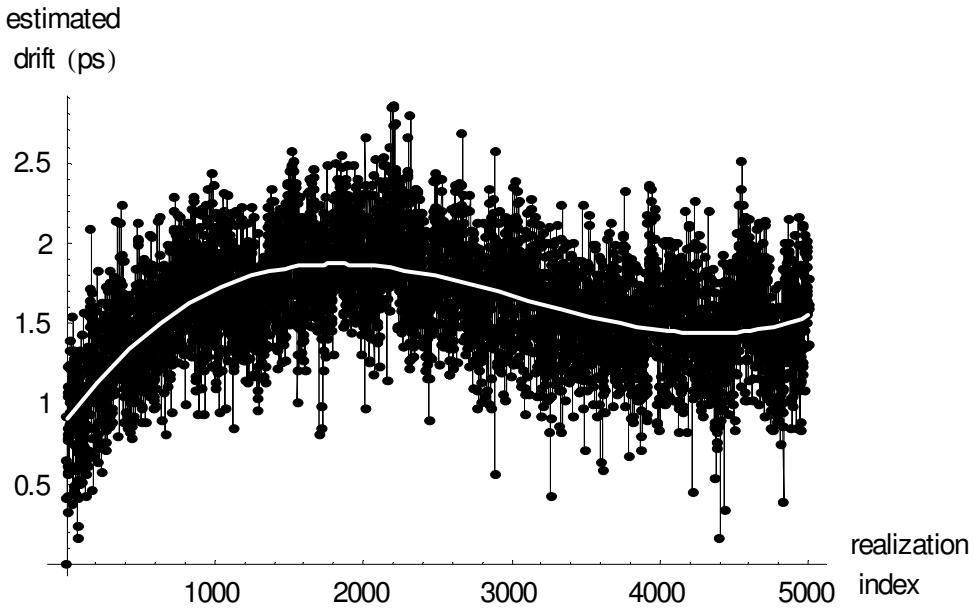

Figure 4-45. Estimated time base drift (white line: smoothed using 4th order polynomial)

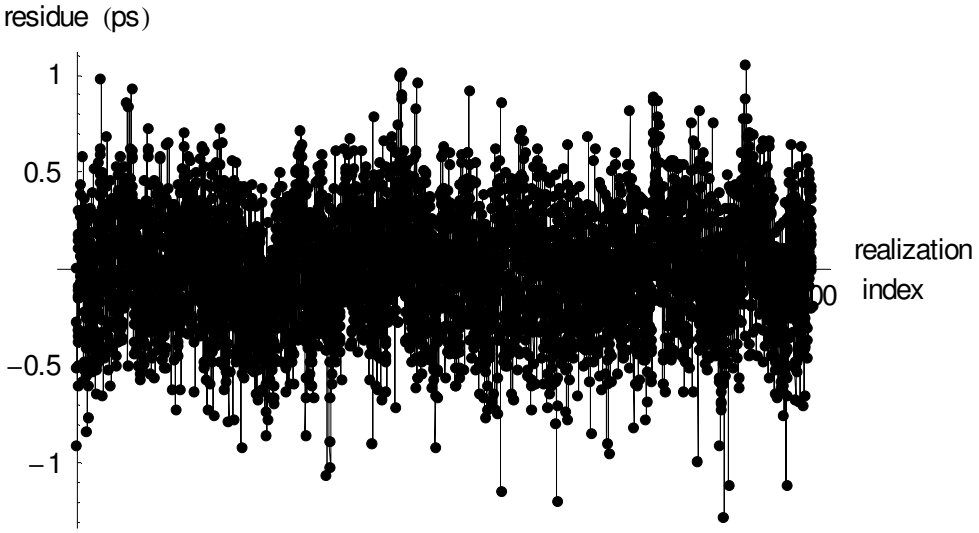

Figure 4-46. Residue (estimated drift minus smoothed drift (4th order polynomial)).

polynomial is preferred. As a verification, the complete processing of the data was repeated using the moving average window. It was found that the resulting estimates of the standard variation of both jitter and additive noise matched within their 95% confidence intervals. As such, for the measurements performed at NIST, the selection of the smoothing method is found not to be critical.

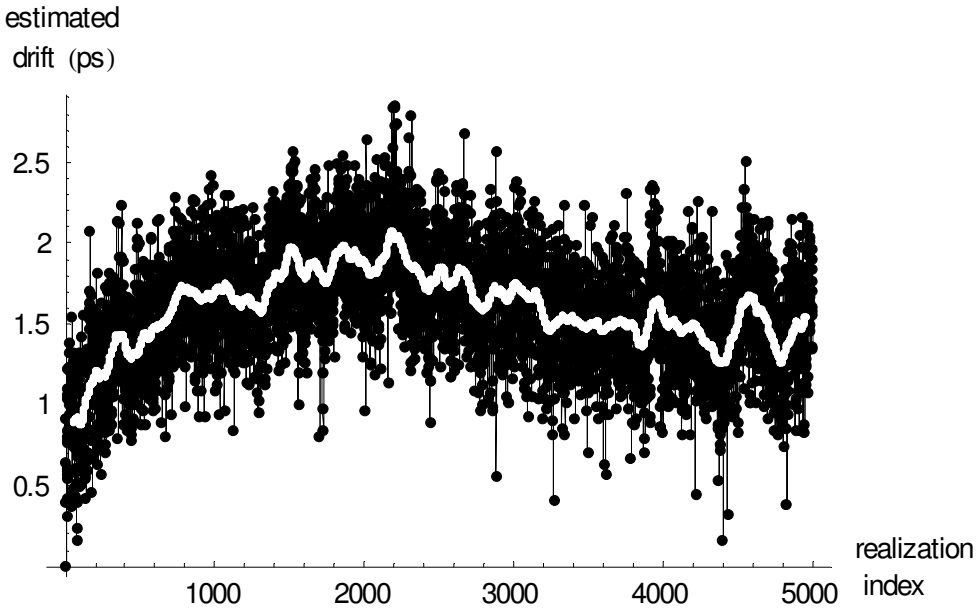

Figure 4-47. Estimated time base drift (white line: smoothed using moving average)

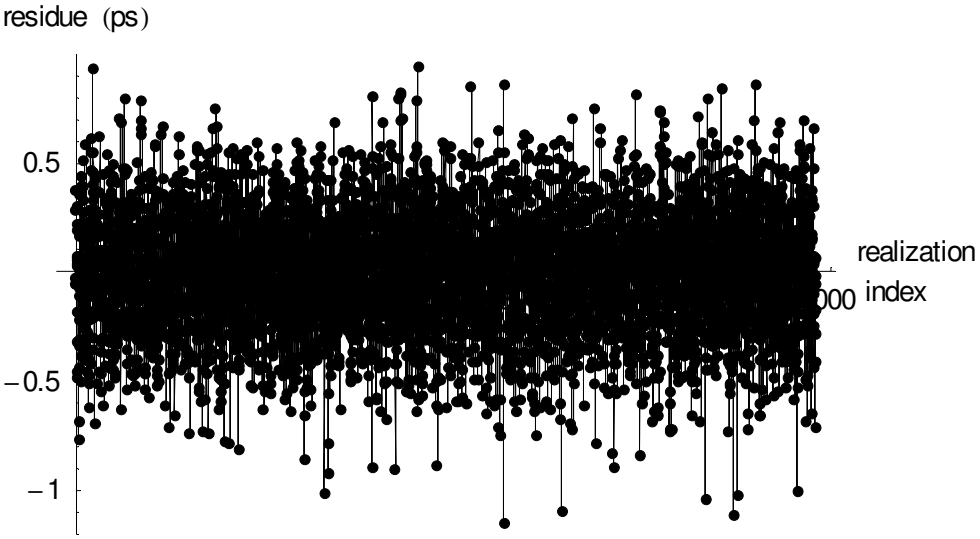

Figure 4-48. Residue (estimated drift minus smoothed drift (moving average)).

Time base drift compensation

The smoothing, based on the 4th order polynomial, is used to align the different realizations. First all realizations are aligned within $\pm 0.5\Delta t$, as shown in figure 4-49. This does not introduce any shaping of the variance, as only shifts over an integer number of samples are used.

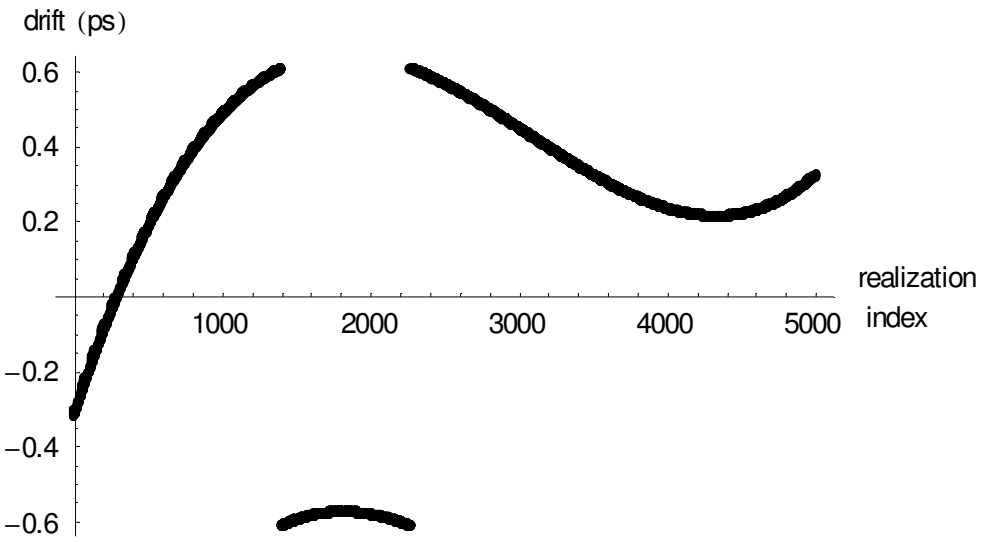

Figure 4-49. Drift compensated within ± 0.5 Δt.

Next, the realizations are divided in 5 buckets, each being $\pm 0.1\Delta t$ wide. All realizations within each bucket are aligned with respect to the center of that bucket. Given a maximum delay of $0.1\Delta t$, the resulting shaping of the variance that is introduced by the alignment can be neglected.

Figure 4-50 shows the unequal distribution of the 5000 realizations over the different buckets.

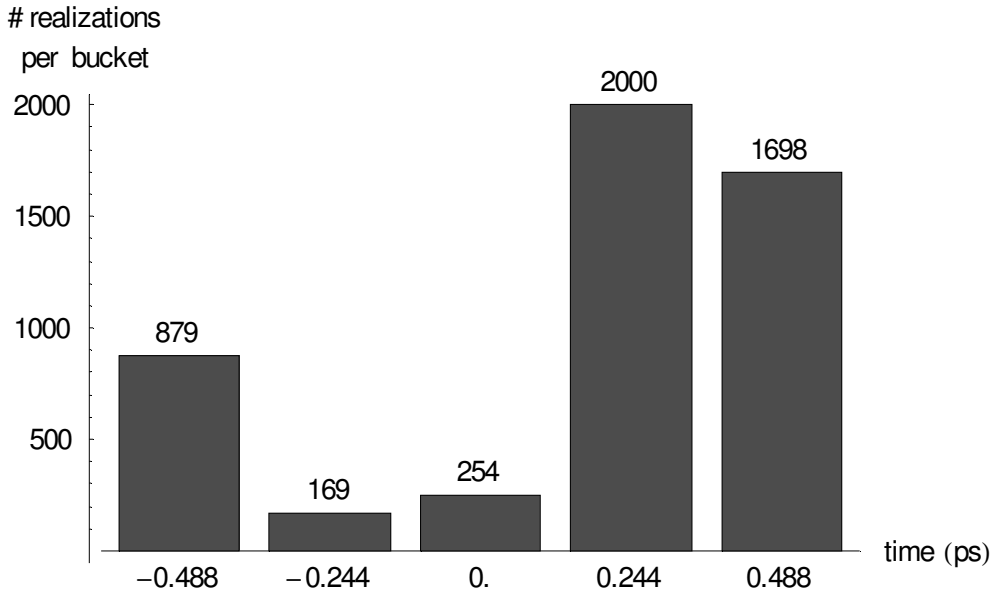

Figure 4-50. Distribution of the 5000 realizations over the different buckets.

In order to obtain variance information, required by the WLS estimator, the realizations per bucket must be divided into different data sets. The effect of using the sample variance, instead of the exact variance, on the expected value and variance of the cost and on the uncertainty of the parameters as function of the number of the data sets, is studied in [6].

In order to minimize the increase of these parameters, the number of the data sets should be sufficiently large. On the other hand, the number of realizations within each data set should be sufficiently large too, because of the $\chi^2_{N_s-1}$ - distribution of the sample variance, N_s representing the number of realizations of the pulse used to calculate the sample variance. This distribution approaches a normal distribution (assumed by [6]), only if N_s is sufficiently large.

A possible compromise is to use a square root law. Based on the smallest number of realizations per bucket, as indicated in figure 4-50, both the number of data sets and the number of realizations per data set is set to $\sqrt{169} = 13$.

Another possible criterion is the allowed increase of the uncertainty of the parameters. In order to limit this increase to 10%, the number of data sets N_w must at least be 25 ($\sqrt{\frac{N_w - 1}{N_w - 5}} \leq 1.1$). The corresponding increase of the expected value of the cost is also smaller than 10%, while the uncertainty on the cost increases by less than 20%.

The jitter was estimated using two different number of data sets, i.e. 13 and 25, based on the above mentioned criteria.

Figure 4-51 shows the sample mean of both the sample mean and sample variance of the 5 times oversampled pulse, each based on 25 data sets. Depending on the bucket index, the number of realizations per data set varies from 6 to 80. As such the uncertainty on the sample variance becomes a function of the bucket index.

Time base distortion estimation and compensation

Due to the time base distortion of the Agilent 83480A Digital Communication Analyzer, the sample mean and the sample variance are specified on a non-equidistant time grid.

The actual time jitter estimation algorithm does not require the sample mean and sample variance to be specified on an equidistant time grid. However, in order to efficiently calculate the derivatives of the mean pulse via the frequency domain using an FFT, cubic interpolation is used to obtain the values of the sample mean on an

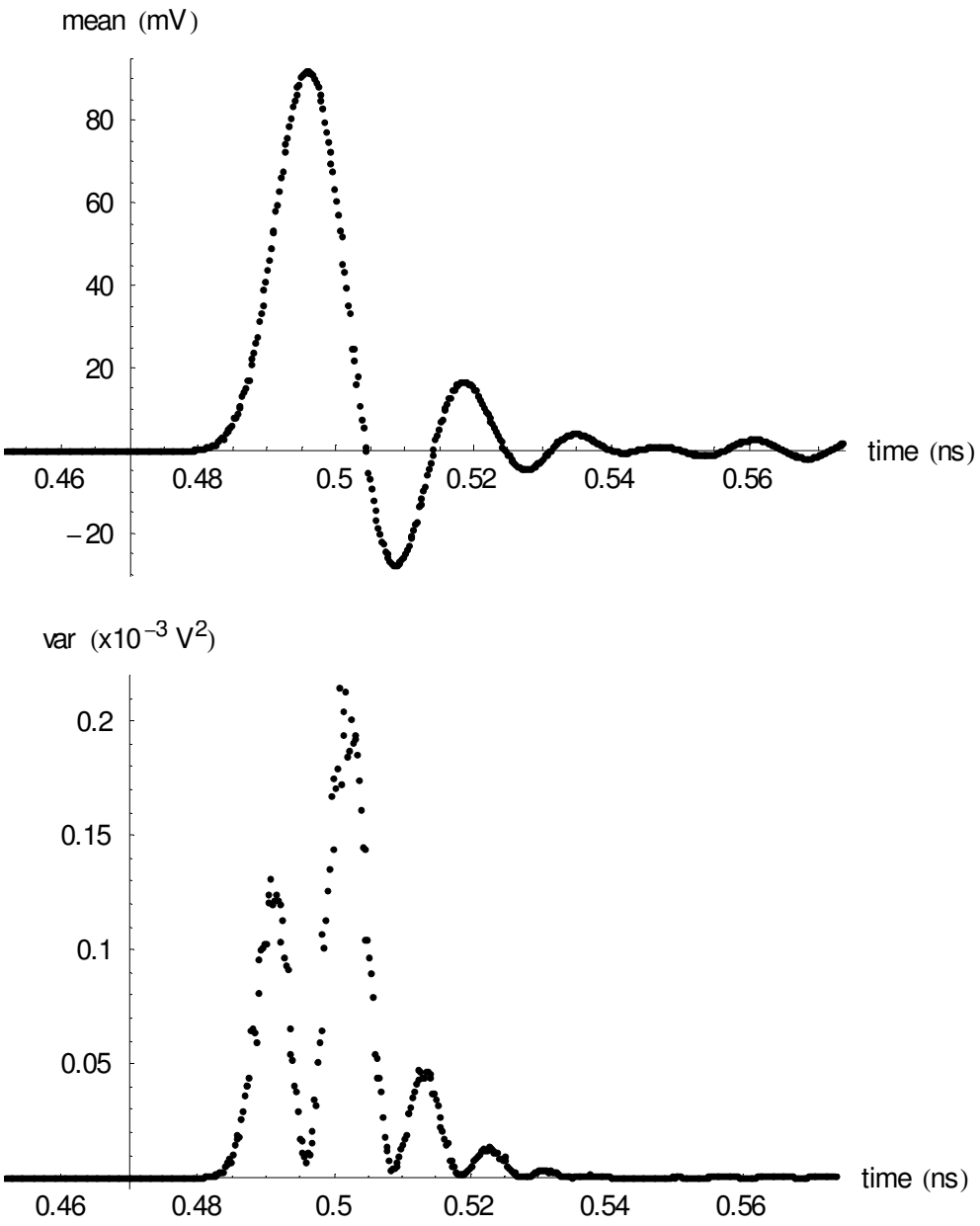

Figure 4-51. Sample mean of sample mean and sample variance of the oversampled pulse, each based on 25 data sets.

equidistant grid, as proposed in [9]. Finally, the derivatives are obtained at the original non-equidistant grid, again by applying the “inverse” cubic interpolation. As a result, some of the points at both edges of the time record may be lost when extrapolation is not allowed.

The time base distortion is estimated first, using [10], after collecting the required data and making sure that both the trigger rate and time base settings are identical to those used during the combined impulse response measurement of the opto-electrical converter and the sampling oscilloscope.

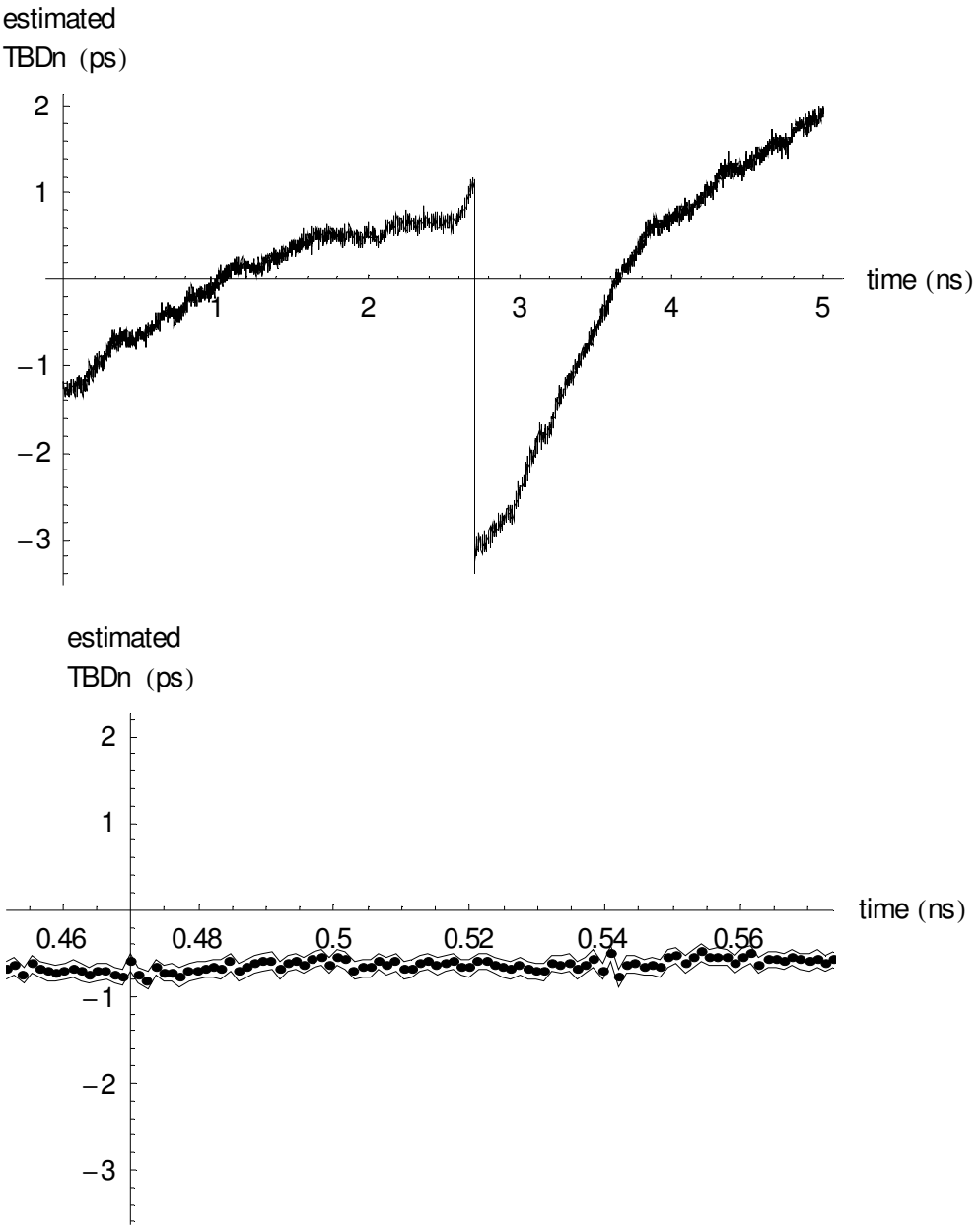

Figure 4-52. Estimated time base distortion (lower plot = zoom into main portion of pulse + 95% confidence interval).

Figure 4-52 shows the estimated time base distortion, defined as $TBDn[i] = t_i - i \cdot \Delta t$, the non-equidistant time stamps being represented by t_i.

For an ideal time base, all values $TBDn[i]$ equal zero. The lower plot zooms into the portion of the time base where the main portion of the pulse is located. All points, contributing to the main pulse, are approximately shifted 0.6 ps to the left with respect to the ideal time base. Therefore their relative distance with respect to each other is very close to Δt. As such it can be concluded that in this case the influence of the

time base distortion is minimal, because the main portion of the pulse does not suffer from any significant time base distortion, it is only shifted over a constant delay.

Applying cubic interpolation to both sample mean y_i and sample variance $\sigma^2_{t_i}$, is statistically incorrect. Indeed, the implemented cubic interpolation uses a linear combination of the four surrounding non-equidistant points to calculate the value at the equidistant grid:

$$y_{k\Delta t} = \sum_{i=1}^{4} a_i \cdot y_i \qquad \text{Eq. 4-39}$$

The correct corresponding variance equals

$$\sigma^2_{k\Delta t} = \sum_{i=1}^{4} a_i^2 \cdot \sigma^2_{t_i} \neq \sum_{i=1}^{4} a_i \cdot \sigma^2_{t_i} \qquad \text{Eq. 4-40}$$

Using the first order approximation to describe the variance $\sigma^2_{t_i}$ (Eq. 4-3), the variance on the equidistant grid becomes

$$\sigma^2_y(k \cdot \Delta t) = \sum_{i=1}^{4} a_i^2 \cdot \sigma^2_{n_y} + \sum_{i=1}^{4} a_i^2 \cdot \left(\frac{dy_0}{dt}\right)^2 \Bigg|_{t=t_i} \cdot \sigma^2_{n_t}. \qquad \text{Eq. 4-41}$$

This still requires to calculate the derivatives at the non-equidistant grid. As such it is easier to remain on the non-equidistant grid.

Time base jitter estimation

Finally, the time base jitter can be estimated using the proposed model (Eq. 4-6) and a weighted least squares estimator, taking the varying variance of the sample variance as function of the bucket index into account, in order to minimize the uncertainty on the estimated parameters.

Another advantage of the weighted least squares is the fact that one is able to calculate the expected value of the cost and its 95% confidence interval as a mean to verify the presence of model errors or other anomalies.

Figure 4-53 shows the correctly aligned “measured” variance on a logarithmic vertical scale. It clearly shows the increased variance due to jitter corresponding to the main pulse and to the reflection about 0.8 ps after the main pulse.

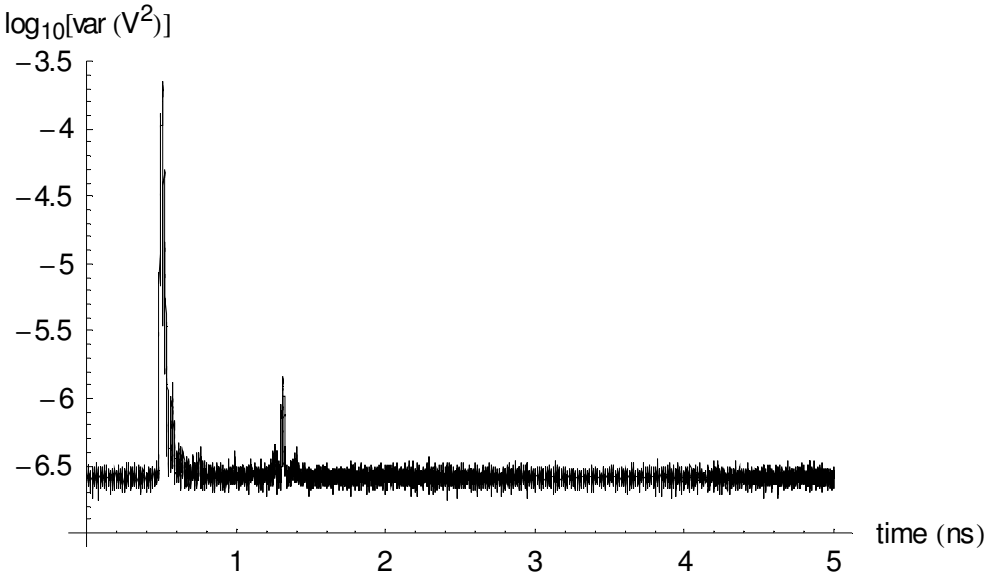

Figure 4-53. Correctly aligned “measured” variance (logarithmic vertical scale).

Given 25 data sets, 4096 time points[1] and 2 parameters, the expected value of the cost and its 95% confidence interval equals 22324 ± 474. The actual value of the cost function for a third order model is found to be 23582. Its value is less than 6% larger than the expected value of the cost and slightly outside the 95% confidence interval. With respect to the first order model, the cost is decreased by more than 13%.

The estimated variance of the additive noise equals $(0.258\ 10^{-6} \pm 0.163\ 10^{-9})\ V^2$ and that of the jitter noise is $(0.923\ 10^{-24} \pm 4.99\ 10^{-27})\ s^2$. The uncertainties correspond to the 95% confidence interval.

The corresponding standard deviation values for the additive and jitter noise and their 95% confidence intervals are 0.508 mV ± 0.16 µV and 0.961 ps ± 2.6 fs.

It is time to compare the modelled variance to the measured variance and its 95% confidence interval based on the sample variance of the measured variance. While figure 4-54 zooms into the main pulse, figure 4-55 demonstrates an equally excellent fit for the first reflection. Figure 4-56 and figure 4-57 show the variance at both edges of the record and demonstrate that also the additive noise portion fits exceptionally well. Finally, figure 4-58 zooms into the maximum value of the variance to show the

1. The actual number of time points turns out to be 20466, based on the oversampling by a factor 5 and the loss of 14 points at both edges based on the cubic interpolation followed by an inverse cubic interpolation to obtain the derivatives.

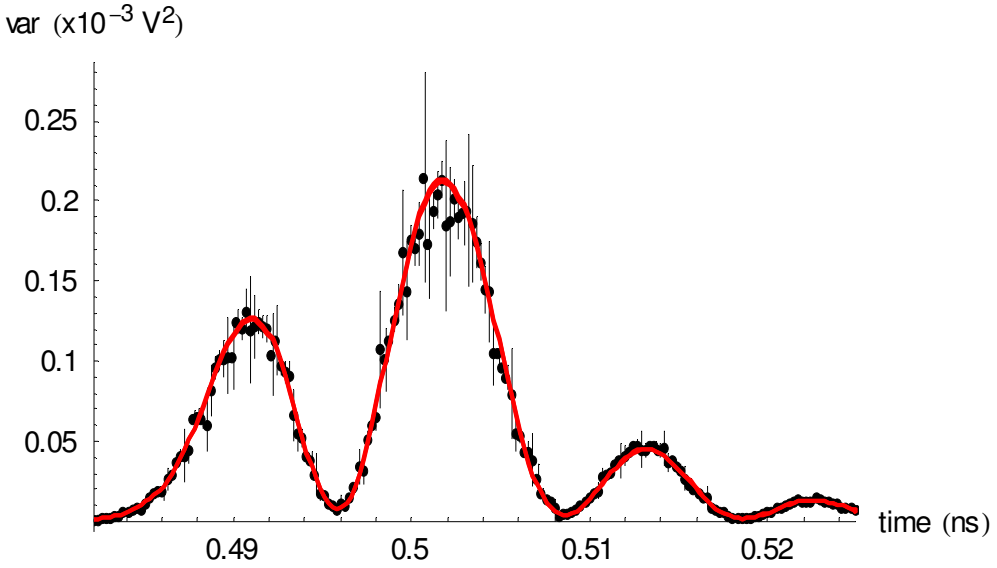

Figure 4-54. Comparing (the boundaries of the 95% confidence interval of) the (3rd order model, WLS estimator) modelled variance (red lines) and the correctly aligned "measured" variance at the main pulse (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the sample variance of the sample variance).

uncertainty of the estimated variance as a result of the uncertainty on the estimated parameters.

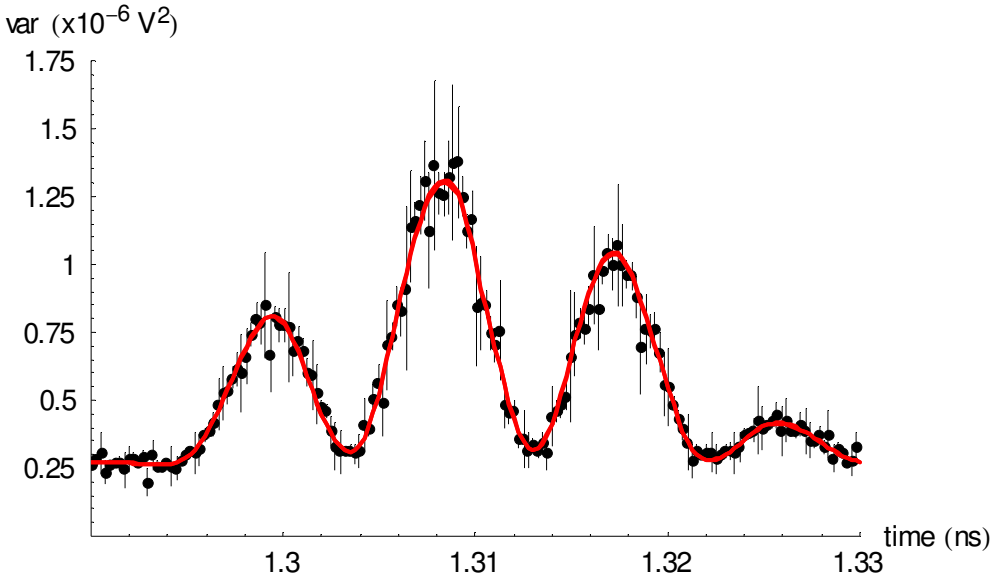

Figure 4-55. Comparing (the boundaries of the 95% confidence interval of) the (3rd order model, WLS estimator) modelled variance (red lines) and the correctly aligned "measured" variance at the first reflection (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the sample variance of the sample variance).

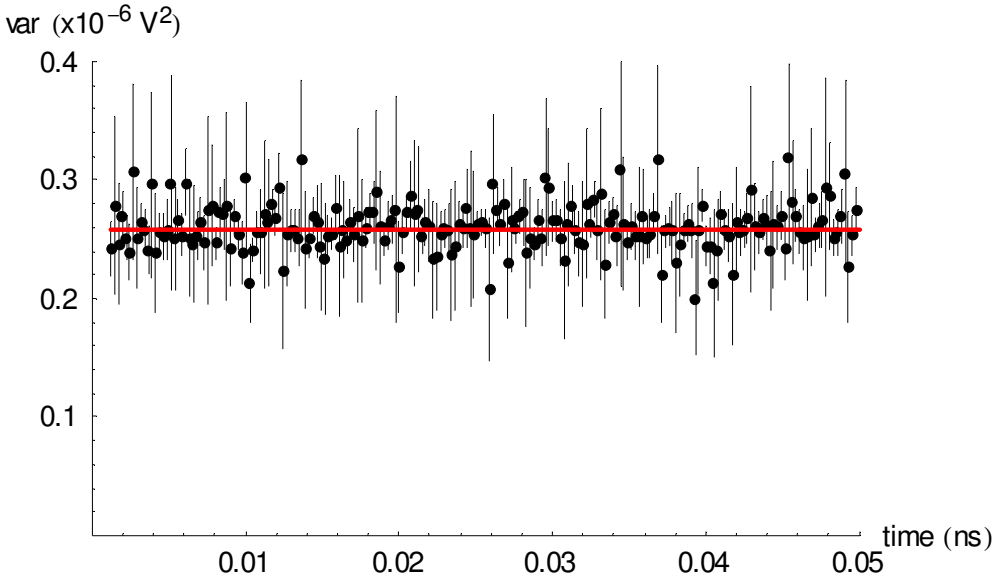

Figure 4-56. Comparing (the boundaries of the 95% confidence interval of) the (3rd order model, WLS estimator) modelled variance (red lines) and the correctly aligned "measured" variance (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the sample variance of the sample variance), zooming into the pedestal corresponding to the variance of the additive noise at the left edge of the record.

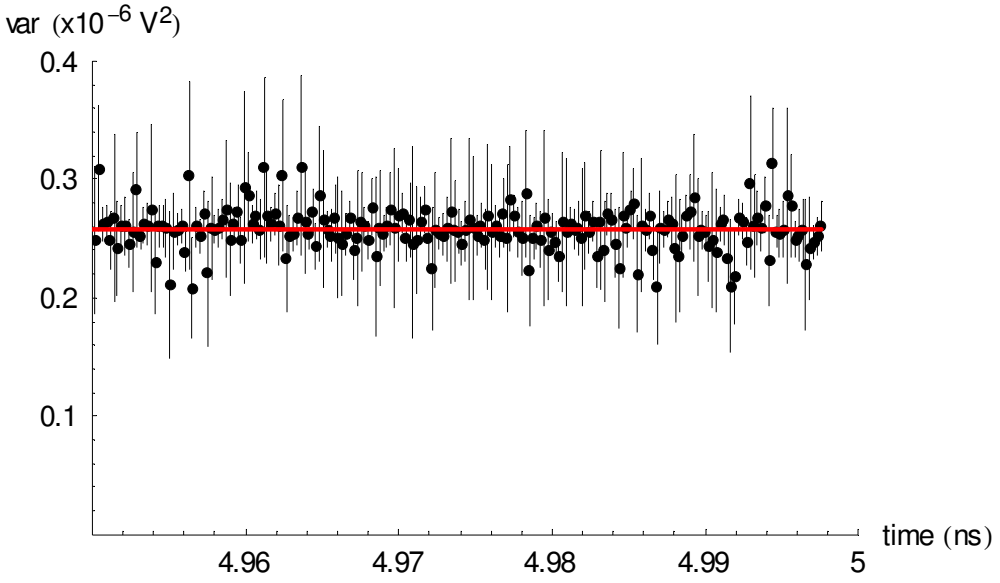

Figure 4-57. Comparing (the boundaries of the 95% confidence interval of) the (3rd order model, WLS estimator) modelled variance (red lines) and the correctly aligned "measured" variance (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the sample variance of the sample variance), zooming into the pedestal corresponding to the variance of the additive noise at the right edge of the record.

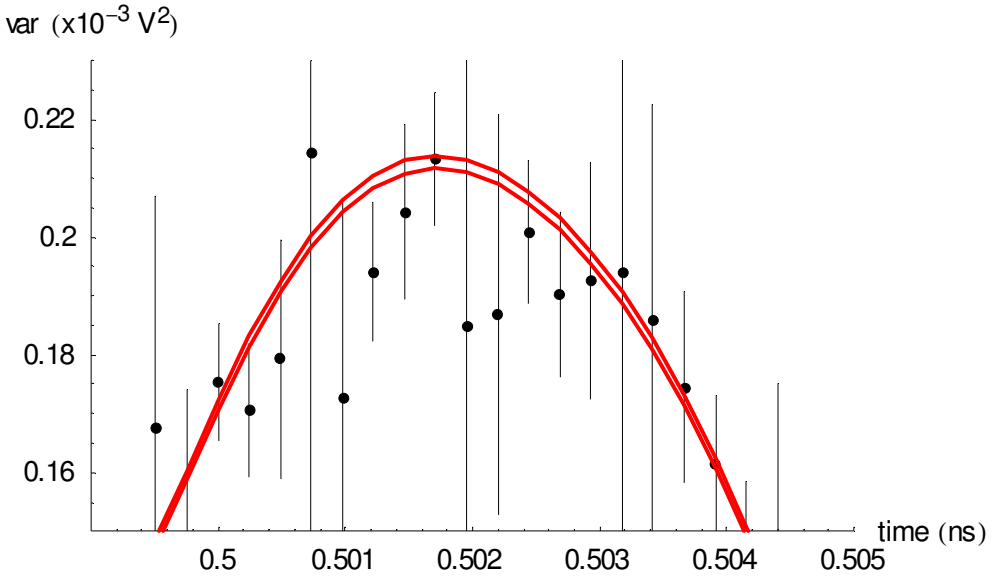

Figure 4-58. Comparing the boundaries of the 95% confidence interval of the (3rd order model, WLS estimator) modelled variance (red lines) and the correctly aligned "measured" variance (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the sample variance of the sample variance), zooming into the maximum value.

The correlation coefficient between the estimated variance of additive and jitter noise turns out to be -0.12. As such there is no statistical evidence of a significant correlation between both estimated parameters.

Repeating the estimation based on 13 data sets, the standard deviation of the additive and jitter noise and their 95% confidence intervals are 0.508 mV ± 0.15 µV and 0.962 ps ± 2.5 fs. As such, selecting a different number of data sets has no effect on the final estimate. The expected and realized value of the cost and its variance however did change, according to [6]. The realized cost turns out to be 25597, 4% above the expected value of the cost and slightly outside the 95% interval: 24557 ± 583.

The power of a solid stochastical framework

It is shown that, using a solid stochastical framework, one is able to detect anomalies like overlooking the shaping of the variance due to drift compensation. This is similar to the power demonstrated by [10] to capture anomalies with respect to the time base distortion: due to the fact that the realized cost was significantly larger than the expected value of the cost, it was found that, for that particular setup, the time base distortion varied with the applied calibration frequency, while the model assumes that this time base distortion is identical for all applied calibration frequencies.

Suppose the shaping of the variance due to time base drift compensation is overlooked. Repeating the above estimation procedure, using 70 data sets according to the "square root" rule starting from 5000 realizations, but not using any

oversampling to limit the drift compensation to $\pm 0.1\Delta t$, the expected value of the cost and its 95% confidence interval equals 4212 ± 188. The realized cost using a third order model turns out to be 14764, which is 3.5 times the expected value of the cost. As can be seen in figure 4-59, this is confirmed by a poor correspondence between the model and the variance based on incorrectly processed measurements.

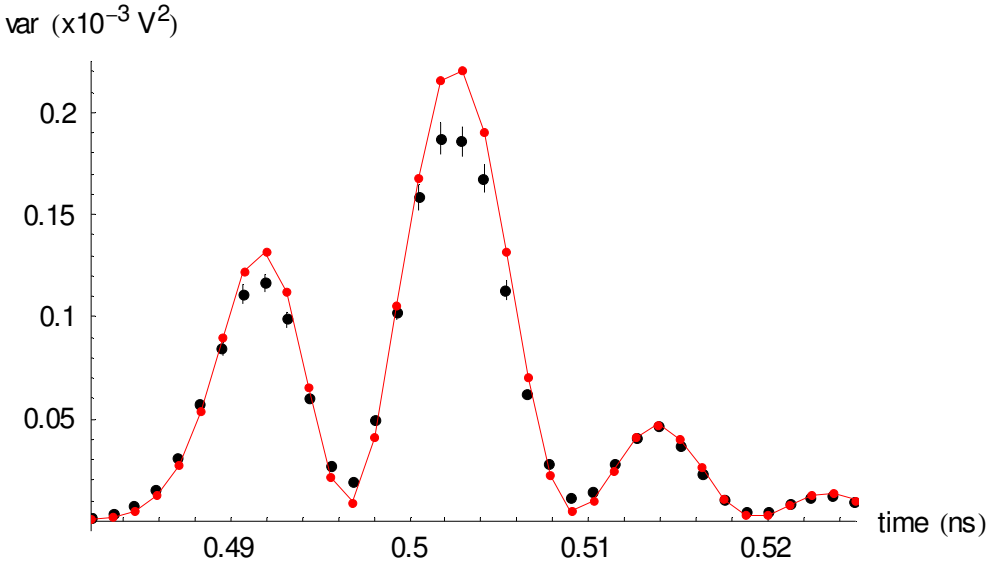

Figure 4-59. Comparing the (3rd order model, WLS estimator) modelled variance (connected red dots) and the incorrectly shaped "measured" variance (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the sample variance of the sample variance).

Using the knowledge of the χ^2 - distribution of the variance of the sample variance of the pulse, one can avoid the calculation of the sample variance of the sample variance and as such reduce the number of required measurements. Figure 4-60 compares the modelled variance to the measured variance and its 95% confidence interval based on a χ^2 - distribution. The value of the cost function turns out to be 7% larger than the expected value of the cost and is located slightly outside the 95% confidence interval of the expected value of the cost. The standard deviation of the estimated additive and jitter noise and their 95% confidence intervals are $0.508\ \text{mV} \pm 0.16\ \mu\text{V}$ and $0.965\ \text{ps} \pm 2.3\ \text{fs}$ and these match the values, obtained by using the sample variance of the sample variance.

ML estimation

Based on Eq. 4-19, the MLE implementation can take care of the unequal distribution of the 5000 realizations over the different buckets. The corresponding varying uncertainty is taken into account by using the appropriate number of degrees of freedom as function of the bucket index.

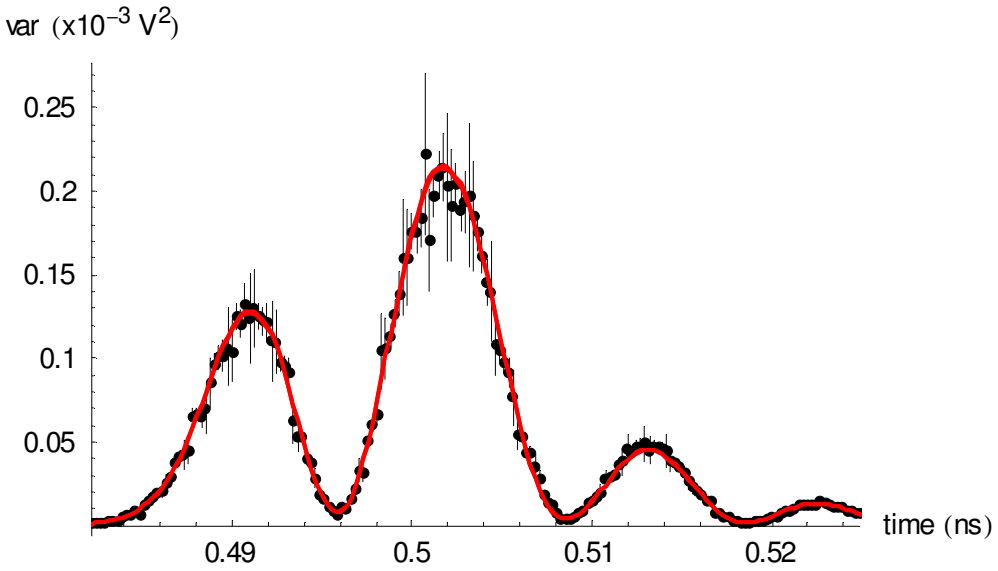

Figure 4-60. Comparing (the boundaries of the 95% confidence interval of) the (3rd order model, WLS estimator) modelled variance (red lines) and the correctly aligned "measured" variance (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the χ^2 - distribution of the sample variance (known number of degrees of freedom)).

Based on a third order model, the estimated jitter standard deviation is found to be 0.970 ps ± 2.3 fs, while the estimated standard deviation of the additive noise is found to be 0.509 mV ± 0.16 µV. Compared to the result of the WLS estimator (0.508 mV ± 0.16 µV and 0.965 ps ± 2.3 fs), it is clear that the 95% confidence intervals almost overlap. Furthermore, the uncertainty of the parameters for the maximum likelihood estimator is found to equal that of the WLS estimates. Also, the crosscorrelation between both parameters turns out to be -0.11 as was the case for the WLS estimator.

As such, excellent correspondence can be claimed between the estimated standard variation of both jitter and additive noise, based on the WLS estimator and the ML estimator. The WLS estimator is to be preferred because the actual value of the cost can be used to detect model errors by comparing it to its expected value. Interpreting the value of the cost based on the ML estimator is found to be less obvious.

LS estimation

In step 2 and 3 of the simulation, the LS estimator provided the best predictions for the jitter standard deviation. Used in combination with the third order model and based on the correctly aligned measured variance, the estimated jitter standard deviation turns out to be 0.955 ps ± 10.5 fs, while the estimated standard deviation of the additive noise is found to be 0.514 mV ± 4 µV. Compared to the result of the WLS estimator (0.508 mV ± 0.16 µV and 0.965 ps ± 2.3 fs), it is clear that the 95% confidence intervals overlap for the jitter estimation and almost overlap for the additive noise

estimation. Furthermore, as expected, the uncertainty of the parameters is obviously larger for the LS estimate than for the WLS estimate. Finally, the crosscorrelation between both parameters turns out to be -0.82, which is significantly larger than -0.11 using the WLS estimator.

Again, excellent correspondence can be claimed between the estimated standard variation of both jitter and additive noise, based on the WLS estimator and the LS estimator. The WLS estimator is to be preferred because the value of the cost can be used to detect model errors, while this is impossible for the LS estimator. Furthermore, the uncertainty on both estimated parameters is significantly smaller using the WLS estimator.

Figure 4-61 and figure 4-62 clearly show the increased uncertainty of the modelled variance with respect to the WLS estimator.

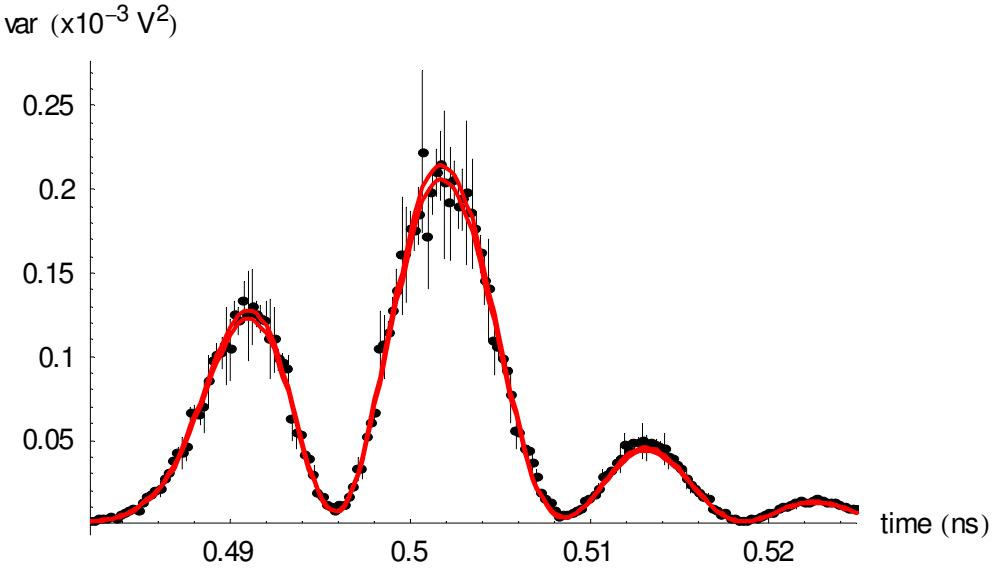

Figure 4-61. Comparing the boundaries of the 95% confidence interval of the (3rd order model, LS estimator) modelled variance (red lines) and the correctly aligned “measured” variance (black dots and vertical black lines, boundaries of the 95% confidence intervals based on the χ^2 - distribution of the sample variance (known number of degrees of freedom)).

Bias in estimation of variance of additive noise

Finally, it is mentioned that overlooking the shaping of the variance, due to time base drift compensation, in combination with a first order model and a (unweighted) least-squares estimator, introduces a bias of more than 10% on the estimate of the variance of the additive noise (figure 4-63). This may explain why in [3] the additive noise is estimated separately.

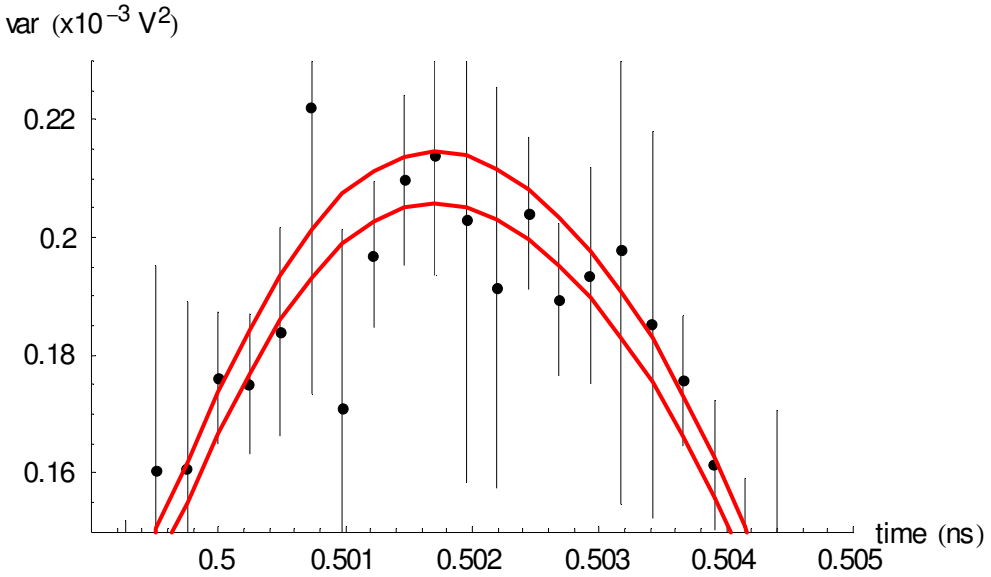

Figure 4-62. Comparing the boundaries of the 95% confidence interval of the (3rd order model, LS estimator) modelled variance (red lines) and the correctly aligned "measured" variance (95% confidence interval of the "measured" variance based on the χ^2 - distribution of the sample variance (known number of degrees of freedom), zooming into the maximum.

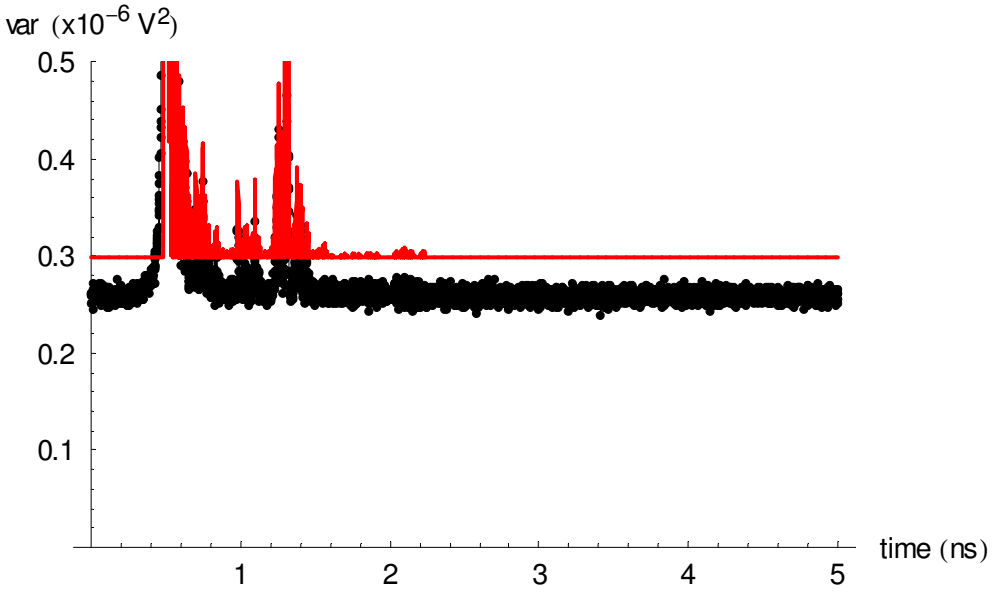

Figure 4-63. Zooming into the additive noise portion contribution, showing a bias on the estimated variance (red line) for a 1st order model, LS estimator and incorrectly aligned "measured" variance.

It is found that this offset is removed by properly aligning the pulses, as is done in the methods proposed here.

Conclusions

The system identification approach described in this chapter and applied to jitter estimation of the combined impulse response of an opto-electrical converter and a high-speed sampling oscilloscope is a major extension of [2], which can be applied to "real" problems. Indeed, the simulation results presented in [2] are based on the rather unrealistic assumption that Eq. 4-9 is the exact representation of the sample variance, while in reality it is only a first order approximation.

The underlying stochastical framework allows to detect model errors and anomalies like the shaping of the variance due to time base drift compensation. Error bounds are provided on both the estimated parameters and the modelled variance. Finally the method allows the simultaneous estimation of the variance of the additive noise and the jitter noise, where other methods [3] fail to do so.

Future research

The described method first aligns all realizations by applying a delay which is an integer multiple of Δt. In order to obtain a more uniform distribution of the realizations over the different buckets, one can shift this time window between (-Δt, 0) and (0, Δt), instead of selecting the default (-$\frac{\Delta t}{2}$, +$\frac{\Delta t}{2}$).

Furthermore, based on its support for non-equidistant time grids, the method can easily be extended to handle situations where one or more buckets have no or an insufficient number of realizations.

References

[1] T. Souders, D.Flach, C. Hagwood and G. Yang, “The Effects of Timing Jitter in Sampling Systems,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 39, No. 1, February 1990

[2] G. Vandersteen and R. Pintelon, “Maximum Likelihood Estimator for Jitter Noise Models,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 49, No. 6, December 2000

[3] K. Coakley, C.-M. Wang, P. Hale and T. Clement, “Adaptive Characterization of Jitter Noise in Sampled High-Speed Signals,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 52, No. 5, October 2003

[4] Mathematica, Wolfram Research

[5] “Vector Statistical Library Notes. Intel® Math Kernel Library,” Version 6.0, July 2004

[6] J. Schoukens, R. Pintelon and Y. Rolain, “Maximum Likelihood Estimation of Errors-In-Variables Models using the Sample Covariance Matrix Obtained from Small Data Sets,” published as part of “*Recent Advances in Total Least Squares Techniques and Errors-In-Variables Modeling*”, Sabine Van Huffel (editor), Siam, Philadelphia, 1997

[7] J. Schoukens, “Inleiding in de waarschijnlijksheidsrekening en de statistiek,” Vrije Universiteit Brussel

[8] K. Coakley and P. Hale, “Alignment of Noisy Signals,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No 1., February 2001

[9] Y. Rolain, J. Schoukens and G. Vandersteen, “Signal Reconstruction for Non-Equidistant Finite Length Sample Sets: a “KIS” approach,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 47, No. 5, October 1998, pp. 1046 - 1052

[10] G. Vandersteen, Y. Rolain and J. Schoukens, “An Identification Technique for Data Acquisition Characterization in the Presence of Nonlinear Distortions and Time Base Distortions,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No. 5, October 2001

CHAPTER 5

System identification approach applied to drift estimation.

Abstract

When collecting a large number of repeated measurement records of the impulse response of a linear time-invariant system using a high-frequency sampling oscilloscope, it was found that the successive measurements of this impulse response slightly shift over time, within the acquisition window. This phenomenon is referred to as time base drift.

Time base drift and its estimation in the presence of both additive and jitter noise has already been mentioned in this work as part of the “Enhancements to the nose-to-nose calibration technique” and during the “System identification approach applied to jitter estimation”. In this initial approach, the first measurement record is used as a reference signal during the alignment of successive measurements. This gives estimates, which are comparable to the ones based on the method proposed in [1]. Here, an enhanced version of the initial approach is proposed. In this method, the aligned average is used as a reference signal instead of the first measurement.

A system identification approach is applied to estimate the time base drift introduced by a high-frequency sampling oscilloscope. First, a new least squares estimator is proposed to estimate the delay of a set of repeated measurements in the presence of additive and jitter noise. Next, the effect of both additive and jitter noise is studied in the frequency domain using simulations.

Special attention is devoted to the covariance matrix of the experiments. The use of this matrix allows to come up with a good estimate of the uncertainty on the estimated delays. Using the covariance matrix, a weighted least squares estimator is implemented to minimize the uncertainty on the estimated delays. Comparative results are shown based on simulations proposed by [1].

Finally, the enhanced method is applied to estimate the drift using the same measurements as those that were used during the jitter estimation. The impulse response of an opto-electrical converter is measured using an Agilent 83480A sampling oscilloscope in combination with a 83484A 50 GHz electrical plug-in.

Modelling and estimating drift in the presence of additive and jitter noise

High-frequency sampling oscilloscopes often use an equivalent-time sampling principle to diminish the needed conversion rate. They suffer from both additive measurement noise $n_y(t_i)$ and timing jitter noise $n_t(t_i)$ at the sampling time instant t_i. Furthermore, it is observed that these oscilloscopes also suffer from time base drift. Time base drift is due to imperfections on the position of the trigger point relative to the signal and results in a movement of the signal in the acquisition window. As a result, successive measurements correspond to delayed versions of the "exact" signal. The delay τ_k models this effect and varies with respect to the realization index k. The resulting signal model is

$$y_k(t_i) = y_0(t_i + n_t(t_i) + \tau_k) + n_y(t_i) \qquad \text{Eq. 5-1}$$

Herein, $y_k(t_i)$ represents the k-th measurement of the exact signal $y_0(t_i)$ when both additive noise and jitter are added as a part of the measurement. Both $n_y(t_i)$ and $n_t(t_i)$ are assumed to be zero mean, normally distributed, independent and stationary with respect to t, and as a result they are also independent with respect to realization index k. Furthermore, it is assumed that $t_i = i\Delta t$. Δt represents the sampling period.

Let $Y_k(\omega_m)$ correspond to the discrete Fourier transform of $y_k(t_i)$. As part of the enhancements to the nose-to-nose calibration technique and the system identification approach applied to jitter estimation, the time base drift of each acquisition was estimated using the first acquisition as a reference signal. In the approach proposed here, the (unknown) exact signal is used as the reference signal. The LS cost function then becomes

$$V_{LS} = \sum_{k,\,m} \left| Y_k(\omega_m) - e^{-j\omega_m\tau_k} \cdot Y_0(\omega_m) \right|^2 \qquad \text{Eq. 5-2}$$

Eq. 5-2 must be minimized with respect to both the unknown delays τ_k and the Fourier coefficients of the unknown exact signal $Y_0(\omega_m)$. Since the spectra $Y_0(\omega_m)$ appear linearly in the equation error, it is possible to eliminate them from the cost. Therefore, the symbolic derivative of Eq. 5-2 with respect to $Y_0(\omega_m)$ is set to zero[1].

$$\frac{\partial V_{LS}}{\partial Y_0(\omega_m)} = 0 \qquad \text{Eq. 5-3}$$

$$\Leftrightarrow \quad \sum_k [Y_k(\omega_m) - e^{-j\omega_m\tau_k} \cdot Y_0(\omega_m)]^* \cdot (-e^{-j\omega_m\tau_k}) = 0 \qquad \text{Eq. 5-4}$$

Here Y^* corresponds to the complex conjugate of Y. $Y_0(\omega_m)$ can now be calculated as follows:

$$\hat{Y_0}(\omega_m) = \frac{1}{K} \sum_{k=1}^{K} Y_k(\omega_m) \cdot e^{j\omega_m\tau_k}. \qquad \text{Eq. 5-5}$$

K represents the number of repeated realizations of delayed versions of $Y_0(\omega_m)$.

Substituting Eq. 5-5 in Eq. 5-2 and using the fact that $|e^{-j\theta}| = 1$, results in a cost function which is only a function of the unknown delays τ_k:

$$V_{LS}(\tau_1, .., \tau_K) = \sum_{k,m} \left| Y_k(\omega_m) \cdot e^{j\omega_m\tau_k} - \frac{1}{K} \sum_{l=1}^{K} Y_l(\omega_m) \cdot e^{j\omega_m\tau_l} \right|^2 \qquad \text{Eq. 5-6}$$

It is possible to introduce an arbitrary delay τ as follows

$$Y_k(\omega) = e^{-j\omega\tau_k} \cdot Y_0(\omega) = e^{-j\omega(\tau_k+\tau)} \cdot [Y_0(\omega) \cdot e^{j\omega\tau}], k = 1..K, \qquad \text{Eq. 5-7}$$

without influencing V_{LS}. If $\tilde{Y_0}(\omega) = Y_0(\omega) \cdot e^{j\omega\tau}$ is considered to be the new reference, it is clear that there is a degeneracy: (only) one delay can be freely chosen. One possibility is to select one delay to be zero. For example, here it is assumed that $\tau_1 = 0$.

1. Since the function V_{LS} is not an analytical form in Y_0 due to the presence of the complex conjugate, this deserves some additional explanation (see page 5-5).

The derivatives of Eq. 5-6 with respect to τ_2 up to τ_K yield the gradient while the second derivatives yield the Hessian of the cost. The Newton-Raphson iteration scheme is then used to find the estimates $\tau_2, .., \tau_K$ which minimize Eq. 5-6, given that $\tau_1 = 0$.

Starting values are readily available from the initial implementation[1] (also imposing $\tau_1 = 0$) which uses the first realization $Y_1(\omega)$ as reference signal, instead of the average signal, as is obtained after the alignment of the records.

Although these starting values reduce the number of iterations that are needed to converge to the solution, it was found that zero starting values also do the job. Even starting values where the sign of the delay is incorrect still lead to convergence to the same solution. This shows the robustness of the method to poor starting values.

Symbolic derivation

To justify the use of symbolic derivatives [2] with respect to Eq. 5-2, consider $f(x, x^*) \in \mathbb{R}$ as an alternative for $f(x_R, x_I)$, where x^* represents the complex conjugate of $x = x_R + jx_I$.

$$\frac{\partial f}{\partial x_R} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial x_R} + \frac{\partial f}{\partial x^*} \cdot \frac{\partial x^*}{\partial x_R} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial x^*} \qquad \text{Eq. 5-8}$$

$$\frac{\partial f}{\partial x_I} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial x_I} + \frac{\partial f}{\partial x^*} \cdot \frac{\partial x^*}{\partial x_I} = j\left(\frac{\partial f}{\partial x} - \frac{\partial f}{\partial x^*}\right) \qquad \text{Eq. 5-9}$$

Based on Eq. 5-8 and Eq. 5-9, and given $\frac{\partial f}{\partial x^*} = \left(\frac{\partial f}{\partial x}\right)^*$, it is clear that

$$\frac{\partial f}{\partial x_R} = \frac{\partial f}{\partial x_I} = 0 \Leftrightarrow Re\left[\frac{\partial f}{\partial x}\right] = Im\left[\frac{\partial f}{\partial x}\right] = 0 \Leftrightarrow \frac{\partial f}{\partial x} = 0 \qquad \text{Eq. 5-10}$$

1. As described as part of the “Enhancements to the nose-to-nose calibration technique” and during the “System identification approach applied to jitter estimation”.

Analysis of the noise sources: additive white noise

Is the noise circular complex distributed?

For personal educational purposes, first it is verified that the assumption of circular complex noise in the frequency domain holds in the case of additive noise, which is stationary with respect to time. To this end, realistic simulations will be used. The covariance matrix of the noise in the frequency domain is obtained as a by-product, and is based on covariance information in the time domain.

In order to keep the simulation as realistic as possible, the analytical expression of the impulse response that has been used during the system identification approach applied to jitter estimation is reused here.

500 realizations of this impulse response are generated in the presence of additive noise, which is chosen to be zero mean, is normally distributed and is stationary with respect to time. The standard deviation is chosen to be 3 mV. The delay is set to zero. A scaling factor of $\sqrt{N}$ is used during the calculation of the discrete Fourier transform (N represents the number of time points within each realization of the simulation). Starting from a variance σ^2 = 9x10^{-6} V^2 at each time instant, it is found that 9x10^{-6} falls inside the 95% confidence interval of the resulting sample variance at each frequency.

As can be expected, the additive noise shows up as circular complex white noise in the frequency domain, with $\sigma_R^2 = \sigma_I^2 = \frac{\sigma^2}{2}$ and $\sigma_{RI}^2 = 0$, i.e. the variance of the real part of the spectrum equals that of the imaginary part, while there is no stochastically significant correlation between the real and the imaginary part. Furthermore, the variance does not vary and is uncorrelated as a function of the frequency.

The calculation of the covariance in the frequency domain is based on the linear relationship between the Fourier coefficients and the time samples. Therefore

$$Cov[Y(\omega)] = J \cdot Cov[y(t)] \cdot J^H, \qquad \text{Eq. 5-11}$$

where J^H represents the hermitian transpose of J, with $J_{mi} = \frac{\partial Y(\omega_m)}{\partial y(t_i)}$.

In the case where $t_i = i \cdot \Delta t$ and $\omega_m = \frac{2\pi m}{N \cdot \Delta t}$, $J_{mi} = \frac{1}{\sqrt{N}} e^{-2\pi j \frac{mi}{N}}$, $j = \sqrt{-1}$.

Using Eq. 5-11 in the special case $Cov[y(t)] = \sigma^2 \cdot \mathrm{I}$, where I represents a $N \times N$ identity matrix, one finds the known result: $Cov[Y(\omega)] = \sigma^2 \cdot J \cdot J^H = \sigma^2 \cdot \mathrm{I}$.

Weighted version of V_{LS}

In the case of only additive white noise, $Cov[Y(\omega)] = \sigma^2 \cdot \mathrm{I}$. It is then simple to obtain the expected value of the cost function, which is easy to interpret if Eq. 5-2 is divided by σ^2. A scale factor of $\sqrt{N}$ is consistently used during the calculation of the FFT. This allows simple validation of the model by comparison of the value of the cost function taken in the estimates to the expected cost.

Given K the number of realizations of the measurement and M the number of frequency components used to estimate the K delays, the expected value of the cost equals the number of measurement data minus the number of parameters, i.e. $E\{V_{LS}\} = K \cdot (M-1)$.

As a quick sanity check, three sets of 500 realizations (all τ_k = 0) of the impulse response are generated, one with σ = 0.6 mV and two with σ = 3 mV. The same signal is used as during the preceding tests. If one takes frequency components up to 100 GHz into account, 500 frequency components are used in the case of a frequency resolution of 200 MHz. The expected value of the cost is 249500 while the 95% confidence interval for the χ^2 - distributed cost equals $2\sqrt{2K \cdot (M-1)}$ = 1413. All three realized cost functions (248899, 248900 and 249757) fall inside the 95% confidence interval of the expected cost.

Verification of the uncertainty on the estimated delays

Again, only additive noise with $Cov[Y(\omega)] = \sigma^2 \cdot \mathrm{I}$ is considered to be present. One obtains an estimate of the parameter covariance matrix [3]:

$$Cov[\hat{\tau}] \cong [2re(J^H J)]^{-1} \cdot 2re(J^H Cov[Y(\omega)]\ J) \cdot [2re(J^H J)]^{-1} \qquad \text{Eq. 5-12}$$

$$Cov[\hat{\tau}] \cong [2re(J^H J)]^{-1} \cdot \sigma^2 \qquad \text{Eq. 5-13}$$

Here J corresponds to the Jacobian $J = \frac{\partial e(\tau)}{\partial \tau}$ of the complex error vector $e(\tau)$; $V_{LS}(\tau) = e^H(\tau) \cdot e(\tau)$ is the vector equivalent of Eq. 5-6.

If one considers 500 realizations of the impulse response and 500 frequency components per experiment, the Jacobian is a 250,000 by 499 complex matrix. Therefore, initially the calculation of the Jacobian is avoided. In the absence of model errors and, when evaluated in the solution, Eq. 5-13 can be approximated by

$$Cov[\hat{\tau}] \cong H^{-1} \cdot \sigma^2, \qquad \text{Eq. 5-14}$$

where H represents the Hessian, i.e. the second derivative of the cost function (Eq. 5-2) with respect to the estimated parameters $\tau_2, .., \tau_K$. The inverse Hessian is readily available as a part of the Newton-Raphson algorithm.

The same sanity check as was mentioned above shows that for the two realizations of the impulse response with σ = 3 mV, the uncertainty on the estimated delays based on the inverse Hessian is found to be 70.9 fs. It lies once within and once just outside the 95% confidence interval of the observed uncertainties (based on a χ^2_{500} - distribution, resp. 61.3 fs .. 69.6 fs and 63.5 fs .. 72.1 fs).

A more detailed inspection of the parameter covariance matrix shows a correlation coefficient between any two delays which approximates 0.5 for all estimated delays. Although initially this might come as a surprise, a correlation coefficient of 0.5 can easily be explained. Indeed, because of the constraint that τ_1 was set to 0, the remaining estimates $\tau_2, .., \tau_K$ in fact correspond to estimates of $\tilde{\tau}_2 - \tilde{\tau}_1, .., \tilde{\tau}_K - \tilde{\tau}_1$ where $\tilde{\tau}_k$, $k = 1, .., K$ are assumed to be identically distributed and uncorrelated. As such, for $k \neq l$

$$\begin{aligned}
\sigma^2_{\tau_k \tau_l} &= E[(\tau_k - E\{\tau_k\})(\tau_l - E\{\tau_l\})] \\
&= E[(\tilde{\tau}_k - \tilde{\tau}_1 - E\{\tilde{\tau}_k - \tilde{\tau}_1\})(\tilde{\tau}_l - \tilde{\tau}_1 - E\{\tilde{\tau}_l - \tilde{\tau}_1\})] \\
&= E[(\tilde{\tau}_1 - E\{\tilde{\tau}_1\})^2] = \alpha
\end{aligned} \qquad \text{Eq. 5-15}$$

while for $k = l$

$$\begin{aligned}\sigma^2_{\tau_k\tau_k} &= E[(\tau_k - E\{\tau_k\})^2] \\ &= E[(\tilde{\tau}_k - \tilde{\tau}_1 - E\{\tilde{\tau}_k - \tilde{\tau}_1\})^2] \\ &= E[(\tilde{\tau}_1 - E\{\tilde{\tau}_1\})^2] + E[(\tilde{\tau}_k - E\{\tilde{\tau}_k\})^2] = 2\alpha\end{aligned}$$

Eq. 5-16

The correlation coefficient therefore equals to $\rho = \frac{\sigma^2_{\tau_k\tau_l}}{\sigma^2_{\tau_k\tau_k}} = \frac{\alpha}{2\alpha} = \frac{1}{2}$.

Analysis of the noise sources: jitter noise

The next logical step is to study the influence of jitter noise on the measurements in the frequency domain. The additive noise is set to zero.

The jitter noise $n_t(t_i)$ is assumed to be zero mean, normally distributed, independent and stationary with respect to t. Its standard deviation is assumed to be known and is set equal to 1 ps.

First, the study is performed based on simulations using the sample variance of the real and the imaginary part of the spectral data corresponding to the realizations of the same known signal[1] as before that is disturbed by jitter noise only.

Next, the covariance matrix in the frequency domain is calculated based on the covariance matrix in the time domain. For the considered jitter noise, the latter reduces to a diagonal matrix. The values on the diagonal correspond to the variance and their value depends on the derivatives of the underlying signal with respect to t.

In order to minimize the variation of the phase spectrum of the considered impulse response as a function of the frequency, the analytical expression of the impulse response is adapted such that its peak value is located at $t = 0$. This way, the deterministic portion of the variation of the real and the imaginary part of the spectrum, when taken as a function of the frequency, is minimized.

Simulation results

Figure 5-1 shows the sample mean of 5000 realizations of the considered impulse response. Figure 5-2 zooms in on the left edge of the time window and clearly shows that the peak value of the impulse response is located at $t = 0$.

Figure 5-3 and figure 5-4 show the corresponding sample variance of the signals shown in figure 5-1 and figure 5-2.

Although it has no physical meaning, it is useful to take a look at the discrete Fourier transform of the sample variance. Later, it will become apparent that there is a relationship between this Fourier transform and the sample variance of the real and the imaginary part of the spectrum of the 5000 realizations. Figure 5-5 and figure 5-6 respectively show the real and the imaginary part of the Fourier transform of the

1. Again, the same analytical expression of the impulse response is used as during the system identification approach applied to jitter estimation.

Figure 5-1. Sample average of the impulse response disturbed by jitter noise.

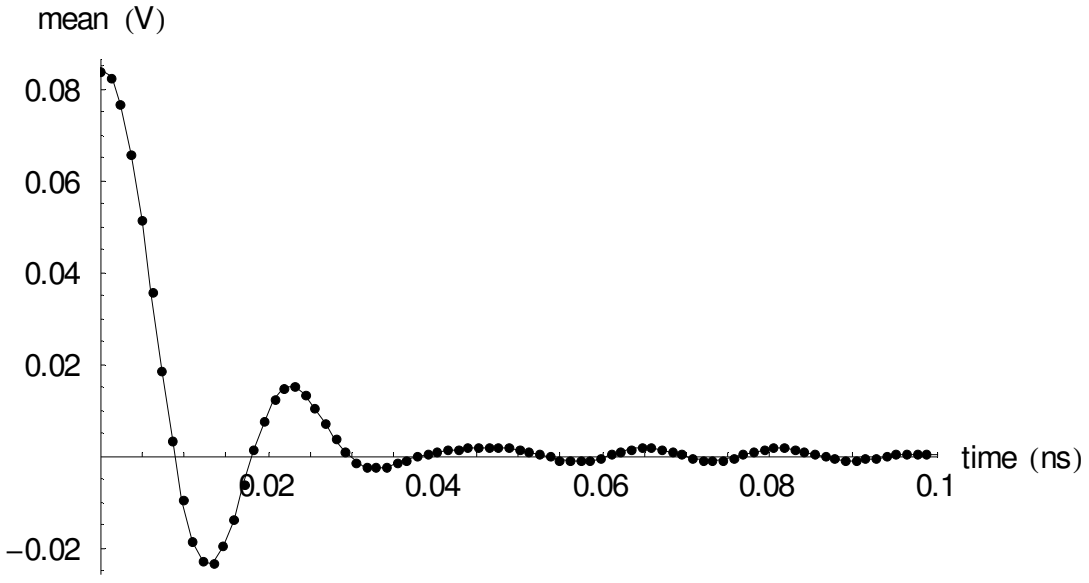

Figure 5-2. Sample average of the impulse response disturbed by jitter noise (zoomed in around $t = 0$).

sample variance (figure 5-3).

Figure 5-7 and figure 5-8 show the sample variance of the real and the imaginary part of the spectrum for the 5000 realizations of the impulse response.

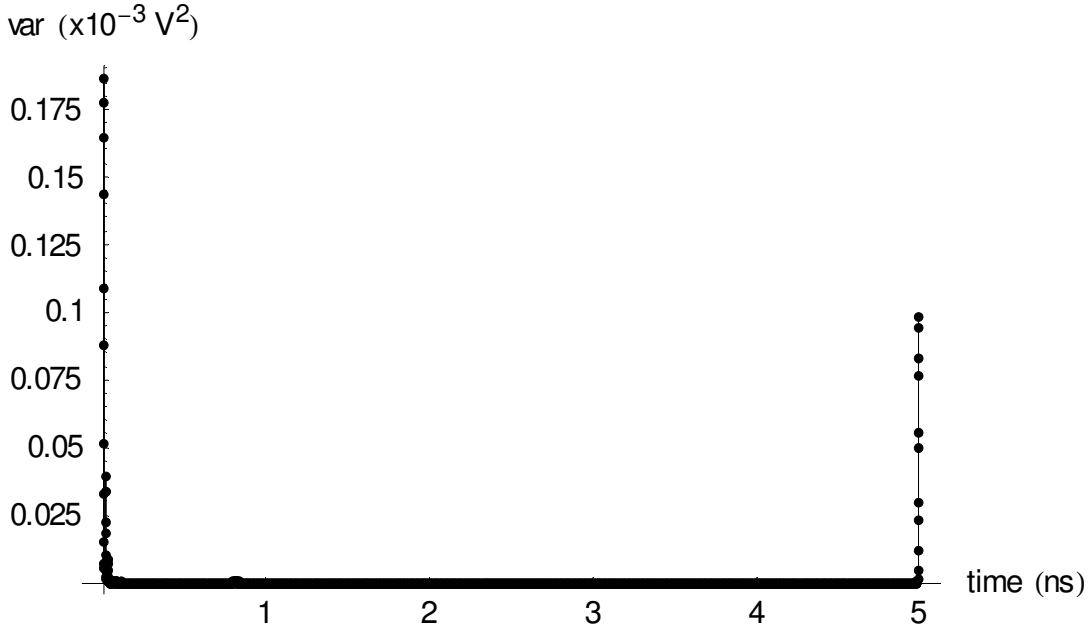

Figure 5-3. Sample variance of the impulse response disturbed by jitter noise.

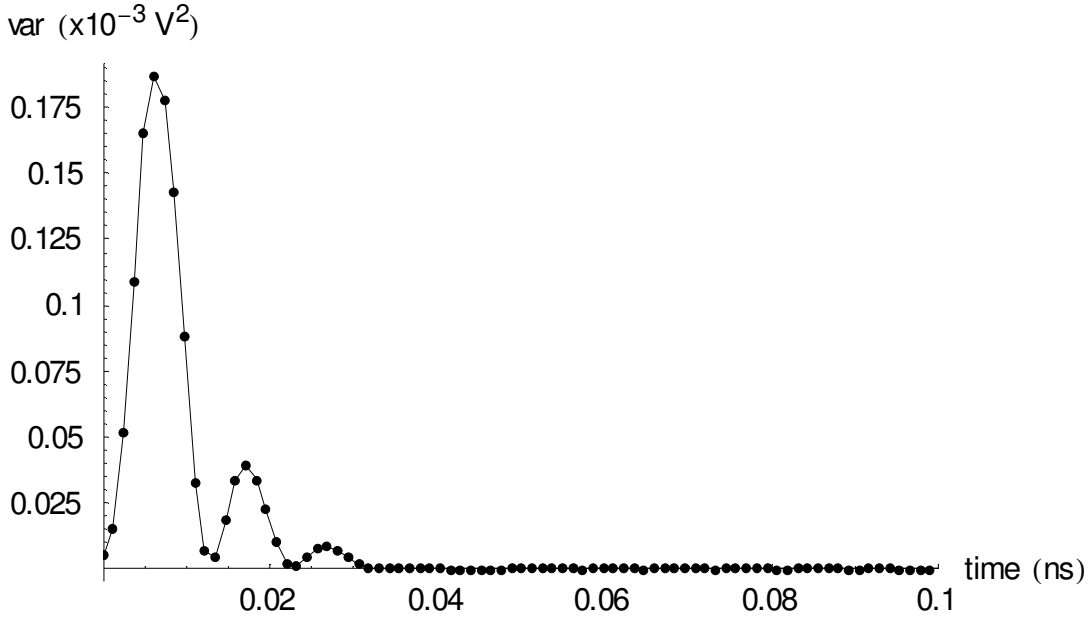

Figure 5-4. Sample variance of the impulse response disturbed by jitter noise (zoomed in around $t = 0$).

Visual comparison of figure 5-5 and figure 5-7 suggests that there is a relationship between the Fourier transform of the sample variance of the signal and the sample variance of the Fourier transform of the signal, especially when taking the folding of the frequency axis by a factor of 2 into account. This is explained in more detail on page 5-20.

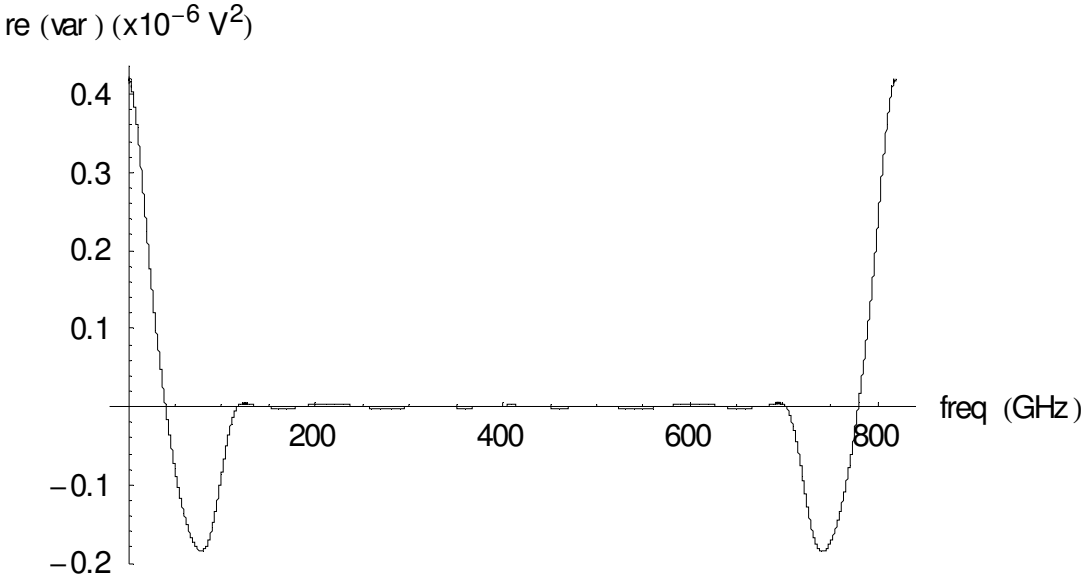

Figure 5-5. Real part of the Fourier transform of the sample variance (figure 5-3).

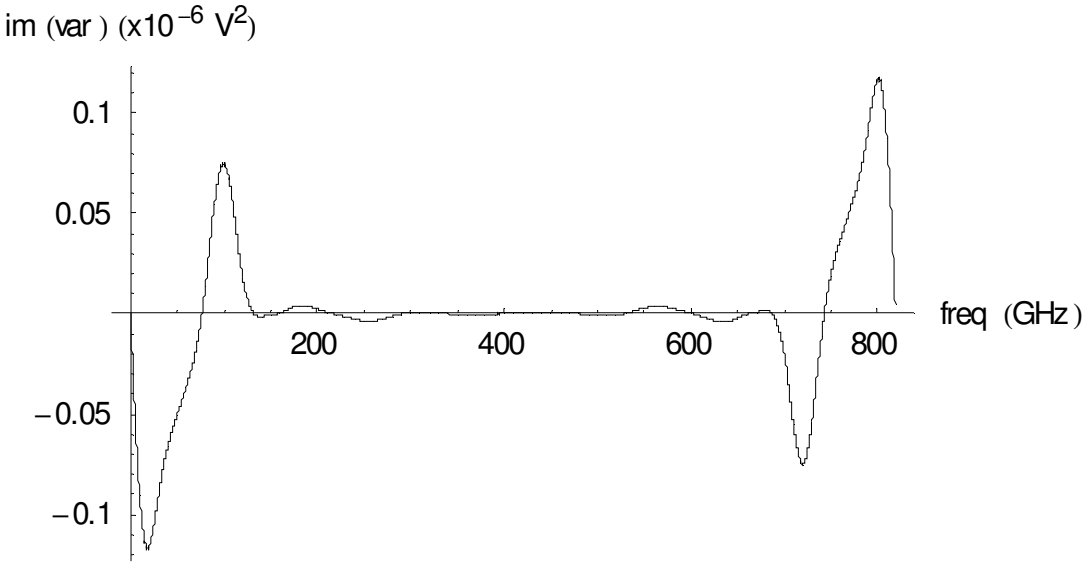

Figure 5-6. Imaginary part of the Fourier transform of the sample variance (figure 5-3).

Calculating the sample variance $\sigma^2_{x,K}$ of K realizations $x(k)$ of a complex number x as $\sigma^2_{x,K} = \frac{1}{K-1}\sum_{k=1}^{K}\left|x(k) - \frac{1}{K}\sum_{l=1}^{K} x(l)\right|^2$, it is clear that $\sigma^2_{x,K}$ equals the sum of the sample variance of the real and the imaginary part of that number.

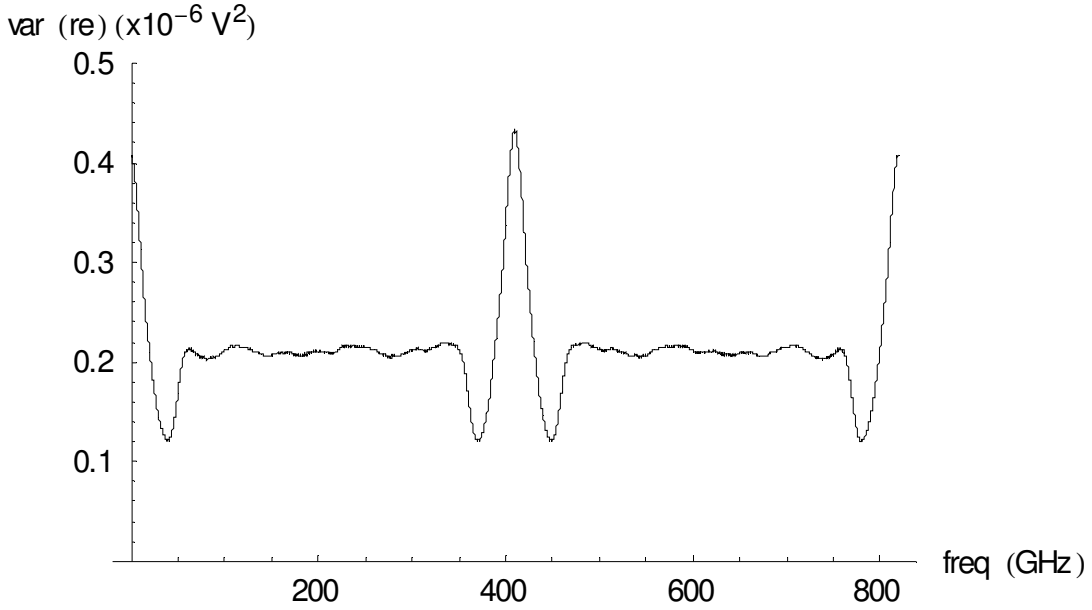

Figure 5-7. Sample variance of the real part of the Fourier transform of the 5000 realizations of the impulse response.

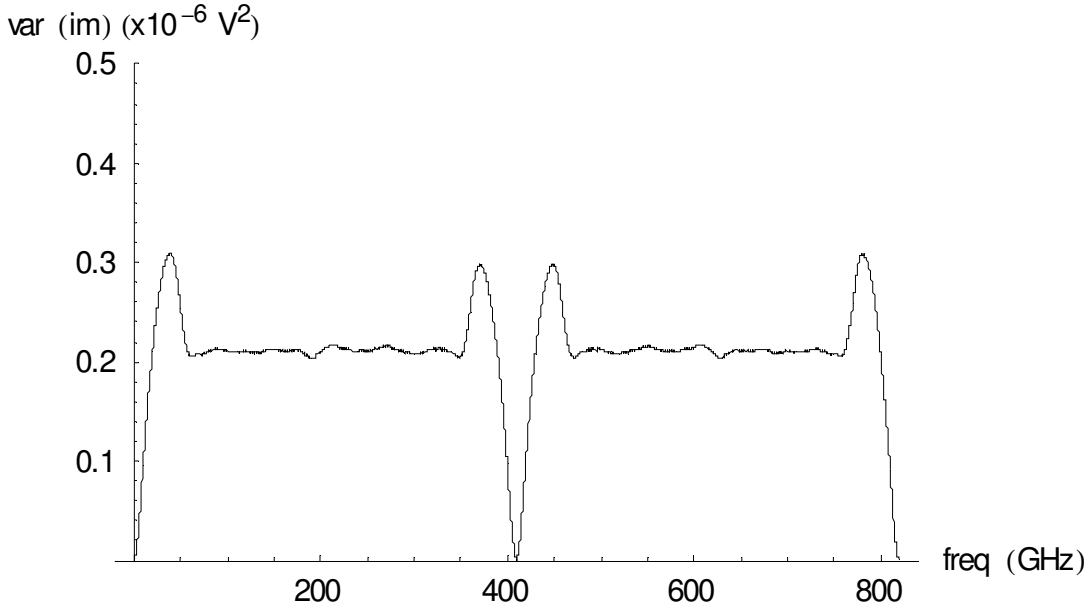

Figure 5-8. Sample variance of the imaginary part of the Fourier transform of the 5000 realizations of the impulse response.

Figure 5-9 shows this sample variance for the spectrum of the 5000 realizations of the impulse response. Taking into account the width of the 95% confidence interval, it is acceptable to consider that this value is a constant as a function of the frequency. Using the 95% confidence interval of a χ^2 distribution, this was verified for the

frequency band where there is no apparent correlation between the noise on the real and the imaginary part. Correct calculation of the 95% confidence interval in case a correlation is present, requires the use of the Wishart distribution [4] and is not dealt with at this moment.

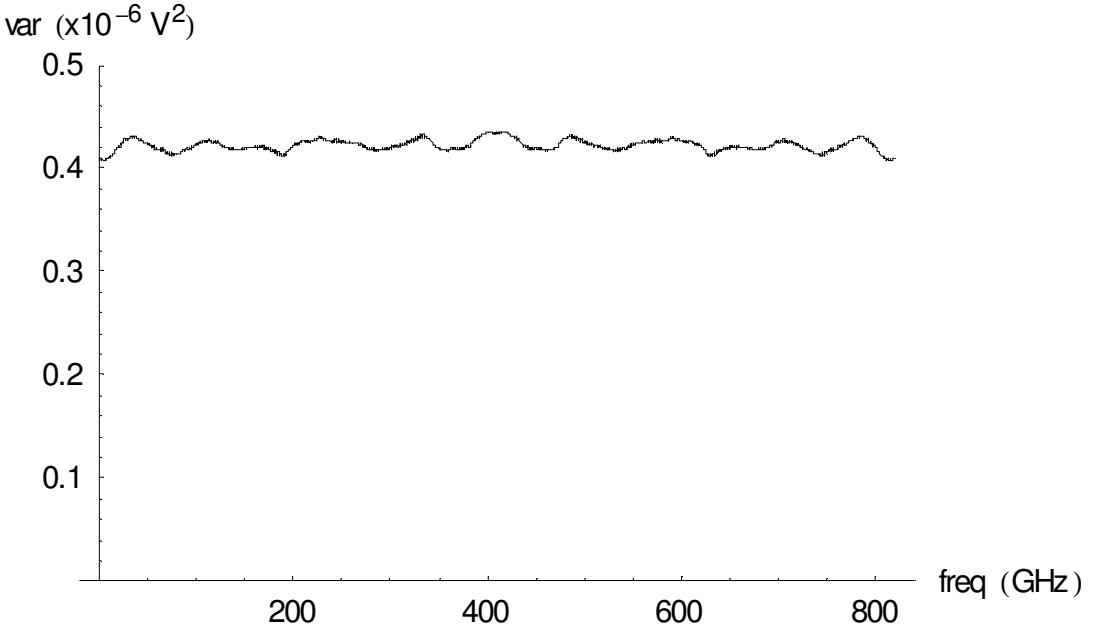

Figure 5-9. Sample variance of the Fourier transform of the 5000 realizations of the impulse response.

Another important verification is to check whether the correlation between the real and the imaginary part of the spectrum is not significant and the result is shown in figure 5-10. The correlation coefficient ρ of two stochastical quantities x and y is defined as

$$\rho = \frac{\sigma_{xy}^2}{\sigma_x \cdot \sigma_y}. \qquad \text{Eq. 5-17}$$

Figure 5-11 zooms in to the lower frequency band and shows the correlation between the noise on the real and the imaginary part of the spectrum up to about 65 GHz. Correct calculation of the 95% confidence interval requires additional work and is not provided[1].

Next it makes sense to look at the 5000 realizations of the real and the imaginary value of the spectrum at some frequencies of interest:

- at 4 GHz: the correlation has its largest negative value

1. However, visual comparison of figure 5-11 and figure 5-18 shows good correspondence, the latter being calculated starting from the variance in the time domain.

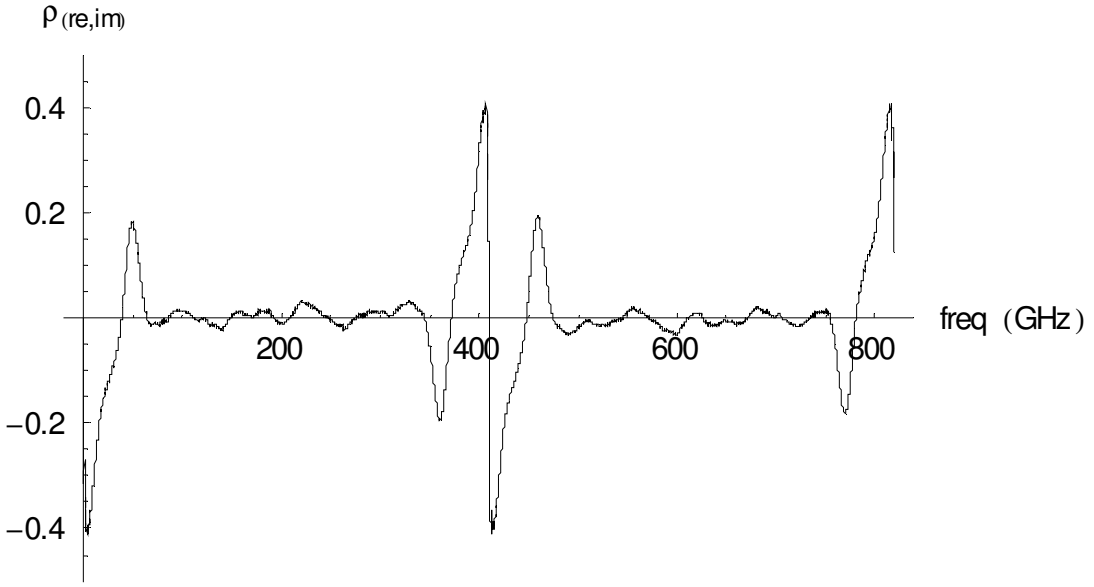

Figure 5-10. Correlation coefficient of the real and the imaginary part of the Fourier transform of the 5000 realizations of the impulse response.

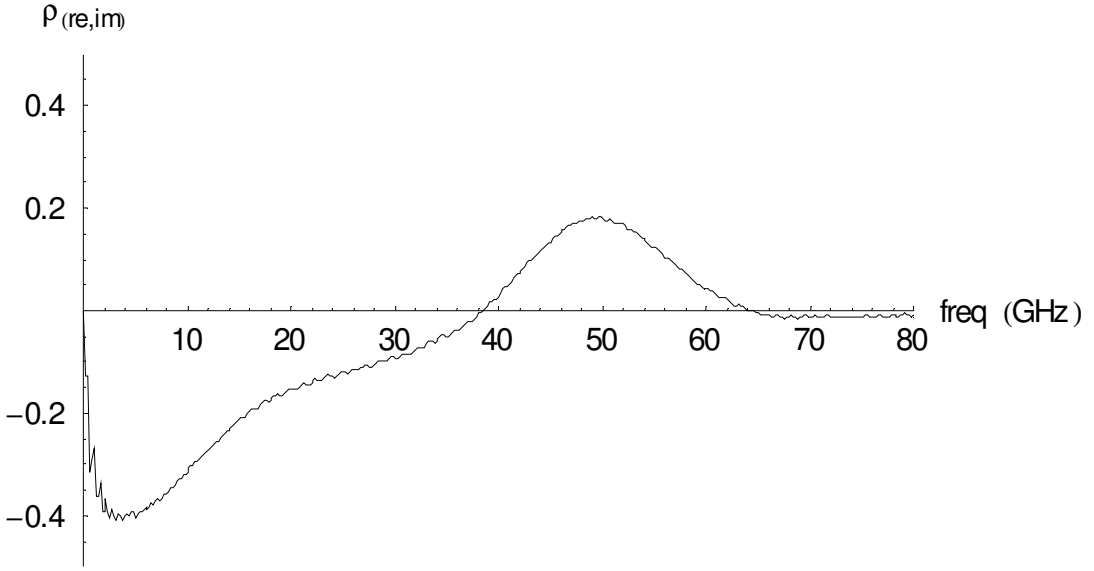

Figure 5-11. Correlation coefficient of the real and the imaginary part of the Fourier transform of the 5000 realizations of the impulse response (DC up to 80 GHz).

- at 38 GHz: the correlation changes from a negative to a positive value
- at 50 GHz: the correlation has its largest positive value
- at 70 GHz: no correlation is present between the real and the imaginary part

Figure 5-12. 5000 realizations of the real and the imaginary value of the spectrum at 4 GHz (maximum negative correlation $\cong -0.4$).

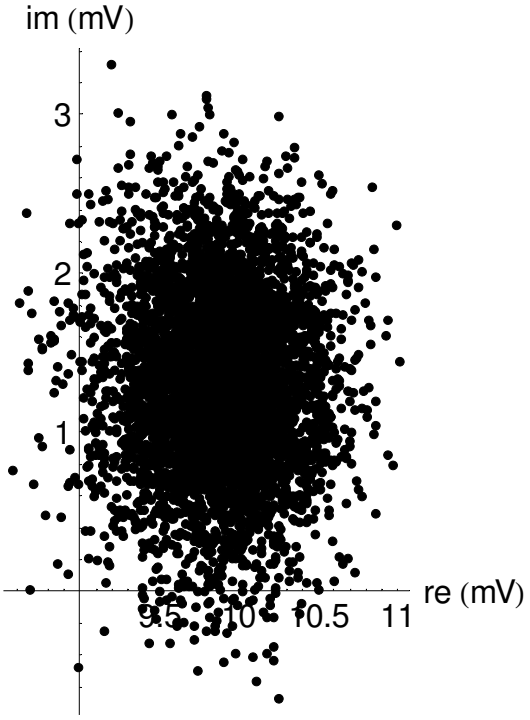

Figure 5-13. 5000 realizations of the real and the imaginary value of the spectrum at 38 GHz (zero correlation, but clearly not circular).

Figure 5-12 up to figure 5-15 show the results using a 1:1 aspect ratio for all plots.

Clearly the assumption that the spectral noise is circular complex is not valid, especially at lower frequencies. The bandwidth where the correlation is significant depends on the measured signal itself.

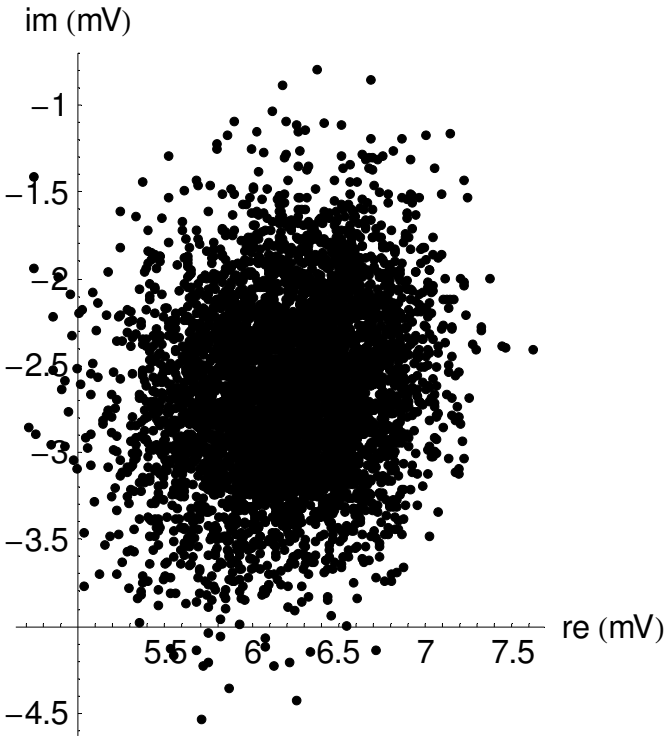

Figure 5-14. 5000 realizations of the real and the imaginary value of the spectrum at 50 GHz (maximum positive correlation $\cong 0.2$).

In the presence of jitter noise, the covariance matrix in the frequency domain must take into account the contributions of the real and the imaginary part of the spectrum separately.

Furthermore, one can take a look at the correlation between spectral contributions originating from different frequencies (real, real), (real, imag), (imag, real) and (imag, imag). Correlation is observed up to about 130 GHz.

Calculation of the covariance matrix of the spectral noise in the frequency domain

It is shown that it is possible to calculate the covariance matrix in the frequency domain, if one has an estimate of the variance in the time domain. The time domain variance has been obtained in the chapter on jitter estimation. There, it is demonstrated that it is indeed possible to estimate the standard deviation of both the additive and the jitter noise, hence also the variance as a function of time. This information can be used to come up with an estimate of the covariance matrix in the frequency domain. This allows to obtain an optimal scaling of each frequency contribution of the cost function in the sense that it becomes possible to know the expected value of the cost in advance and to minimize and estimate the uncertainty on the estimated delays.

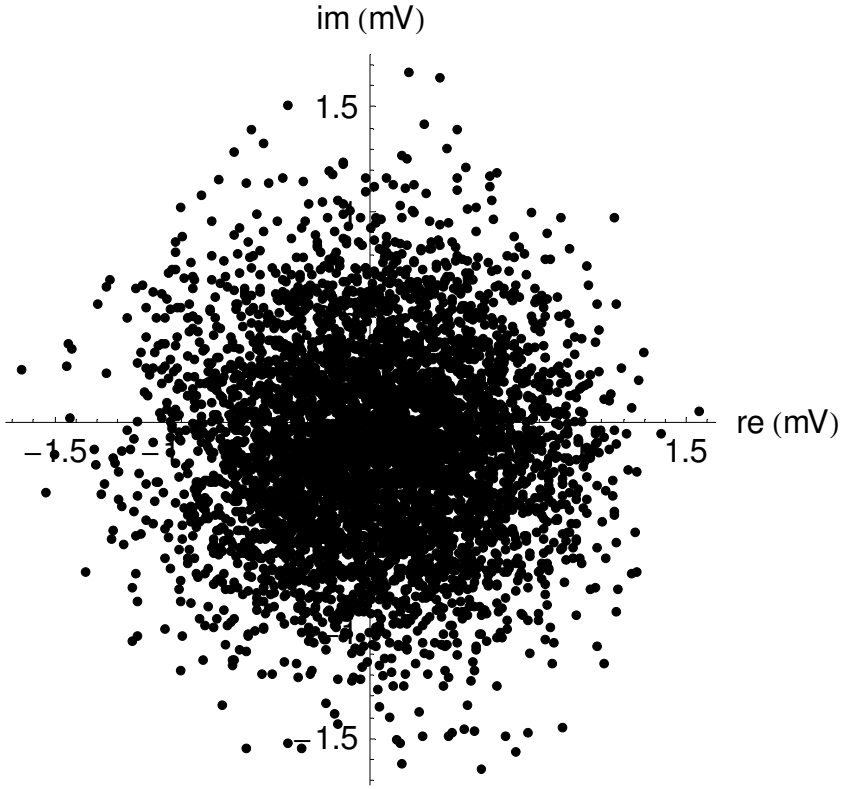

Figure 5-15. 5000 realizations of the real and the imaginary value of the spectrum at 70 GHz (no correlation, circular complex noise).

Because the above simulations show that the noise is not circular complex, the frequency covariance matrix must take the real and the imaginary part of the spectrum into account separately. The covariance matrix in the frequency domain can then be calculated similar to Eq. 5-11, provided that the real and the imaginary part are separated.

$$Cov[Y(\omega)] = J \cdot Cov[y(t)] \cdot J^T \qquad \text{Eq. 5-18}$$

In Eq. 5-18, J^T represents the transpose of the Jacobian matrix $J \in |\mathrm{R}^{2M \times N}$,which is the matrix notation of the discrete Fourier transform where the odd rows correspond to the real part and the even rows to the imaginary part of the Fourier coefficient. M represents the number of frequency components of interest and N is the number of time points of one realization of the impulse response. The covariance matrix in the time domain $Cov[y(t)] \in |\mathrm{R}^{N \times N}$ reduces to a diagonal matrix in case the jitter is not correlated as a function of time. The values on this diagonal correspond to the variance as function of time. Based on the analytical expression of the impulse response and the standard deviation of the jitter, it is possible to calculate the required derivatives of the signal with respect to time and hence to obtain the variance as a function of time.

$$\tilde{\sigma}^2_{y_3}(t_i) = \left(\frac{dy_0}{dt}\right)^2\Bigg|_{t=t_i} \cdot \sigma^2_{n_t} + \left\{\frac{1}{2}\cdot\left(\frac{d^2 y_0}{dt^2}\right)^2 + \frac{dy_0}{dt}\cdot\frac{d^3 y_0}{dt^3}\right\}\Bigg|_{t=t_i} \cdot \sigma^4_{n_t} + \frac{5}{12}\cdot\left(\frac{d^3 y_0}{dt^3}\right)^2\Bigg|_{t=t_i} \cdot \sigma^6_{n_t}$$

Eq. 5-19

In the chapter on jitter estimation, for the considered impulse response, Eq. 5-19 was found to provide a very good approximation for the variance in the case of Gaussian jitter that has a standard deviation less than or equal to 1 ps.

The rows $2k, 2k+1$ of the Jacobian matrix J contain the coefficients of the discrete Fourier transform that yield the real and the imaginary part of the k^{th} frequency component. In the case $Cov[y(t)]$ is a diagonal matrix with $\tilde{\sigma}^2_{y_3}(t_i)$ on the diagonal, the sum of each row of $J \cdot Cov[y(t)]$ corresponds to the real, resp. the imaginary part of the Fourier transform of the variance as function of time.

$$\tilde{\Sigma}^2_{Y_3}(m) = \frac{1}{\sqrt{N}}\sum_{i=0}^{N-1} \tilde{\sigma}^2_{y_3}(i)\cdot e^{-j2\pi mi/N},\ m = 0\ ..\ N-1 \qquad \text{Eq. 5-20}$$

For the diagonal terms of $Cov[Y(\omega)]$, the matrix product $J \cdot Cov[y(t)] \cdot J^T$ yields $\frac{1}{N}\sum_{i=0}^{N-1} \tilde{\sigma}^2_{y_3}(i)\cdot\cos^2(2\pi ki/N)$ for the even rows, $\frac{1}{N}\sum_{i=0}^{N-1} \tilde{\sigma}^2_{y_3}(i)\cdot\sin^2(2\pi ki/N)$ for the odd rows.

This is not equal to the real and the imaginary part of $\frac{1}{N}\sum_{i=0}^{N-1} \tilde{\sigma}^2_{y_3}(i)\cdot e^{-j4\pi ki/N}$. The latter corresponds to a scaled version of Eq. 5-20, except for the fact that the complex exponential $e^{-j2\pi kn/N}$ is now replaced by $e^{-j2\pi kn/N'}$ with $N' = \frac{N}{2}$. It looks as if N is divided by a factor of 2 and therefore the frequency axis appears to be folded by

a factor of 2. This explains both the apparent similarity and the difference (see figure 5-5 up to figure 5-8) between the sample variance of the spectral data and the spectral data corresponding to the sample variance.

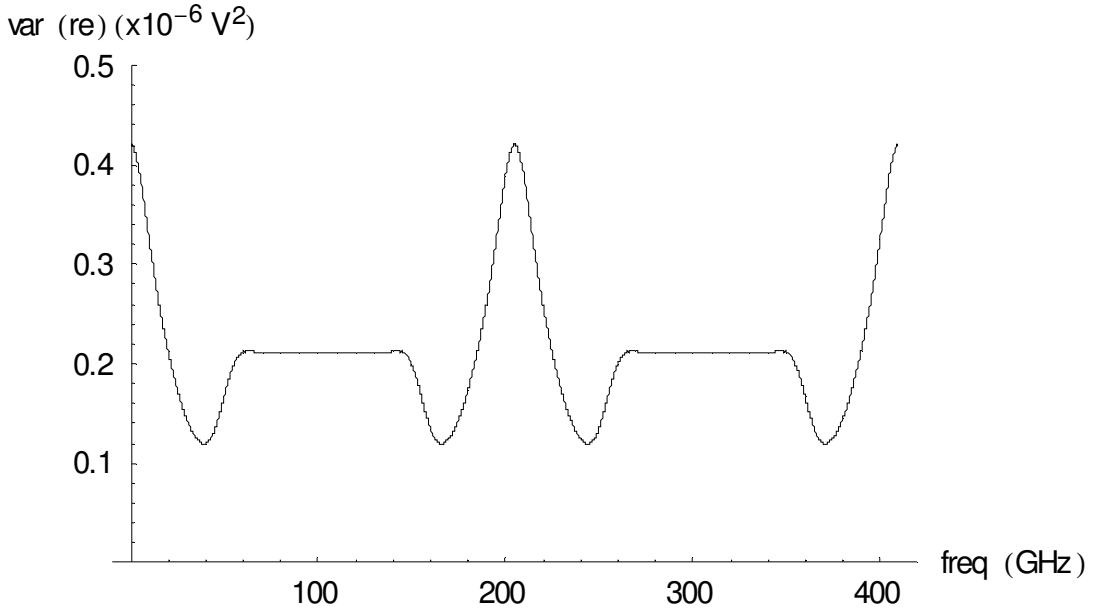

Figure 5-16. Variance as a function of the frequency using Eq. 5-18 and Eq. 5-19 (even rows).

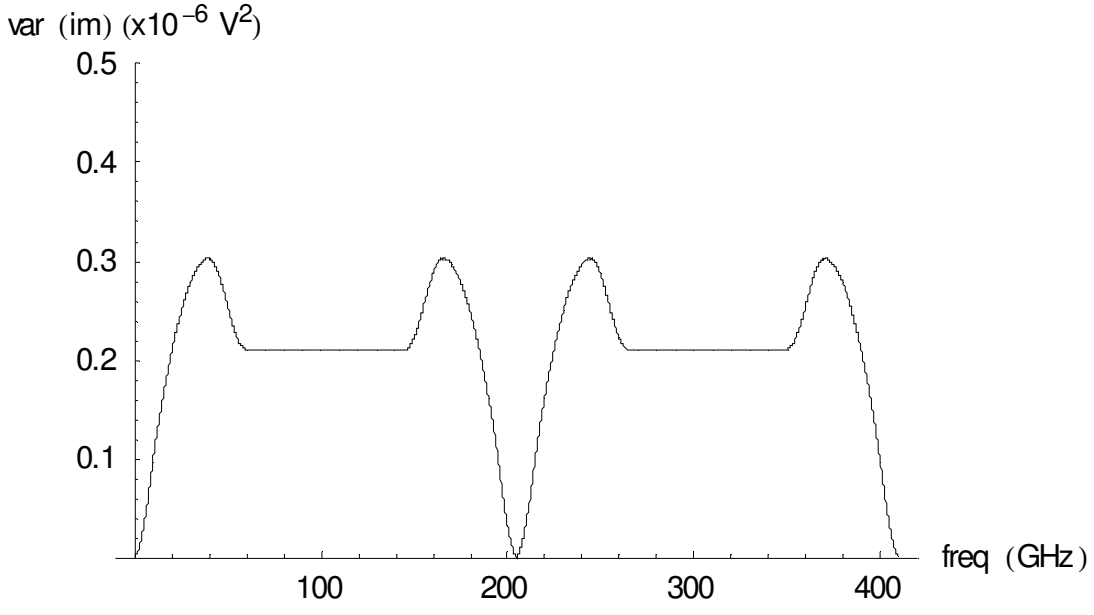

Figure 5-17. Variance as a function of the frequency using Eq. 5-18 and Eq. 5-19 (odd rows).

Figure 5-16 and figure 5-17 show the variance as function of frequency, based on Eq. 5-18 and Eq. 5-19 for the even and odd rows of the diagonal of $J \cdot Cov[y(t)] \cdot J^T$. The sum of both turns out to be 0.422 10^{-6} V^2, is constant as a function of the frequency and corresponds to the mean value of the sample variance shown in figure 5-9. When comparing figure 5-16 and figure 5-17 to figure 5-7 and figure 5-8, it should be noted that, due to memory limitations, the number of time points was reduced from 4096 to 2048. This explains the decrease of the sampling frequency with respect to figure 5-7 and figure 5-8.

Figure 5-18 is the equivalent of figure 5-11, but is now calculated based on the covariance matrix, obtained using Eq. 5-18.

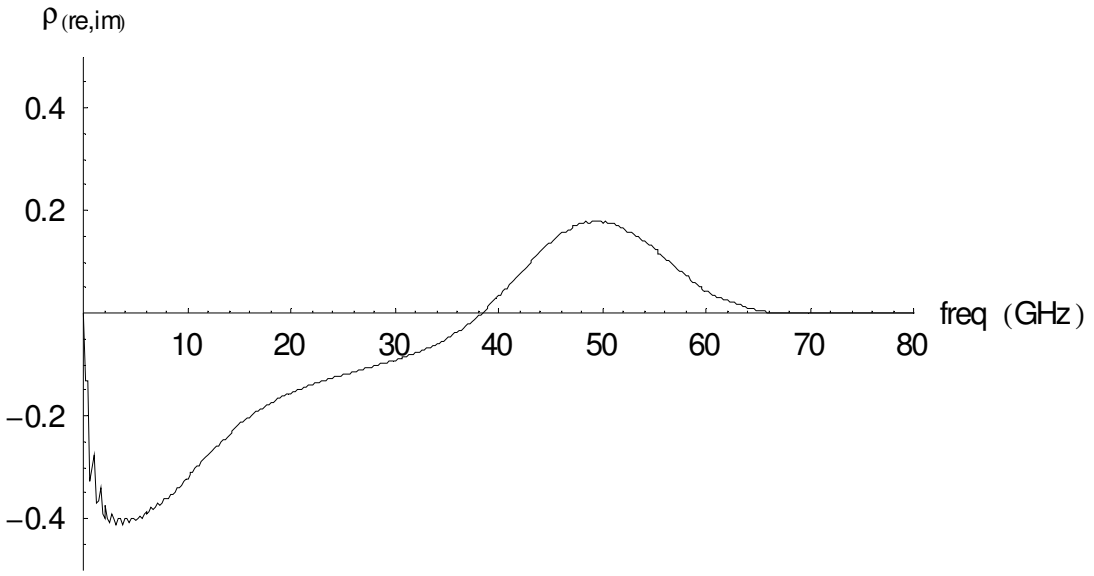

Figure 5-18. Correlation coefficient of the real and the imaginary part using Eq. 5-18 and Eq. 5-19.

Finally, the artwork resulting from the contour plots of each portion ((re,re), (re,im), (im,re) and (im,im)) of the frequency covariance matrix is shown in figure 5-19 up to figure 5-22. White is mapped onto the largest value and black to the smallest value.

Covariance	min. value (V^2)	max. value (V^2)
(re,re)	-0.183 μ	0.422 μ
(im,im)	-0.303 μ	0.303 μ
(re,im) and (im,re)	-0.117 μ	0.117 μ

Table 5-1. Min. and max. values of the different portions of the frequency covariance matrix.

It can be concluded that an estimate of the variance in the time domain (Eq. 5-19) allows to construct the full covariance matrix in the frequency domain.

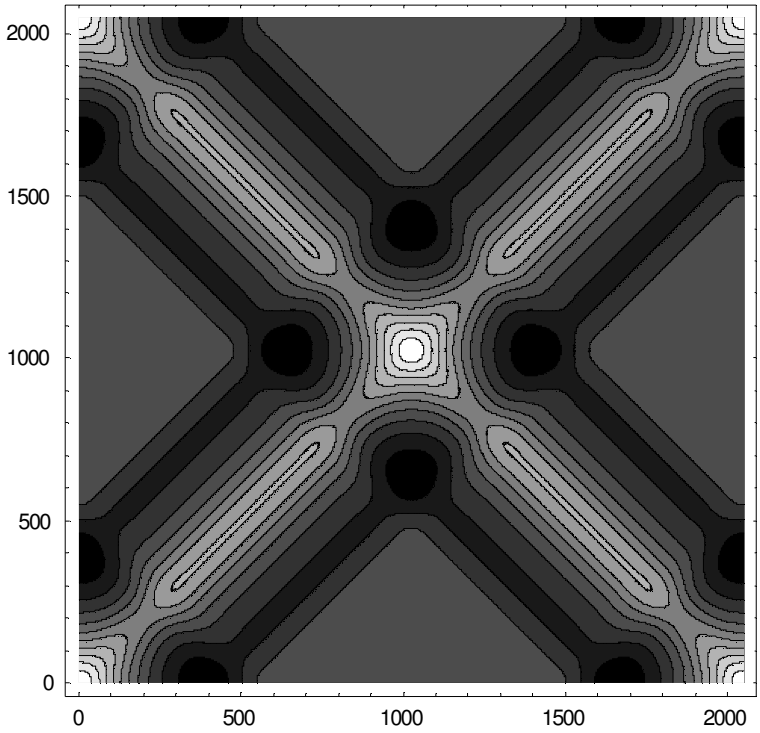

Figure 5-19. Contour plot of the (re, re) portion of the frequency covariance matrix (both axes correspond to the DFT index (N = 2048)).

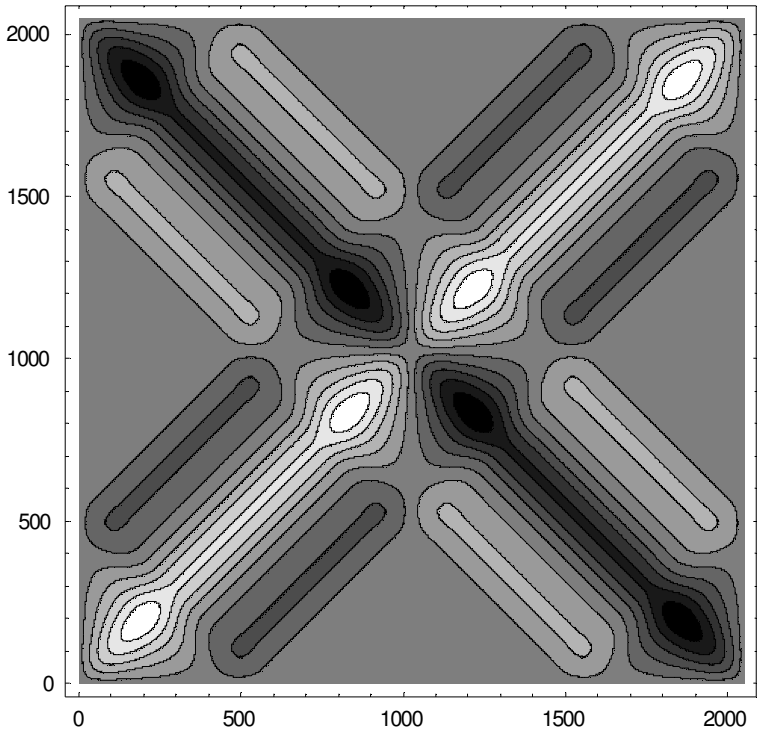

Figure 5-20. Contour plot of the (im, im) portion of the frequency covariance matrix (both axes correspond to the DFT index (N = 2048)).

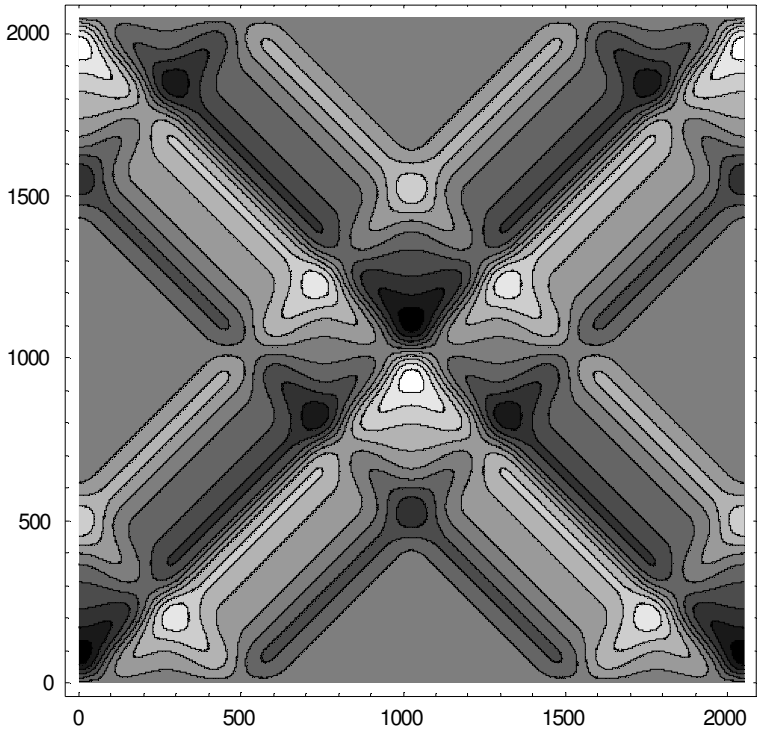

Figure 5-21. Contour plot of the (re, im) portion of the frequency covariance matrix (both axes correspond to the DFT index (N = 2048)).

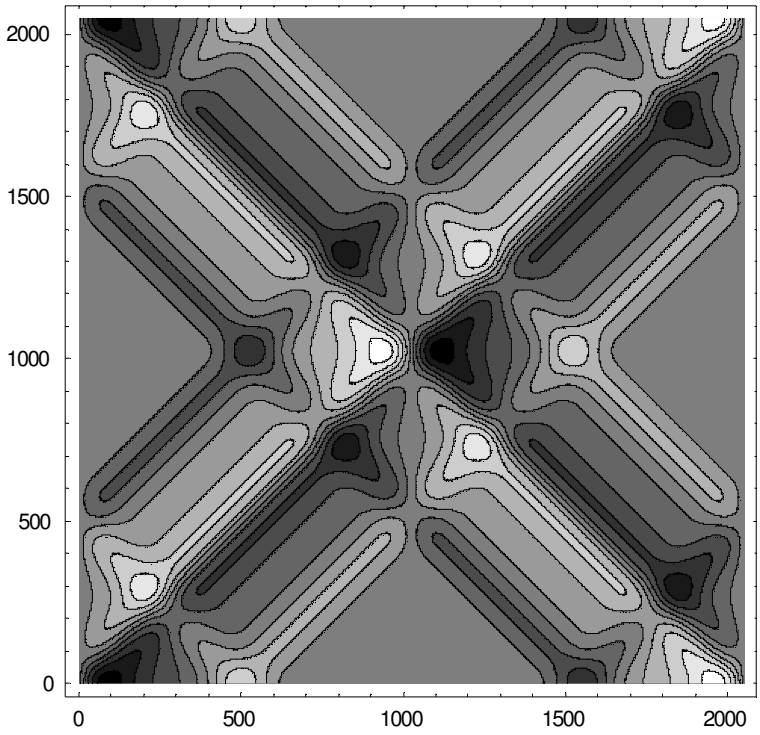

Figure 5-22. Contour plot of the (im, re) portion of the frequency covariance matrix (both axes correspond to the DFT index (N = 2048)).

The added value of the covariance matrix for the WLS

The knowledge of the covariance matrix and a clever implementation of the Jacobian-based Gauss-Newton method allow to obtain several additional improvements.

LS parameter covariance matrix

First, it is possible to estimate the uncertainty of the estimated delays in the case of a least squares estimator. Taking into account the fact that the noise on the real and the imaginary part of the spectral components must be treated separately, Eq. 5-6 can be rewritten as

$$V_{LS} = \sum_{k,\,m} \left\{ [re(e_{k,\,m})]^2 + [im(e_{k,\,m})]^2 \right\}, \text{ with} \qquad \text{Eq. 5-21}$$

$$e_{k,\,m} = Y_k(\omega_m) \cdot e^{j\omega_m \tau_k} - \frac{1}{K} \sum_{l\,=\,1}^{K} Y_l(\omega_m) \cdot e^{j\omega_m \tau_l} \qquad \text{Eq. 5-22}$$

Let $e = \left[e_1 \;..\; e_K \right]^T$, $e_k = \left[re(e_{k,\,1}) \;\; im(e_{k,\,1}) \;\; .. \;\; re(e_{k,\,M}) \;\; im(e_{k,\,M}) \right]$.

In matrix notation, Eq. 5-21 becomes $V_{LS} = e^T(\tau) \cdot e(\tau)$, where $e(\tau) \in |\mathrm{R}^{K \cdot 2M \times 1}$. Here, K corresponds to the number of realizations of the unknown signal and M corresponds to the number of frequency components taken into account to estimate the K delays, that are grouped in the vector $\tau = \left[\tau_1 \;..\; \tau_K \right]^T$.

An estimate of the uncertainty on the LS estimates of the delays is given by

$$Cov(\hat{\tau}) \cong (J^T J)^{-1} \; J^T Cov(Y(\omega)) \; J \; (J^T J)^{-1} \qquad \text{Eq. 5-23}$$

where J represents the Jacobian and equals $\frac{\partial e(\tau)}{\partial \tau}$. Based on the degeneracy mentioned earlier (Eq. 5-7), τ_1 is again fixed equal to zero and the Jacobian matrix looses one column: $J \in |\mathrm{R}^{K \cdot 2M \times (K-1)}$, while $Cov(Y(\omega)) \in |\mathrm{R}^{2M \times 2M}$.

Typical values for both K and M in an experimental environment are 500. This results in a large Jacobian matrix of size 500,000 x 499 and a covariance matrix of 1000 x 1000. Fortunately, there is some structure in the Jacobian matrix derived from Eq. 5-22, such that the problem can be solved in blocks of size $2M \times (K-1)$.

WLS estimator

Secondly, it is possible to construct a weighted least squares estimator which minimizes the uncertainty on the estimated delays and, at the same time, has a known expected value of the cost function, such that the obtained cost function can be compared to its expected value. This criterion can then be used to detect model errors.

$$V_{WLS} = e^T(\tau) \cdot [Cov(Y(\omega))]^{-1} \cdot e(\tau) \qquad \text{Eq. 5-24}$$

The advantages of this method will become apparent when it will be applied to both simulations and measurements.

Simulations

A first set of simulations is based on the same analytical expression of the impulse response as the one that is used during the simulations of the jitter estimation. The drift is estimated in the presence of both additive and jitter noise. The jitter noise and the additive noise are set to have realistic values: both are Gaussian, zero mean with a standard deviation of respectively 1 ps and 0.6 mV. The exact signal is assumed to be known. Therefore, using Eq. 5-18 and Eq. 5-19, one can calculate $Cov(Y(\omega))$.

Each simulation consists of 500 realizations of the signal. Hence, $K = 500$. Each realization contains 4096 time samples of the impulse response, which spans an acquisition window of 5 ns. If the spectral components are taken into account up to 100 GHz, then $M = 500$.

Estimators

Different estimators are compared:

- based on Eq. 5-2, using the first realization as a reference instead of the aligned sample average. Inspired by the terminology used in [1], this estimator will be referred to as the “naive” LS estimator.
- the “enhanced” LS estimator, corresponding to Eq. 5-6, uses the aligned sample average as reference.
- the “enhanced” WLS estimator, based on Eq. 5-24, also uses the aligned sample average as reference and adds weighting based on the inverse full covariance matrix. This takes the uncertainty on the real and the imaginary part into account separately.

Zero drift

During the first simulation, the drift is set to be exactly zero. Due to the jitter noise and the additive noise, the estimated drift is not exactly zero, however, but rather becomes a stochastical variable. Because of the degeneracy, demonstrated by Eq. 5-7, only its standard deviation is shown in Table 5-2.

Estimator	**Uncertainty σ_τ of estimated τ (ps)**	**95% confidence interval (ps) of σ_τ**
naive LS	0.260	0.245 .. 0.277
enhanced LS	0.253	0.238 .. 0.270
enhanced WLS	0.132	0.124 .. 0.141

Table 5-2. Uncertainty of the estimated drift (case where the exact drift is zero).

Table 5-2 clearly shows that the naive and the enhanced LS estimator obtain the same performance. Their uncertainties lie inside their respective confidence intervals. The enhanced WLS estimator reduces the uncertainty on the estimated delay by about a factor 2.

If, in the presence of jitter, the noise in the frequency domain is assumed to be circular complex noise, the uncertainty on the estimated delay for the enhanced LS estimator based on the parameter covariance matrix (Eq. 5-14) equals 0.021 ps. Hence, it underestimates the obtained uncertainty by a factor of more than 10. Using the full covariance matrix (Eq. 5-23), the estimated uncertainty on the delay turns out to be 0.242 ps and this value falls within the 95% confidence interval of the obtained uncertainty.

In the case of the enhanced WLS estimator, the expected value of the cost is 499500 ± 1999. The realized cost turns out to be 497250 and falls within the 95% confidence interval of the expected value of the cost. It can thus be concluded that there are no detectable model errors[1].

Linear drift

During the second simulation, the drift is known to be a linear function of the realization index. Applying a delay of 0.01 ps per realization, the drift of the first realization is zero, while that of the 500th realization is known to be 4.99 ps.

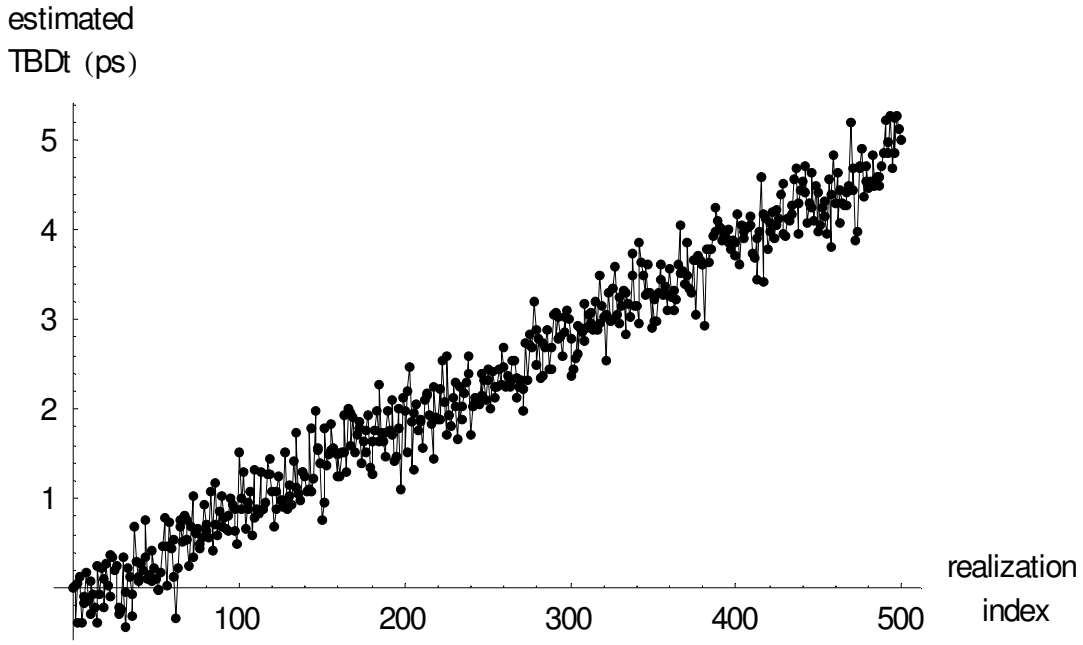

Figure 5-23. Using the enhanced LS estimator (case of linear drift).

1. relative to the variance of the measurements.

Figure 5-23 shows the results for the enhanced LS estimator, which turns out to be visibly indistinguishable from that of the naive LS estimator. Figure 5-24 clearly shows the reduced uncertainty for the enhanced WLS estimator.

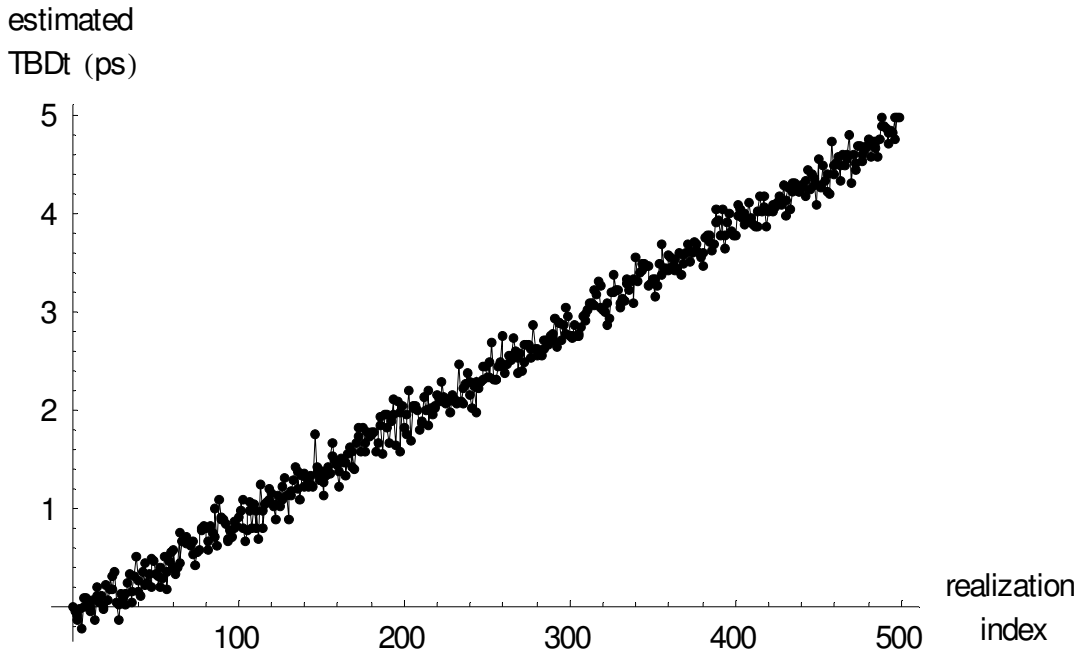

Figure 5-24. Using the enhanced WLS estimator (case of linear drift).

The residuals for each estimator are found by fitting a best linear approximation with slope 0.01 ps per realization through the estimated delay and subtracting the corresponding value from the estimated delay for each realization.

Estimator	**Uncertainty σ_τ of estimated τ (ps)**	**95% confidence interval (ps) of σ_τ**
naive LS	0.264	0.249 .. 0.281
enhanced LS	0.257	0.242 .. 0.274
enhanced WLS	0.123	0.116 .. 0.131

Table 5-3. Uncertainty on the estimated drift (case of linear drift).

The same conclusions can be drawn as for the zero drift. In this case the realized value of the cost for the enhanced WLS estimator turns out to be 498240 and again falls inside the 95% confidence interval of the expected value of the cost.

The general conclusion is that in realistic situations, corresponding to nose-to-nose and EOS-based measurements, the effect of using the aligned average instead of the first realization as a reference signal is minimal. However, use of proper weighting based on the full covariance matrix of the measurements has significant impact.

Comparison to state-of-the-art methods

Here the performance of the implemented estimators is compared to two of the estimators described in [1]. The comparison is based on the simulated signal that is described in Appendix I of [1]. Figure 5-25 shows this noise-free signal, where both the time and amplitude are given in arbitrary units. In fact, the time scale is expressed in integer multiples of the sampling period Δt.

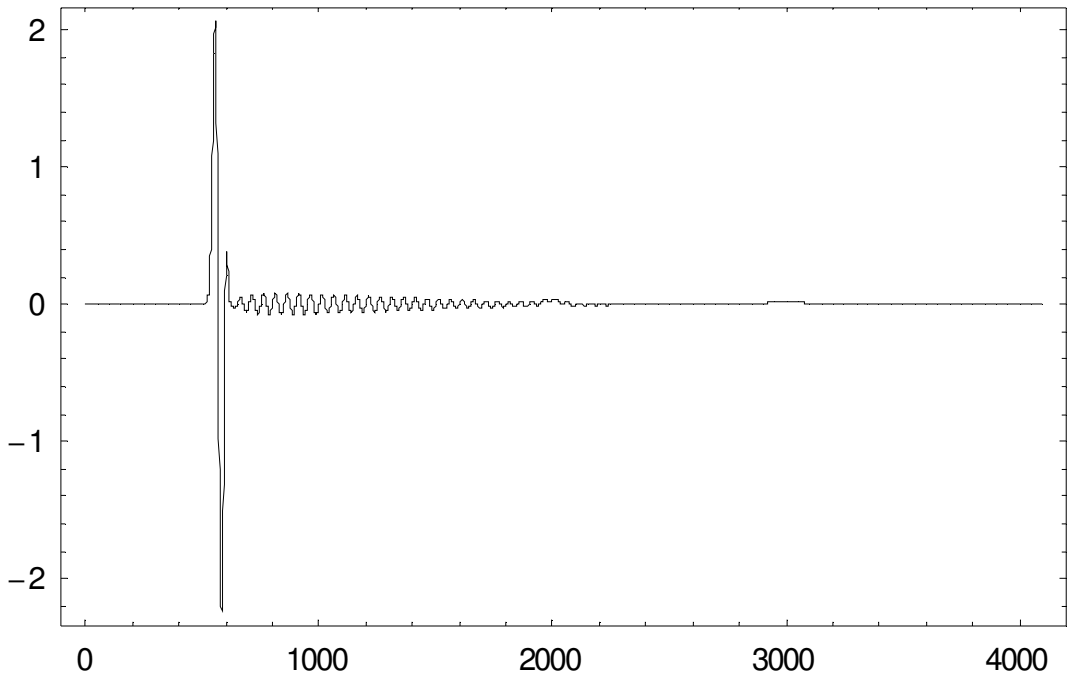

Figure 5-25. Noise-free simulated signal as described in [1] (both axes in arbitrary units).

The paper mainly compares four estimators. Two of them are described below.

The “naive cross-correlation” method can be compared to the naive LS estimator described in this work, because both use the first realization as a reference signal. However, the naive method described in [1] is based on a cross-correlation technique performed in the time domain and as such restricted to a grid corresponding to integer multiples of Δt. In order to overcome this limitation, [1] searches for the global minimum about the grid value which maximizes the cross-correlation based on a golden search and parabolic interpolation (sic).

The “complete cross-correlation” method calculates the relative drift between any combination of the K realizations to come up with an averaged drift of all realizations with respect to the first realization. As such, this method is similar to the enhanced LS estimator described in this chapter.

The comparison is based on the simulations proposed in [1]. The complete cross-correlation method performs best and the naive cross-correlation in general performs

worst for these simulations and are therefore selected as reference methods to compare them to the estimators proposed in this work.

For different values of the standard deviation of the jitter noise σ_{n_t} and the standard deviation of the additive noise σ_{n_y}, that are both expressed in arbitrary units[1], a set of 100 misaligned signals are realized. The standard deviation of the random drift associated with each signal is 2.5. In fact the value of σ_{n_y} should be compared to a peak-to-peak value of about 4.2, while both σ_{n_t} and the random drift are expressed as multiples of Δt.

The simulation results are summarized in figure 4 of [1]. The RMS prediction error is used as performance criterion and is defined in Appendix II of [1]. For selected values of σ_{n_t} and σ_{n_y}, the RMS prediction errors are retrieved as well as possible from figure 4 of [1] and are used in this comparison.

The RMS prediction error as defined[2] in [1] is shown in Eq. 5-25.

$$RMS = \sqrt{\sum_{k=1}^{K} [(\delta_k - \bar{\delta}) - (\hat{d}_k - \hat{\bar{d}})]^2} \qquad \text{Eq. 5-25}$$

Here δ_k represents the true absolute drift and $\hat{d}_k$ the estimated drift. Both values are compared to their respective sample mean $\bar{\delta} = \frac{1}{K}\sum_{k=1}^{K} \delta_k$ and $\hat{\bar{d}} = \frac{1}{K}\sum_{k=1}^{K} \hat{d}_k$.

The limited selection of values for σ_{n_t} and σ_{n_y} is motivated in Table 5-4. One value is taken for each kind of behaviour.

In order to get an idea of what is meant by “moderate” jitter $\sigma_{n_t} = 1$ and “moderate” additive noise $\sigma_{n_y} = 0.1$, figure 5-26 shows the noise-free signal, while figure 5-27

1. This allows easy comparison to [1]. In that paper σ_{jit} represents the standard deviation of the jitter noise and σ_{add} represents the standard deviation of the additive noise.

2. Normally the summation under the root sign should be divided by K.

Situation	σ_{n_t}	σ_{n_y}
no jitter, small additive noise	0	0.02
no jitter, moderate additive noise	0	0.1
significant jitter, moderate additive noise	3	0.1
moderate jitter, small additive noise	1	0.02
moderate jitter, moderate additive noise	1	0.1

Table 5-4. Selected values of σ_{n_t} and σ_{n_y} (in arbitrary units).

shows one realization of such a noisy signal while zooming in to the main portion of the pulse to show the jitter.

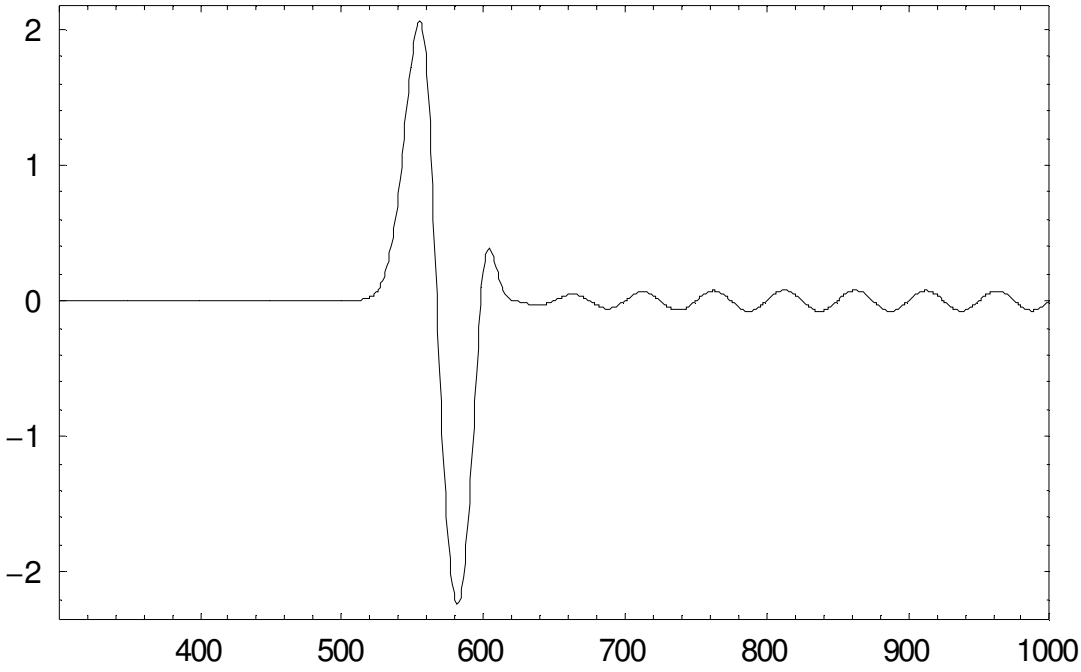

Figure 5-26. Zoomed version of the noise-free simulated signal (both axes are labelled in arbitrary units).

The RMS prediction error as defined by Eq. 5-25 is also expressed in arbitrary units. In fact, to obtain an absolute measure, its value has to be multiplied by Δt.

The values filled in below for the naive and complete cross-correlation method are retrieved from figure 4 of [1] and therefore are approximate values. Given the fact that the RMS value is based on 100 realizations and the corresponding 95% confidence interval (0.877σ .. 1.163σ), the retrieved values are sufficiently accurate to allow comparison.

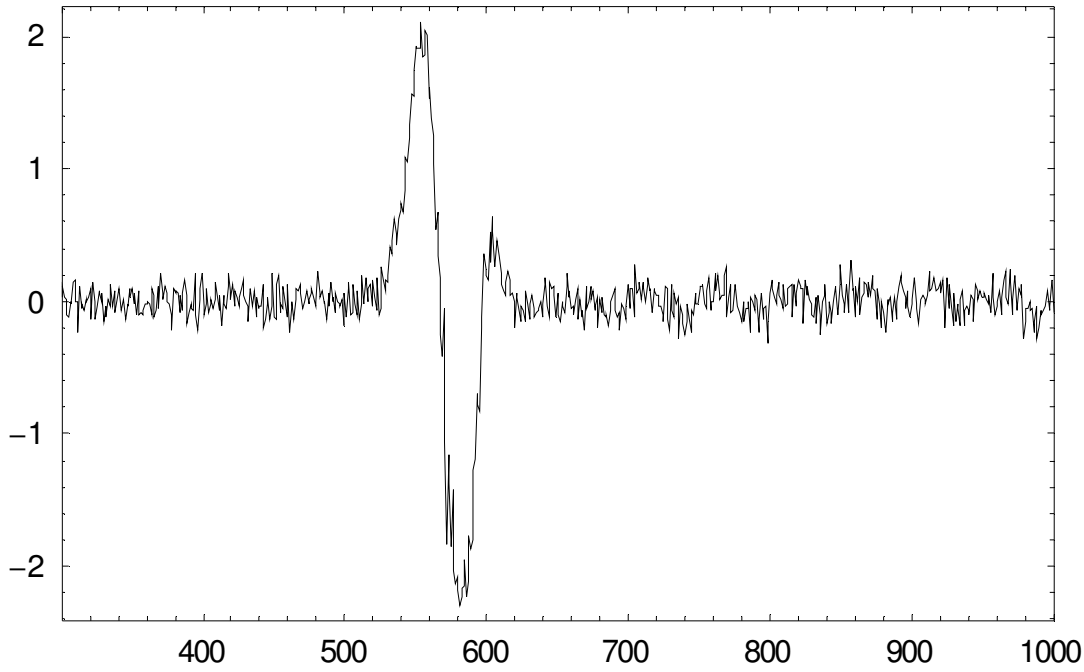

Figure 5-27. Zoomed version of the simulated signal in case of moderate jitter and moderate additive noise (both axes are labelled in arbitrary units).

The enhanced WLS estimator is only added for the case of moderate jitter. In order for this estimator to yield better results, jitter must be present and it must not be excessive in order for the third order model (Eq. 5-19) to remain valid.

All values below are given in arbitrary units, unless stated otherwise.

Comparison #1: no jitter, small additive noise

Estimator	RMS prediction error	95% confidence interval
naive cross-correlation	~0.03	
complete cross-correlation	~0.02	
naive LS	0.0164	0.0144 .. 0.0191
enhanced LS	0.0161	0.0141 .. 0.0187

Table 5-5. RMS prediction error for a simulation where $\sigma_{n_t} = 0$ and $\sigma_{n_y} = 0.02$.

It is clear that in the case of zero jitter and small additive noise, the naive LS method performs equally well as both the complete cross-correlation method and the enhanced LS method. The performance of the naive cross-correlation method is worse.

Comparison #2: no jitter, moderate additive noise

In the case of zero jitter and moderate additive noise, the naive LS method is not much worse than the complete cross-correlation method and performs significantly better than the naive cross-correlation method. It is believed that this is due to the fact that the latter suffers from interpolation problems. Indeed, an interpolation technique is required to increase the time resolution to a value that is smaller than Δt, and this may cause problems in the presence of noise.

In this case, using the aligned average as reference signal instead of the first realization reduces the RMS prediction error by a factor 1.75.

Estimator	RMS prediction error	95% confidence interval
naive cross-correlation	~0.55	
complete cross-correlation	~0.10	
naive LS	0.14	0.12 .. 0.16
enhanced LS	0.08	0.07 .. 0.09

Table 5-6. RMS prediction error for a simulation where $\sigma_{n_t} = 0$ and $\sigma_{n_y} = 0.1$.

Comparison #3: significant jitter, moderate additive noise

Estimator	RMS prediction error	95% confidence interval
naive cross-correlation	~1.10	
complete cross-correlation	~0.60	
naive LS	0.60	0.53 .. 0.70
enhanced LS	0.50	0.44 .. 0.58

Table 5-7. RMS prediction error for a simulation where $\sigma_{n_t} = 3$ and $\sigma_{n_y} = 0.1$.

One can question the experimental relevance of a jitter process, whose standard deviation $\sigma_{n_t} = 3$, relative to Δt. This means that the jitter may mix up the position of the samples over a range of 6 samples with a probability of 67%. Clearly, this is the kind of performance one wants to avoid in a practical setup. The comparison is added for completeness only.

Again, it is clear that the naive LS method outperforms the naive cross-correlation method. The naive LS and the complete cross-correlation method perform equally well. Although the improvement of the enhanced LS method is measurable, it remains limited.

Comparison #4: moderate jitter, small additive noise

Both the naive, the enhanced LS method and the complete cross-correlation method perform equally well. They outperform the naive cross-correlation method.

The enhanced WLS method reduces the RMS prediction error by a factor of 2 and clearly outperforms all other methods altogether.

The most important improvement of the WLS method is that it allows to compare the expected value of the cost to the realized value of the cost. Based on 100 realizations and 500 spectral component, the expected value of the cost is 99900 ± 894. The realized cost turns out to be 99525 and falls within the 95% confidence interval. Hence, it can be concluded that there are no detectable model errors, given the levels of noise of the simulation. The other methods provide no information to the user to draw such conclusions.

Estimator	RMS prediction error	95% confidence interval
naive cross-correlation	~0.25	
complete cross-correlation	~0.15	
naive LS	0.16	0.14 .. 0.19
enhanced LS	0.16	0.14 .. 0.19
enhanced WLS	0.08	0.07 .. 0.09

Table 5-8. RMS prediction error for a simulation where $\sigma_{n_t} = 1$ and $\sigma_{n_y} = 0.02$.

Comparison #5: moderate jitter, moderate additive noise

Estimator	RMS prediction error	95% confidence interval
naive cross-correlation	~0.60	
complete cross-correlation	~0.22	
naive LS	0.21	0.18 .. 0.24
enhanced LS	0.18	0.16 .. 0.21
enhanced WLS	0.16	0.14 .. 0.19

Table 5-9. RMS prediction error for a simulation where $\sigma_{n_t} = 1$ and $\sigma_{n_y} = 0.1$.

In this case, the performance of the naive LS method equals that of the complete cross-correlation method and is significantly better than the naive cross-correlation method.

Again, the enhanced WLS method has the best performance. However, the improvement with respect to the enhanced LS method is rather limited. The enhanced LS method in its turn shows a limited improvement over the naive LS method.

The realized cost of the WLS method turns out to be 99562 and lies within the 95% confidence interval of the expected value of the cost (99900 ± 894).

Conclusions

During all comparisons, the naive LS method clearly outperforms the naive cross-correlation method in estimating the drift in the presence of noise. It is believed that this is due to the fact that the latter suffers from interpolation problems. An interpolation technique is required to increase the time resolution to a value that is smaller than Δt, and this may cause problems in the presence of noise.

In the case of small additive noise or whenever the jitter contribution is dominant, the naive LS method and the complete cross-correlation method have the same performance.

In the case of moderate but dominant additive noise (comparison #2), the performance of the naive LS method is slightly worse than that of the complete cross-correlation method. At the same time, the performance gain of the enhanced LS method is meaningful.

In the case of moderate but dominant jitter noise (comparison #4), the performance gain of the WLS method with respect to the complete cross-correlation method and both the naive and enhance LS method is significant. This could be expected, as this is the case where the noise is maximally non-circular distributed.

In the presence of both moderate jitter and additive noise, where neither of both is dominant, some limited improvement can be observed of the enhanced WLS method over the enhanced LS method. In its turn, the latter is slightly better than both the naive LS and complete cross-correlation method.

Measurements

The same set of impulse response measurements is processed as in the chapter on jitter estimation.

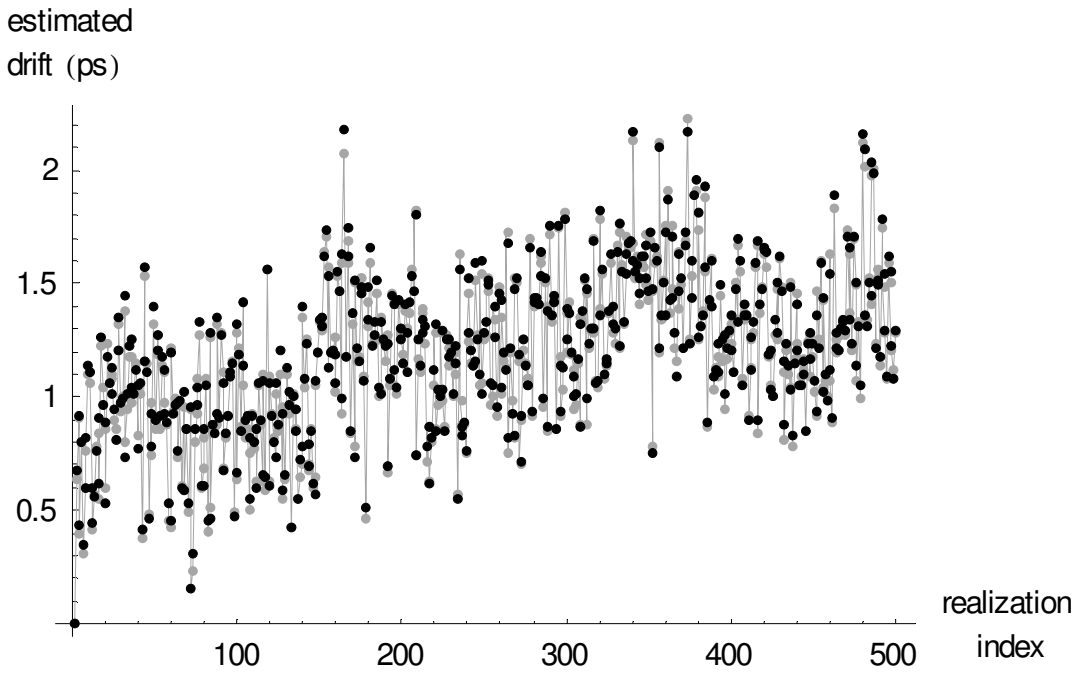

Figure 5-28. Comparison of the estimated drift using the naive LS method (connected gray dots) and using the enhanced LS method (black dots) for the first 500 measured realizations.

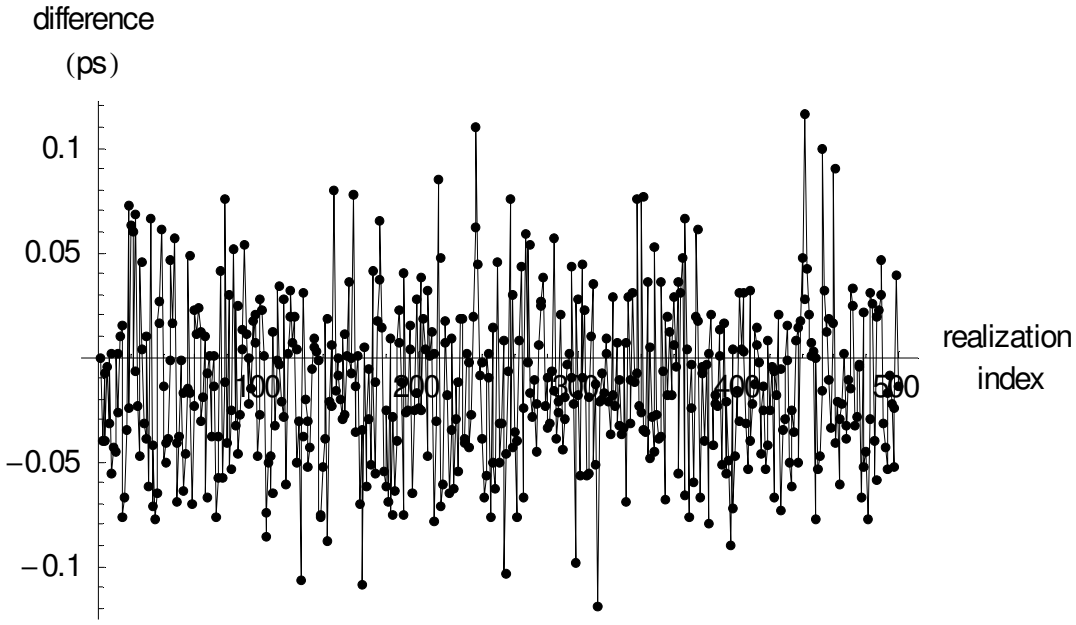

Figure 5-29. Difference of the estimated drift using the naive LS method and using the enhanced LS method for the first 500 measured realizations.

The data corresponds to the impulse response of an O/E converter, which was measured by Tracy Clement at NIST using an Agilent 83480A sampling oscilloscope in combination with a 83484A 50 GHz electrical plug-in.

The jitter standard deviation was estimated to be about 1 ps and the standard deviation of the additive noise was estimated to be about 0.5 mV. This information can be used to calculate the signal variance as a function of time. Using Eq. 5-18, it is then possible to construct the full covariance matrix in the frequency domain, where the real and the imaginary spectral contributions are considered separately.

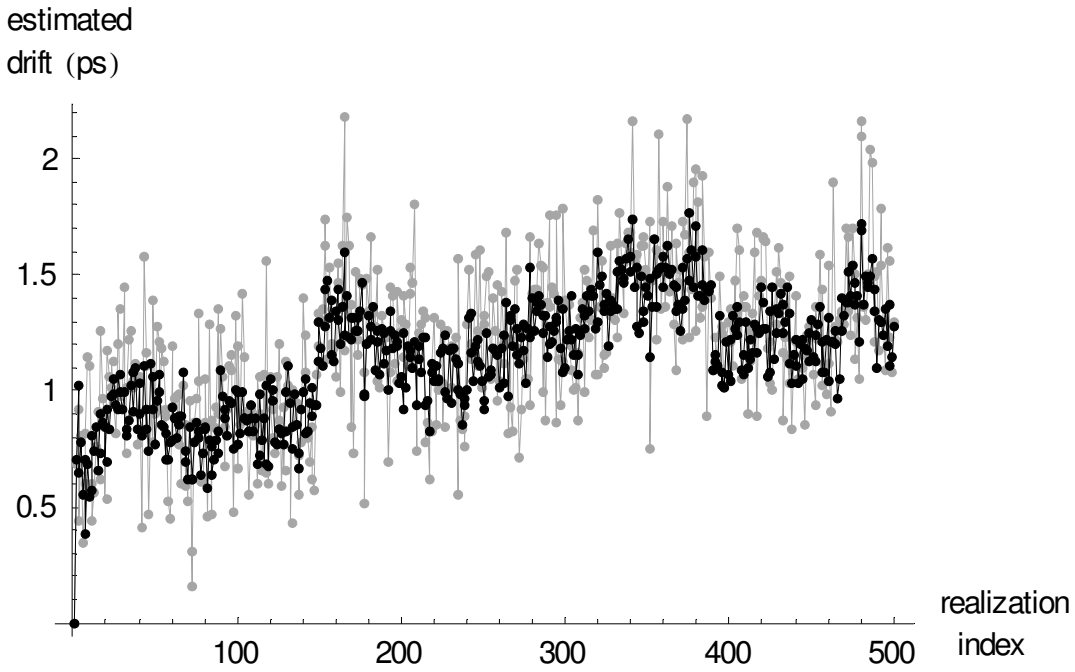

Figure 5-30. Comparison of the estimated drift using the enhanced LS method (connected gray dots) and using the enhanced WLS method (black dots) for the first 500 realizations.

Figure 5-28 compares the drift estimated by the naive LS method and the enhanced LS method. Although there is some difference, the difference is rather limited as is shown by figure 5-29.

However, when the full covariance matrix based on the estimated jitter and additive noise standard deviation is used to construct the WLS estimator, figure 5-30 shows a significantly smoother characteristic, resulting from a decreased uncertainty on the estimated drift. This is an expected property of the WLS, and is due to the use of a proper weighting of the frequency components in the cost function.

These results are consistent with those of simulation #4, where both the additive noise and jitter noise are moderate, but where the latter is dominant.

Finally, it makes sense to compare the standard deviation of the residual after applying a moving average window, to the uncertainty on the estimated drift. The latter is obtained using the values found on the diagonal of the parameter covariance

matrix. The latter is found to be 0.234 ps in the case of the enhanced LS estimator. Figure 5-31 and figure 5-32 show the smoothing based on a moving average window, which is 51 points wide.

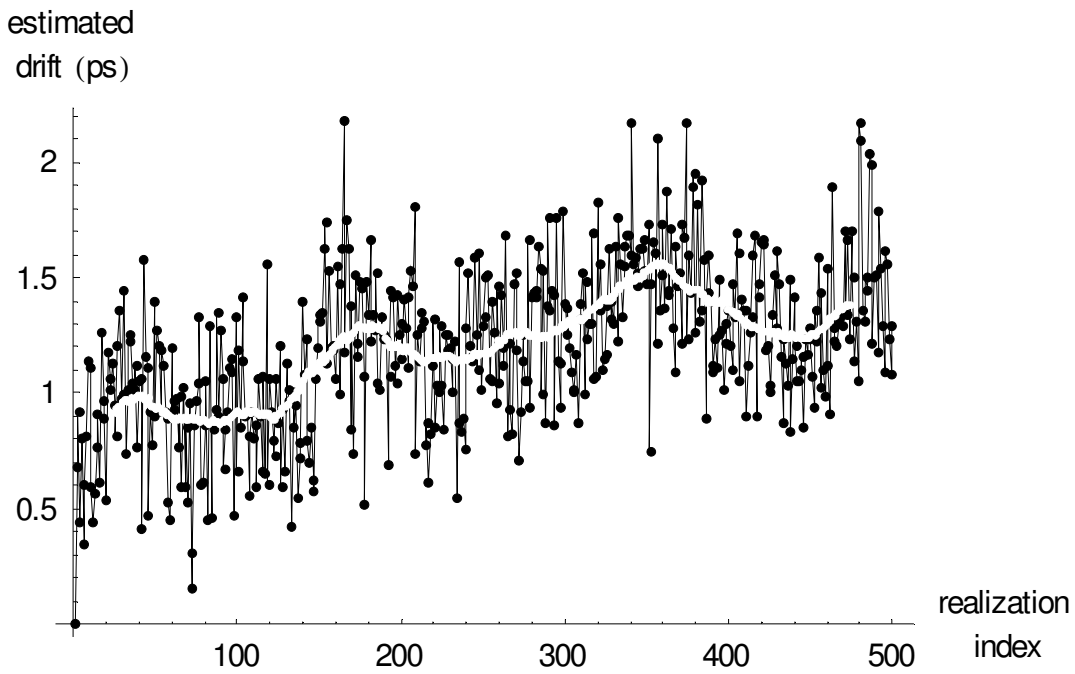

Figure 5-31. Estimated time base drift, using the enhanced LS method (white line: smoothed version using moving average window of width 51 samples).

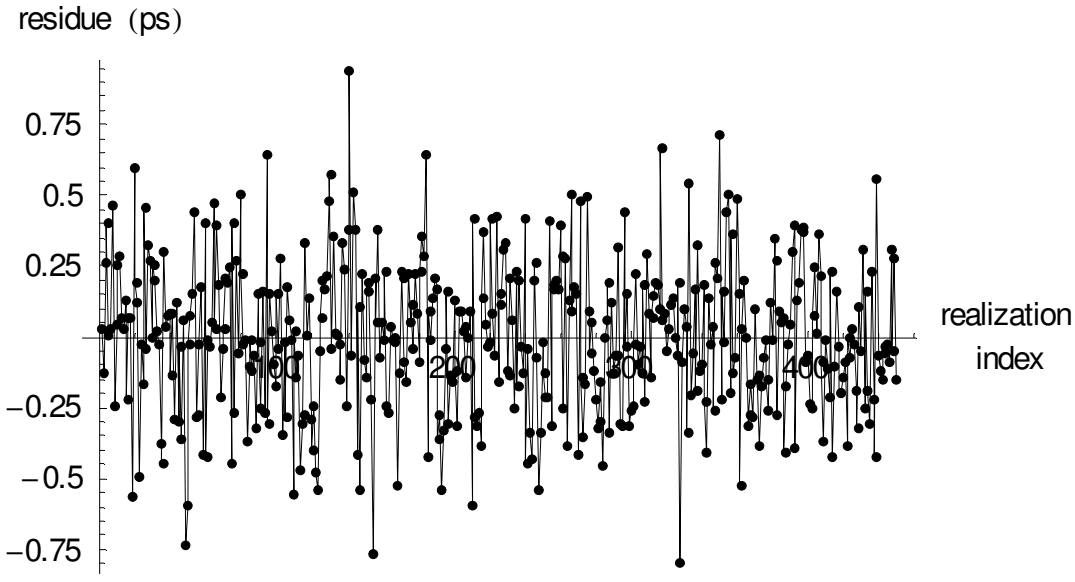

Figure 5-32. Residual of the estimated time base drift, using the enhanced LS method, after applying a moving average window of width 51 samples.

The residual has a standard deviation of 0.265 ps, which is only 13% larger than the uncertainty indicated by the parameter covariance matrix.

Using the enhanced WLS method, the residue is found to have a standard deviation of 0.128 ps, which confirms the reduction of the uncertainty by a factor 2 as was found based on simulation #4.

The above justifies the smoothing applied to the estimated drift in the chapter on jitter estimation.

Conclusions

Using proper system identification techniques, it is possible to come up with an estimate of time base drift in the presence of both additive and jitter noise. To the knowledge of the author, this method performs better than any other published technique.

The use of a proper weighting of the contribution of the individual spectral components to the cost function not only provides a relevant value for the cost function. It also reduces the uncertainty on the estimated drifts in the presence of a realistic quantity of additive and jitter noise by a factor of 2. The weighting is based on the full covariance matrix, where the contributions of the real and the imaginary part are separately taken into account.

Future research

The uncertainty on the estimated drift based on the parameter covariance matrix in case of the WLS estimator is about 0.6 times the realized uncertainty. For the time being, it is not understood why this discrepancy exists.

It is also possible to estimate a parametric model for τ, named $\tau(\theta)$, instead of a non-parametric one as above. The cost function is then minimized with respect to θ instead of with respect to τ_k using the fact that $\frac{\partial}{\partial \theta} = \frac{\partial}{\partial \tau_k} \cdot \frac{\partial \tau_k}{\partial \theta}$.

This approach is expected to work fine in a very stable measurement environment, resulting in a smooth drift characteristic as observed for the impulse response measurements performed at NIST.

References

[1] K. Coakley and P. Hale, “Alignment of Noisy Signals,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No. 1, February 2001
[2] A. Van den Bos, “Estimation of complex parameters,” Sysid ’94, X. IFAC/IFORS International Symposium on System Identification and Parameter Estimation, Copenhagen, 1994, Vol. 3, pp. 495–9
[3] H. W. Sorenson, “*Parameter estimation. Principles and problems.*,” Marcel Bekker Inc., 1980.
[4] E. W. Weisstein, "Wishart Distribution.", MathWorld, A Wolfram Web Resource. http://mathworld.wolfram.com/WishartDistribution.html

CHAPTER 6 Volterra-based behavioural modelling.

Abstract

The behaviour of a class of nonlinear devices, the PISPO[1] systems, can be described by the Volterra theory [1] [2] [3]. During the initial phase of the research, a black-box model for this type of devices was developed in the frequency domain: the Volterra input-output map, a.k.a. “VIOMAP”. The name refers to the fact that the model maps products of spectral components at one or more inputs onto the resulting spectral components at one or more outputs using a set of complex valued kernel values.

Initially, an optimized C program was written to generate all kernels of the model. It was able to deal with a large number of input frequencies applied to one or more inputs and to generate all the unique contributions for all the resulting output frequencies at one or more outputs for a given degree of nonlinearity in a very efficient way. It quickly became apparent that the effort required to generate these contributions was nothing compared to the determination of their values.

Nevertheless, it was possible to show the usability of the method for some applications, based on the fact that the VIOMAP is a natural extension of S-parameters for weakly nonlinear RF and microwave devices. The VIOMAP was used as alternative for load-pull measurements [4]. It was demonstrated that one is able to predict the overall behaviour of cascaded nonlinear RF and microwave devices [5]. The VIOMAP that was extracted using a CW[2] experiment was also used to predict the output of a nonlinear system excited by narrowband input signals [6]. Modelling a nonlinear device, based on measurements at one carrier frequency only, it is possible to predict its response to narrowband signals like 16QAM.

The next logical step is to try to compensate for the nonlinear behaviour of components like mixers and amplifiers, based on that model. Using an inverse of this model to predistort the IQ signal may result in poor reduction of spectral regrowth due to unmodelled subtle side effects. An extension of the experiment to use two-tone excitations, allows to extract a better model. The linearity of the overall compensated system is then significantly enhanced.

This work was presented at the International Microwave Symposium (IMS) in Orlando in 1995. A less theoretical version of the paper [7] received the “best conference paper award” at the 45th Automatic RF Techniques Group (ARFTG) conference, which was held in conjunction with IMS. A revised copy of the unpublished IMS paper is added here, as the method and its application are still up-to-date, even after 11 years.

1. A PISPO system is a system that, when excited with a periodic waveform, outputs a periodic waveform with the same periodicity.
2. continuous wave

Introduction

Mixers and amplifiers are indispensable components in telecommunication links, but are also the cause of potential waste of frequency spectrum through their nonlinear behaviour. On the other hand, being unable to use a power amplifier beyond its region of linear operation results in an inefficient use of the available DC power. An increased back-off from the 1 dB compression point results in a lower power efficiency.

In this work, it is shown that one can compensate for the nonlinear behaviour of a power amplifier operated under a narrowband excitation. First, the behaviour of the amplifier is measured at the carrier frequency using a one-tone experiment. Next, a VIOMAP model is extracted based on this experiment and the inverse model is calculated. This inverse model is then used to predistort the modulated data. Measurement of the response of the compensated system show that both the distortion of the constellation diagram and the pollution of out-of-band frequencies (spectral regrowth) are reduced when predistortion is applied to the base-band signal. On the other hand, it turns out that the inverse VIOMAP based on a one-tone measurement results in an overcompensation of the base-band signal. A more detailed study reveals a variation of the nonlinear behaviour of the amplifier in the immediate neighbourhood of the carrier frequency. This behaviour becomes apparent when exciting the amplifier with a two-tone and by varying the frequency spacing between the tones. This additional effect fully explains the overcompensation, which is visualized by integrating the VIOMAP model in an existing harmonic balance simulator. It provides an enhanced inverse VIOMAP resulting in a more efficient reduction of the distortion of the constellation diagram and the corresponding spectral regrowth.

In order to prevent confusion, the model extrapolation referred to in this work, is related to the covered range of input frequency components and not to the covered input power range, unless explicitly mentioned otherwise.

The one-tone VIOMAP and its inverse

Although the VIOMAP has a solid theoretical basis [1] [2] [3], it is not the aim of this work to stress its mathematical derivation but rather to show one of its applications. The VIOMAP has been developed for n-port devices [5], but in this work only the single input - single output case will be considered. This is possible because of the chosen frequency range and the fixed 50 Ω impedance.

First consider a nonlinear device which is excited by a one-tone signal at frequency f_c and for which one is only interested in the response at that frequency. Based on the Volterra theory, one can write that

$$Y(\omega_c) = H_1(\omega_c) \cdot X(\omega_c) + \sum_{i=1}^{N} \frac{(2i+1)!}{(i+1)! \cdot i!} \cdot H_{2i+1}\left(\underbrace{\omega_c, \ldots}_{i+1}, \underbrace{-\omega_c, \ldots}_{i} \right) \cdot \left|X(\omega_c)\right|^{2i} \cdot X(\omega_c) \qquad \text{Eq. 6-1}$$

where $\omega_c = 2\pi \cdot f_c$.

Based on a measured set of $\{X(\omega_c), Y(\omega_c)\}$, covering the input power range of interest, one can calculate the complex kernel values of the VIOMAP. $H_1(\omega_c)$ represents the small-signal gain and $H_{2i+1}(\omega_c, \ldots, -\omega_c, \ldots)$, $i = 1 \,..\, N$, represents the nonlinear behaviour of the device.

A VIOMAP which is extracted based on one-tone measurements only, will be referred to in the remaining of this chapter as a “one-tone VIOMAP”.

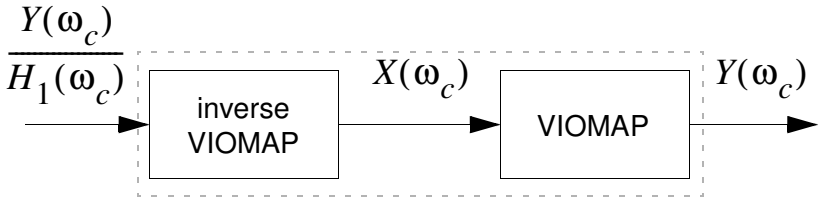

Figure 6-1. Generation of the inverse VIOMAP.

Predistorting the input signal $X(\omega_c)$, it is possible to compensate the nonlinear behaviour of the device at that frequency, as is illustrated by figure 6-1. The system will have an overall gain that is equal to the small-signal gain $H_1(\omega_c)$ of the device

under test, if for each measured output $Y(\omega_c)$ the corresponding input equals $\frac{Y(\omega_c)}{H_1(\omega_c)}$. One can now generate a second data set $\left\{\frac{Y(\omega_c)}{H_1(\omega_c)}, X(\omega_c)\right\}$ starting from the measured data set $\{X(\omega_c), Y(\omega_c)\}$ and use it to calculate the complex kernel values of the "inverse VIOMAP". It is important to notice that using this approach the behaviour of the overall system will only be linearized for input powers in the power range covered by $\frac{Y(\omega_c)}{H_1(\omega_c)}$, and not for the powers in the original range covered by $X(\omega_c)$.

Predistortion of narrowband signals based on an inverse VIOMAP

The one-tone VIOMAP, extracted based on measurements at the center frequency, can be extended to describe the response of weakly nonlinear systems to narrowband signals [6]. The verification of the validity of this “extrapolation” is one of the main topics of this work and will be discussed in detail further on.

A narrowband band-pass signal $x(t)$ can always be written as $a(t) \cdot \cos(\omega_c t + \phi(t))$. The function $a(t)$ is called the envelope of $x(t)$ and $\phi(t)$ is referred to as the phase modulation function. The corresponding output $y(t) = b(t) \cdot \cos(\omega_c t + \psi(t))$ can be predicted based on the VIOMAP, measured at f_c, as is shown below. Using the complex notation, one finds

$$X(\omega_c, t) = \frac{1}{2} \cdot a(t) \cdot e^{j\phi(t)}$$

$$Y(\omega_c, t) = \frac{1}{2} \cdot b(t) \cdot e^{j\psi(t)} = H_1(\omega_c) \cdot X(\omega_c, t) \qquad \text{Eq. 6-2}$$

$$+ \sum_{i=1}^{N} \frac{(2i+1)!}{(i+1)! \cdot i!} \cdot H_{2i+1}(\omega_c, .., -\omega_c, ..) \cdot \left|X(\omega_c, t)\right|^{2i} \cdot X(\omega_c, t)$$

In the case of IQ modulation, $x(t) = i(t) \cdot \cos(\omega_c t) - q(t) \cdot \sin(\omega_c t)$, such that $i(t)$ corresponds to $a(t) \cdot \cos[\phi(t)] = 2 \cdot re(X(\omega_c, t))$ and $q(t)$ to $a(t) \cdot \sin[\phi(t)] = 2 \cdot im(X(\omega_c, t))$. This way, it is easy to understand how to predistort $i(t)$ and $q(t)$: if $X(\omega_c, t) = \frac{1}{2} \cdot [i(t) + j \cdot q(t)]$ is changed to become $X'(\omega_c, t) = H^{inv}[X(\omega_c, t)] = \frac{1}{2} \cdot [i'(t) + jq'(t)]$, then the response of the system to the modulation $i'(t) + jq'(t)$ is the desired modulated signal. Here, H^{inv} represents the inverse VIOMAP, $i'(t)$ and $q'(t)$ represent the predistorted versions of $i(t)$ and $q(t)$, and $j = \sqrt{-1}$.

One-tone and two-tone VIOMAP: some theory

Consider a one-tone excitation experiment at a frequency f_c that is applied to a smooth[1] device. Next, consider a two-tone excitation experiment, centered around that frequency. The first tone is located at $f_c - \Delta f$, the second one at $f_c + \Delta f$ (figure 6-2). For the sake of simplicity, it is assumed that the nonlinear behaviour of the device can be described by a third degree nonlinearity. The VIOMAP will be generated for the one-tone and the two-tone case in order to find out when the VIOMAP extracted based on one-tone experiments, can be used to predict the outcome of two-tone experiments. Of course, the total power in both experiments is normalized to the same value.

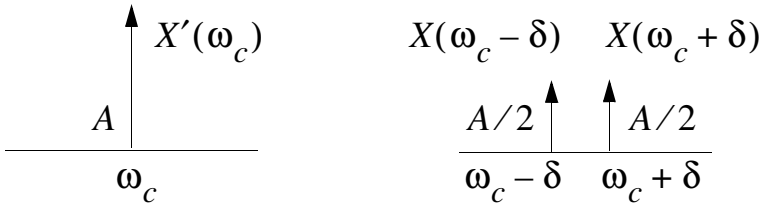

Figure 6-2. One-tone and two-tone experiments.

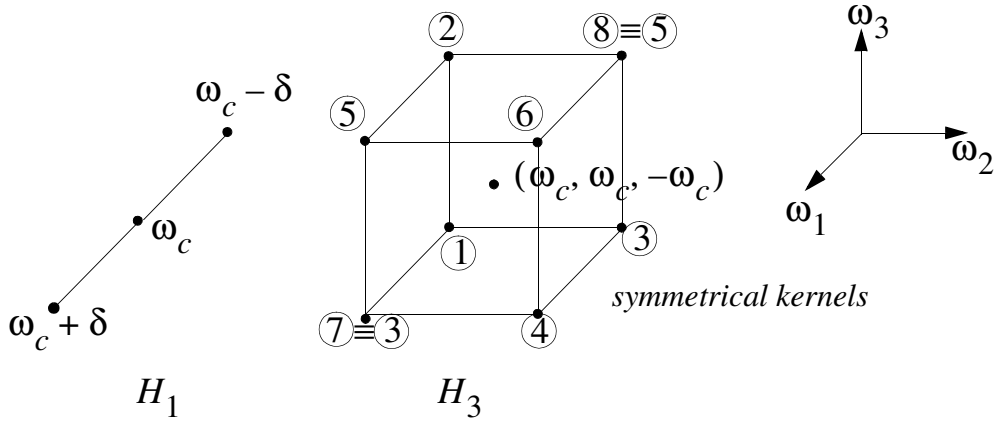

Figure 6-3. Visualization of VIOMAP kernels (numbering of kernels based on Eq. 6-4).

For the one-tone excitation, one obtains:

$$Y'(\omega_c) = H_1(\omega_c) \cdot X'(\omega_c) + 3H_3(\omega_c, \omega_c, -\omega_c) \cdot X'(\omega_c) \cdot X'(\omega_c) \cdot X'^*(\omega_c) \quad \text{Eq. 6-3}$$

1. The derivatives of a smooth system never become neither discontinuous, neither infinite.

Let $\delta = 2\pi\Delta f$, $\omega_+ = \omega_c + \delta$ and $\omega_- = \omega_c - \delta$, then the response for the two-tone experiment is:

$$
\begin{aligned}
&(1)\quad Y(\omega_c - 3\delta) = 3H_3(\omega_-, \omega_-, -\omega_+) \cdot X(\omega_-) \cdot X(\omega_-) \cdot X^*(\omega_+) \\
&\quad Y(\omega_c - \delta) = H_1(\omega_c - \delta) \cdot X(\omega_c - \delta) \\
&(2)\quad\quad + 3H_3(\omega_-, \omega_-, -\omega_-) \cdot X(\omega_-) \cdot X(\omega_-) \cdot X^*(\omega_-) \\
&(3)\quad\quad + 6H_3(\omega_-, \omega_+, -\omega_+) \cdot X(\omega_-) \cdot X(\omega_+) \cdot X^*(\omega_+) \\
&\quad Y(\omega_c + \delta) = H_1(\omega_c + \delta) \cdot X(\omega_c + \delta) \\
&(4)\quad\quad + 3H_3(\omega_+, \omega_+, -\omega_+) \cdot X(\omega_+) \cdot X(\omega_+) \cdot X^*(\omega_+) \\
&(5)\quad\quad + 6H_3(\omega_+, \omega_-, -\omega_-) \cdot X(\omega_+) \cdot X(\omega_-) \cdot X^*(\omega_-) \\
&(6)\quad Y(\omega_c + 3\delta) = 3H_3(\omega_+, \omega_+, -\omega_-) \cdot X(\omega_+) \cdot X(\omega_+) \cdot X^*(\omega_-) \\
&\quad \text{---} \sum(\delta \to 0) \text{---} \\
&\quad Y(\omega_c) = 2H_1(\omega_c) \cdot X(\omega_c) \\
&\quad\quad + 24H_3(\omega_c, \omega_c, -\omega_c) \cdot X(\omega_c) \cdot X(\omega_c) \cdot X^*(\omega_c)
\end{aligned}
\qquad \text{Eq. 6-4}
$$

Now let $X'(\omega_c) = 2 \cdot X(\omega_c)$ to normalize the power for both experiments. One then obtains:

$$
\begin{aligned}
Y(\omega_c) &= H_1(\omega_c) \cdot X'(\omega_c) \\
&\quad + 3H_3(\omega_c, \omega_c, -\omega_c) \cdot X'(\omega_c) \cdot X'(\omega_c) \cdot X'^*(\omega_c)
\end{aligned}
\qquad \text{Eq. 6-5}
$$

which is equal to Eq. 6-3.

Eq. 6-4 and figure 6-3 show under which conditions it becomes possible to use the kernels of the one-tone VIOMAP at a given center frequency f_c to predict the output in the case of two-tone measurements at the frequencies $f_c - \Delta f$ and $f_c + \Delta f$. This requires the monitoring of the variation of the complex kernels around $H_{2i+1}(\omega_c, ..., -\omega_c, ...)$ in the $(2i+1)$th-dimensional space. Here $(2i+1)$ represents the degree of nonlinearity. The more constant these kernels remain in the

neighbourhood of the point at coordinates $(\omega_c, ..., -\omega_c, ...)$, the better the prediction of the spectral components at the output will be. Hence, this results in a quasi-static hypothesis.

Due to the ill-conditioned nature of power series approximations, verification based on constant values of the VIOMAP kernels may turn out not to be a practically usable solution. A much simpler and more practical approach is to take a look at the variation of the generated spectral components at the output, both in amplitude and in phase, based on two-tone measurements with constant input power and varying frequency spacing. A concrete example of this approach can be found further on in this work.

Measurement setup and results

Measurement setup

Figure 6-4 shows a simple communication link consisting of an IQ modulator, an amplifier and an IQ demodulator [8]. The distortion introduced by mixers will not be considered in this work, although it can be characterized using the VIOMAP model too.

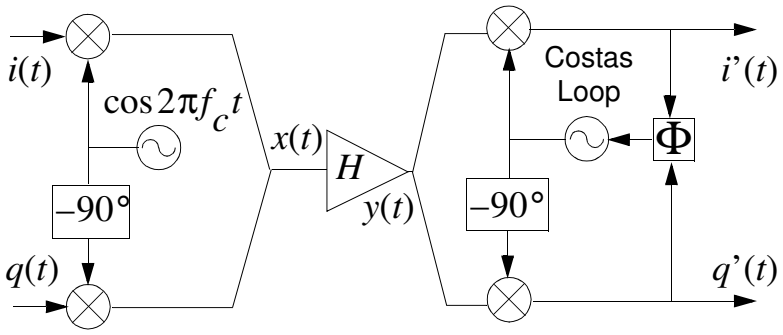

Figure 6-4. A simple communication link.

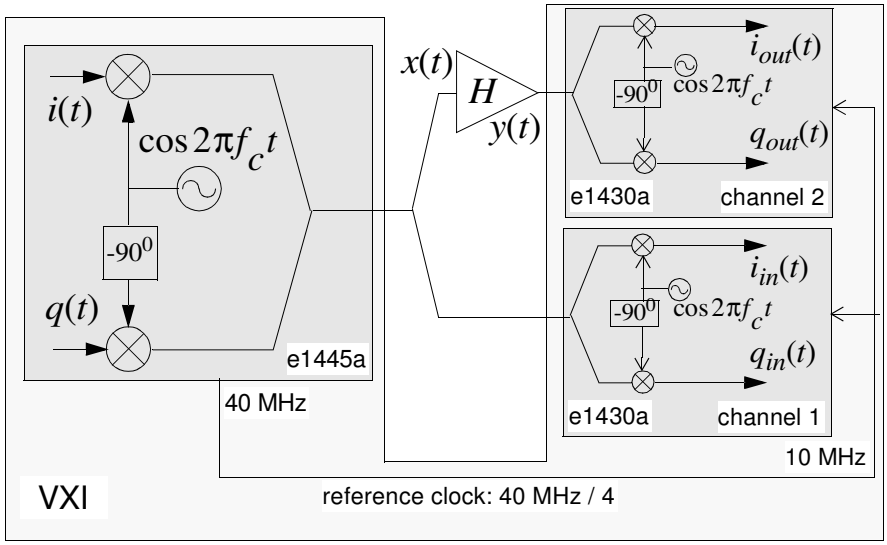

Figure 6-5. The VXI-based measurement setup.

To demonstrate the use of the VIOMAP for narrowband signals, a measurement setup (figure 6-5) is realized where the modulator and demodulator are implemented in software and converted to real-word signals using an “arbitrary waveform generator” VXI card (HP/Agilent E1445A) and two “analog-to-digital convertor” VXI cards (HP/Agilent E1430A). A calibration process using a stepped sine wave is performed to eliminate the effect of the 10 MHz reconstruction filter of the arbitrary waveform generator (AWG) and the characteristics of the anti-aliasing filters at the inputs of the

A/D convertors on the frequency grid of interest. The AWG runs at a sampling frequency of 40 MHz and generates the 10 MHz master clock of the acquisition cards putting a '0011' marker sequence on the VXI backplane. The sequence which is downloaded into the AWG is repeated continuously. At the end of each sequence a 25 ns pulse is generated via the 'Marker Out' front panel connector and is used to trigger the acquisition cards, allowing repeated triggered measurements.

First the one-tone VIOMAP of the broadband (20 kHz - 2 GHz) Sonoma amplifier is determined using a sine wave excitation at the carrier frequency f_c = 1.25 MHz. The input power is swept from -26 dBm to -7 dBm. The linear (small-signal) gain of this amplifier is found to be 20.75 dB. At an input power of -7 dBm this gain drops to 19.15 dB, resulting in a gain compression of 1.6 dB. The corresponding phase shift of 179.6 degrees is not sensitive to a change of the input power. In fact this is to be regretted, because the phase distortion is also described by the VIOMAP and therefore can be corrected for by the inverse VIOMAP.

As will be explained later, also two-tone signals are generated in the close vicinity of the carrier frequency. Fixing the input power and sweeping the frequency spacing reveals subtle side effects.

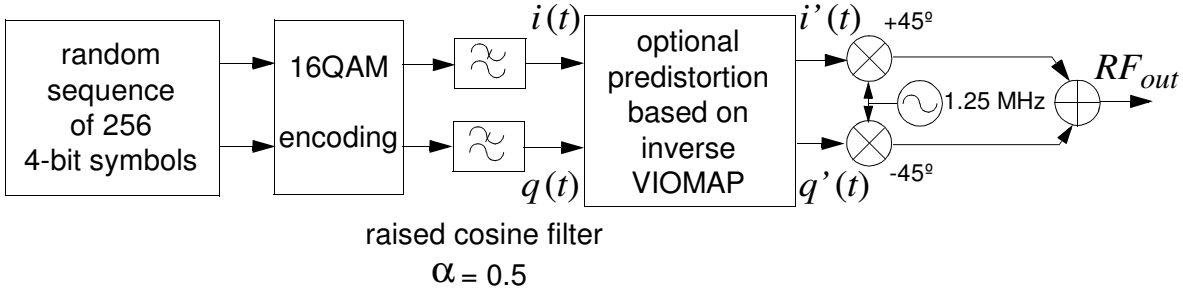

Figure 6-6. The generation of a (predistorted) 16QAM signal.

A random sequence of 256 4-bit words is generated in software (figure 6-6). The symbol rate in set at 10 MHz / 256 = 39 kHz and 16QAM encoding is selected. Choosing powers of 2 for the record length allows the use of Fast Fourier Transforms to transform waveforms between the time domain and the frequency domain. In order to keep the signals bandlimited, a raised-cosine filter [8] is used with a roll-off factor α = 0.5. This way, the bandwidth of the signal is limited to $(1 + \alpha)$ times the signal rate, which results in 58.6 kHz. The filtered $i(t)$ and $q(t)$ are optionally predistorted by the inverse VIOMAP to compensate the nonlinear behaviour of the Sonoma amplifier. Finally these signals are modulated with a carrier frequency of 10 MHz / 8 = 1.25 MHz. The resulting sequence is downloaded into the arbitrary waveform generator and is continuously repeated. The 38.75 MHz alias component of the carrier frequency is suppressed by the 10 MHz reconstruction filter. This is very important when dealing with nonlinear devices. The maximum levels of the baseband signals are scaled to correspond to an input power of -8.1 dBm, driving the amplifier

1.1 dB in compression. This input power must be covered by the VIOMAP in order to prevent extrapolation of the series approximation.

Model extraction.

A VIOMAP is extracted based on one-tone measurements at the carrier frequency, that cover an input power range of -26 dBm to -7 dBm. This one-tone VIOMAP contains 6 complex parameters $(H_1, H_3, ..., H_{11})$ and describes the power of the fundamental spectral component at the output as a function of the input power within 0.01 dB. The maximum phase deviation is smaller than the measurement noise. Predistortion based on the corresponding inverse VIOMAP reduces the distortion of the constellation diagram and the corresponding spectral regrowth, but the improvement is not as drastic as could be expected. Measurement of the constellation diagram at the output reveals the presence of overcompensation.

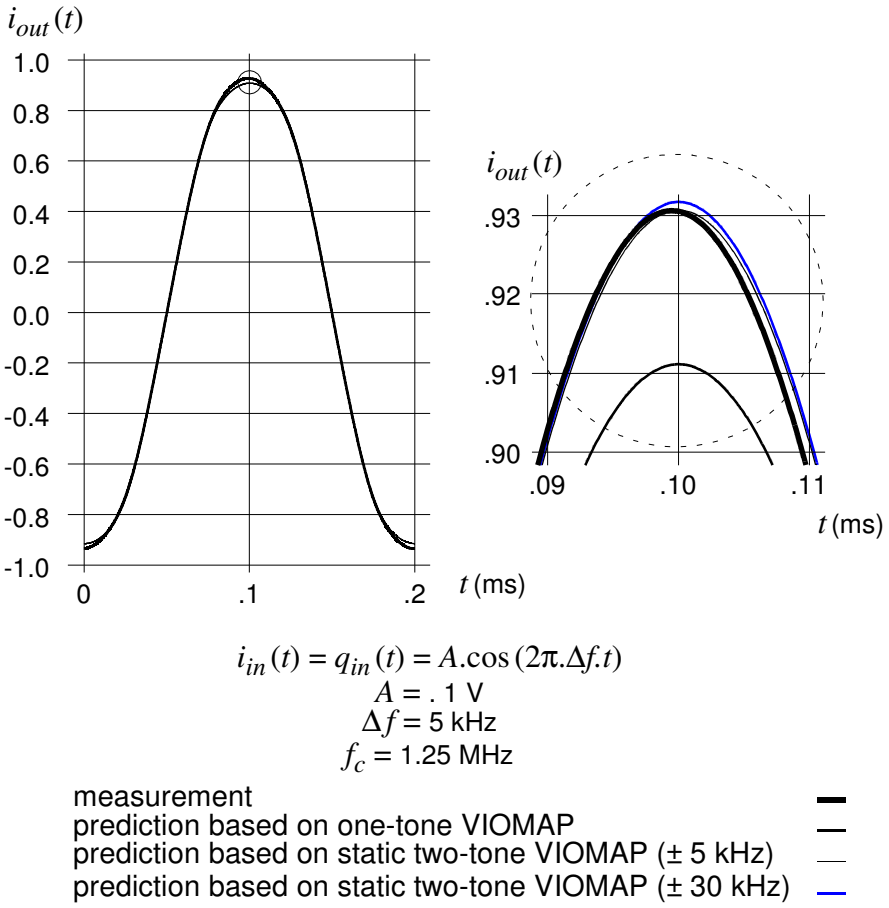

Figure 6-7. Measurement and prediction of $i_{out}(t)$ in the case of a simple IQ signal.

To investigate the effect in more detail, a simple IQ signal is generated resulting in a two-tone excitation of the amplifier. Choosing an appropriate baseband signal $i(t) = q(t) = A \cdot \cos(2\pi\Delta f \cdot t)$ the corresponding modulated signal $x(t)$

becomes $\frac{\sqrt{2}}{2} \cdot A \cdot \left\{ \cos\left[(\omega_c - \delta) \cdot t + \frac{\pi}{4} \right] + \cos\left[(\omega_c + \delta) \cdot t + \frac{\pi}{4} \right] \right\}$, with

$\delta = 2\pi\Delta f$. The Δf of 5 kHz falls within the bandwidth covered by the 16QAM signal and the amplitude of 0.1 V corresponds to an input power of the modulated signal of -7 dBm and falls within the input power range covered by the one-tone VIOMAP. Figure 6-7 shows the corresponding baseband signals at the output. Clearly the one-tone VIOMAP predicts more compression than the actual measurement shows.

First the level of the output of the one-tone measurement is verified around the carrier frequency by sweeping the input frequency, meanwhile keeping the input power fixed at -7 dBm. This experiment does not reveal any variation, neither in amplitude nor in phase, of the response in the neighbourhood of the carrier. Clearly, the deviations between measurement and prediction in the case of the simple IQ signal cannot be explained by such variations.

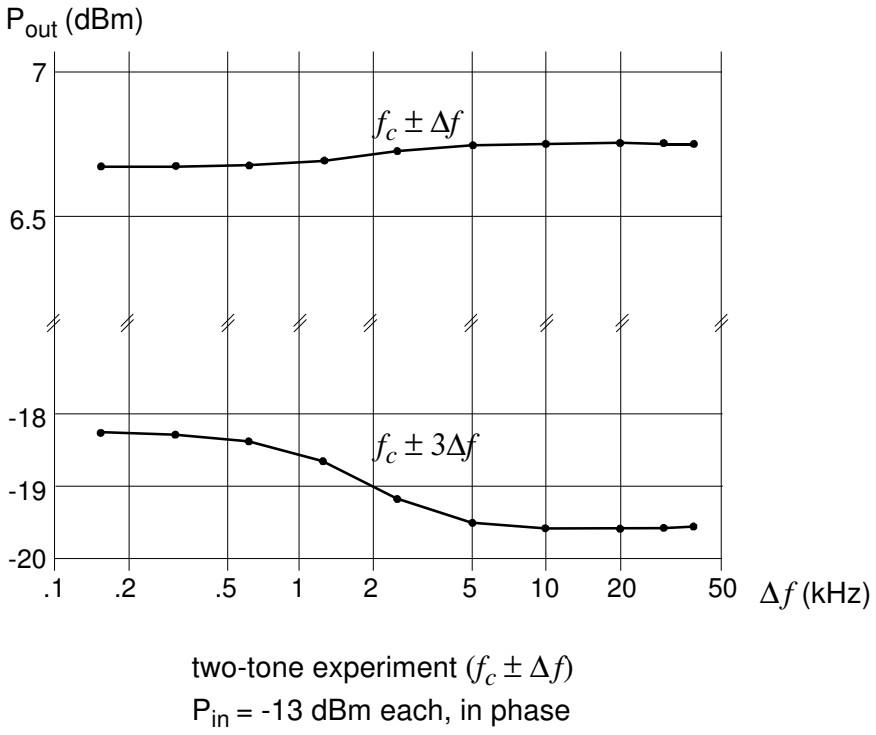

Figure 6-8. Variation of the nonlinear behaviour (in amplitude) of the amplifier as a function of the frequency spacing (logarithmic frequency scale).

Based on the fact that the above simple IQ signal corresponds to a two-tone excitation at the level of the device under test, it is decided to repeat the above experiment for a two-tone input signal. The first tone is located at $f_c - \Delta f$, the second one at $f_c + \Delta f$. The input power of each tone is fixed to -13 dBm, which is 6 dB below the input power of the one-tone (A becomes $\frac{A}{2} + \frac{A}{2}$). Δf is swept over the frequency range covered

by the modulated 16QAM signal. This experiment clearly shows a variation of the nonlinear behaviour of the amplifier as a function of Δf, as indicated by the variation of both amplitude (figure 6-8) and phase of the fundamental tones and their closest side lobes. Repeating the experiment at a carrier frequency f_c of 1.23 MHz, which corresponds to a Δf of 20 kHz when using a carrier at 1.25 MHz, results in identical plots. This way it is proven that the phenomenon is independent of small variations of the carrier frequency itself. The measurement of the linear gain of the amplifier at low frequencies (figure 6-9) reveals the presence of a transition zone starting at 10 kHz. This corresponds to twice the transition at a Δf of 5 kHz in the case of the two-tone experiment (figure 6-8). At that moment the frequency spacing of the two-tone equals 10 kHz. Assuming that the low-pass characteristic of the DC bias circuit has a comparable bandwidth and that its cut-off frequency is larger than that of the AC coupling of the Sonoma amplifier, the variation of the nonlinear behaviour can be explained by the self-biasing effect of the amplifier which is mirrored around the carrier frequency.

Under a one-tone excitation, the power of the RF input signal will determine the shift of the bias point, but it will only induce a shift in the DC bias settings. As soon as one modulates the one-tone signal, the DC bias will try to track the modulation. Because a DC bias network is a low-pass circuit, it can't always track the modulation immediately. Therefore, low-frequency voltages and currents are induced which reflect the low-frequent (LF) characteristics of the bias circuitry. This is referred to as (one possible source of) memory effects. Unfortunately, these effects get multiplied back to the RF signal such that one can observe the LF behaviour in the modulation of the RF signal. This leads to different input-output characteristics depending on the test signal.

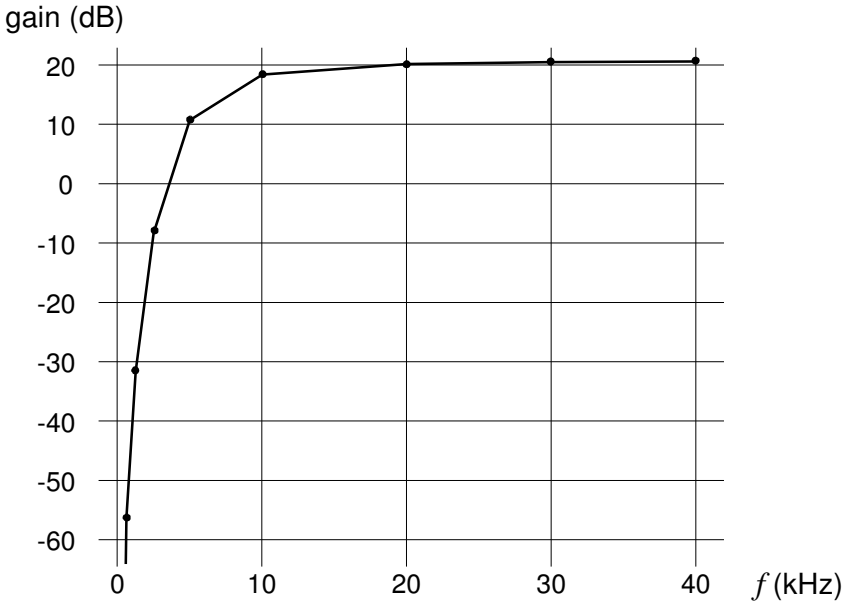

Figure 6-9. Linear gain of the broadband amplifier (P_{in} = -35 dBm).

Static two-tone VIOMAP and its inverse.

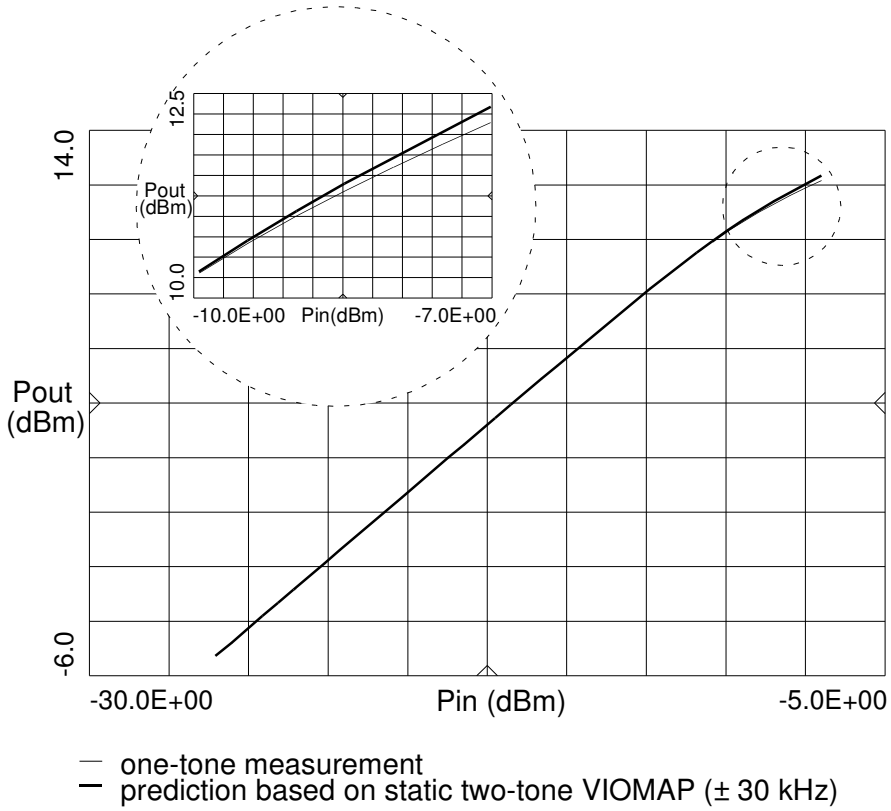

Figure 6-10. Compression characteristic of the amplifier at 1.25 MHz (one-tone vs. two-tone).

Based on the two-tone measurement that reveals the variation of the nonlinear behaviour of the amplifier as a function of the frequency spacing $2\Delta f$ of the two-tone, a better model can be extracted to predict the response of the amplifier in the case of the 16QAM signal. Because the 16QAM signal uniformly covers a frequency range up to a Δf of 29.3 kHz, 83% (= $\frac{29.3 - 5}{29.3}$) of the signal is well described by a two-tone measurement with a Δf higher than 5 kHz. Therefore, a static VIOMAP is extracted based on two-tone measurements with a Δf of 30 kHz. The input power of each tone is raised from -32 dBm to -13 dBm. The individual input powers of each tone are combined such that an equidistant two-dimensional grid is covered, each axis corresponding to the input power of each tone, expressed in dBm. The VIOMAP is chosen to be static in order to map its kernels $(H_1, H_3, ..., H_{11})$ on the ones found based on one-tone measurements at the carrier frequency. As such, these new kernel values are able to predict the outcome $Y_{pred}(\omega_c)$ of the set of one-tone experiments, which were used to generate the one-tone VIOMAP, thus creating a set of $\{X(\omega_c), Y_{pred}(\omega_c)\}$ pairs. As expected, the resulting compression characteristic

(figure 6-10), which is generated by implementing the VIOMAP model in an harmonic balance simulator, shows a lower level of compression than the actual one-tone measurements. Based on the corresponding set of $\left\{ \frac{Y_{pred}(\omega_c)}{H_1(\omega_c)}, X(\omega_c) \right\}$ a new inverse VIOMAP can be calculated. This model needs only 5 complex parameters to match the expansion characteristic (figure 6-11) within 0.01 dB. The phase deviation is again below the measurement noise.

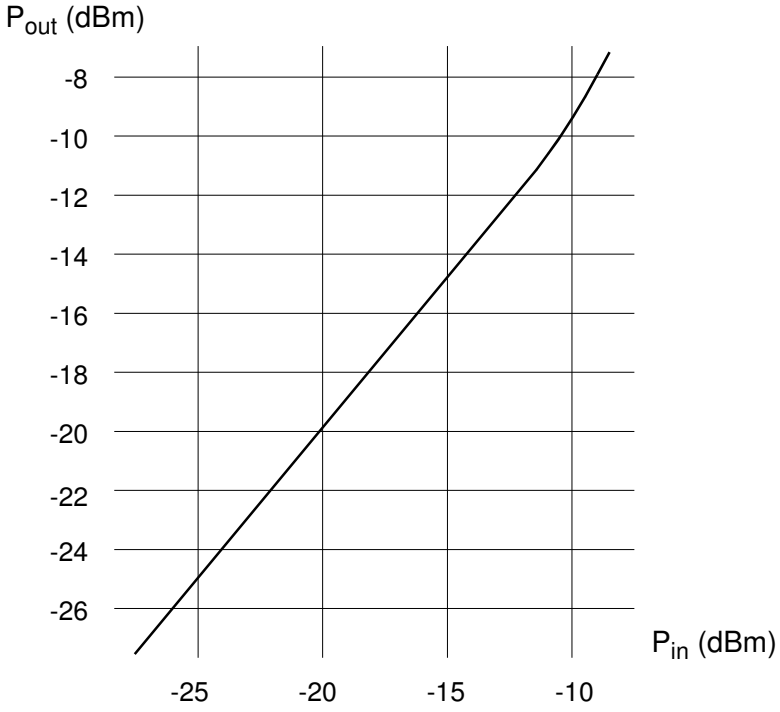

Figure 6-11. Predistortion characteristic based on the inverse (static two-tone) VIOMAP.

Predistortion based on static two-tone VIOMAP vs. one-tone VIOMAP.

In order to show the enhanced predistortion that is obtained when using the static two-tone VIOMAP instead of the one-tone VIOMAP, the 16QAM signal first is applied to the amplifier without predistortion. Figure 6-12 shows the spectrum of the modulated 16QAM signal at the input and at the output of the amplifier. Due to the nonlinear operation of the amplifier, the corresponding spectrum at its output spreads into the adjacent frequency bands. This phenomenon is referred to as “spectral regrowth”.

After demodulation, the compression of $i(t)$ as a result of the nonlinear behaviour of the amplifier is clearly visible on the generated eye diagram (figure 6-13). The outer levels do not reach a value of 0.78 V (based on the small-signal gain). Instead, two distinct levels appear due to crosstalk: if the simultaneous value of $q(t)$ at the input of

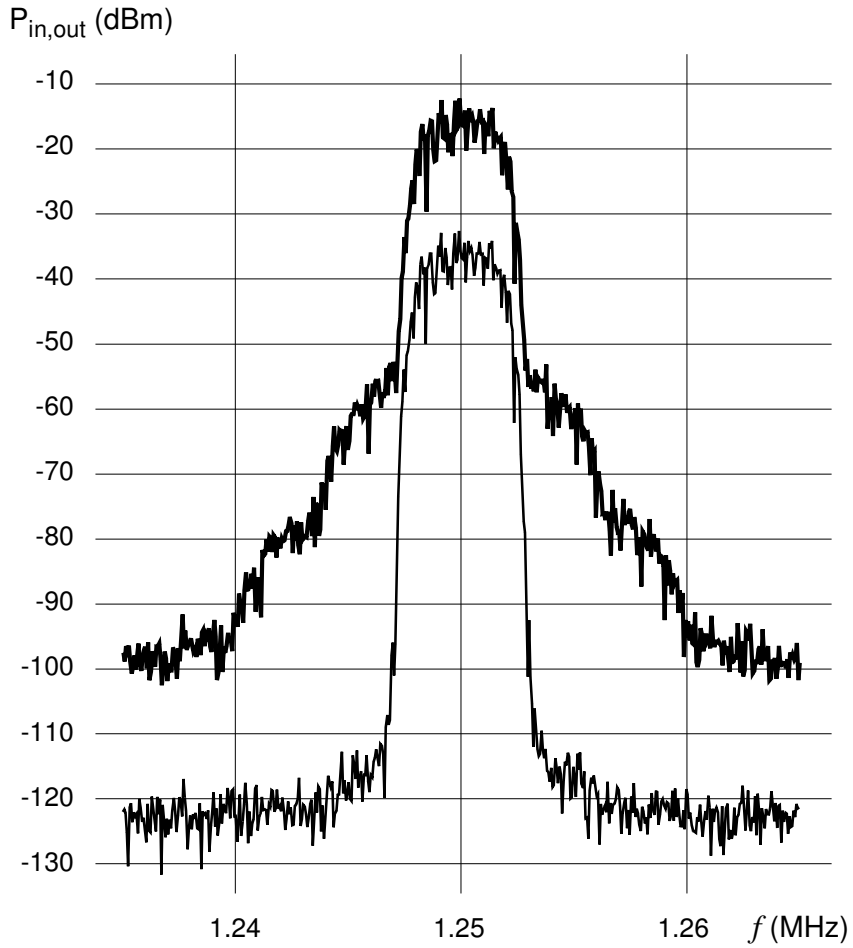

Figure 6-12. Original 16QAM input (thin) and output (thick) spectrum.

the amplifier correspond to one of both outer levels too, the amplifier will be driven more into compression and the resulting $i(t)$ at the output will be smaller than in the case the value of $q(t)$ at the input of the amplifier correspond to one of both inner levels. This is clear when looking at the corresponding constellation diagram (figure 6-14).

It should be noted that figure 6-13 and figure 6-14 correspond to measurements performed as part of [6]. In that case, the outer levels should reach a value of 1.2 V (based on the small-signal gain). During the predistortion experiment, the maximum levels of $i(t)$ and $q(t)$ have been reduced.

The effect of predistortion on the spectrum at the input of the amplifier is shown for both the one-tone and two-tone VIOMAP (upper portion of figure 6-15) and the resulting spectrum at the output clearly shows the enhancement when the predistortion is based on a static two-tone VIOMAP instead of a one-tone VIOMAP (lower portion of figure 6-15). The spectral regrowth is reduced an extra 10 dB,

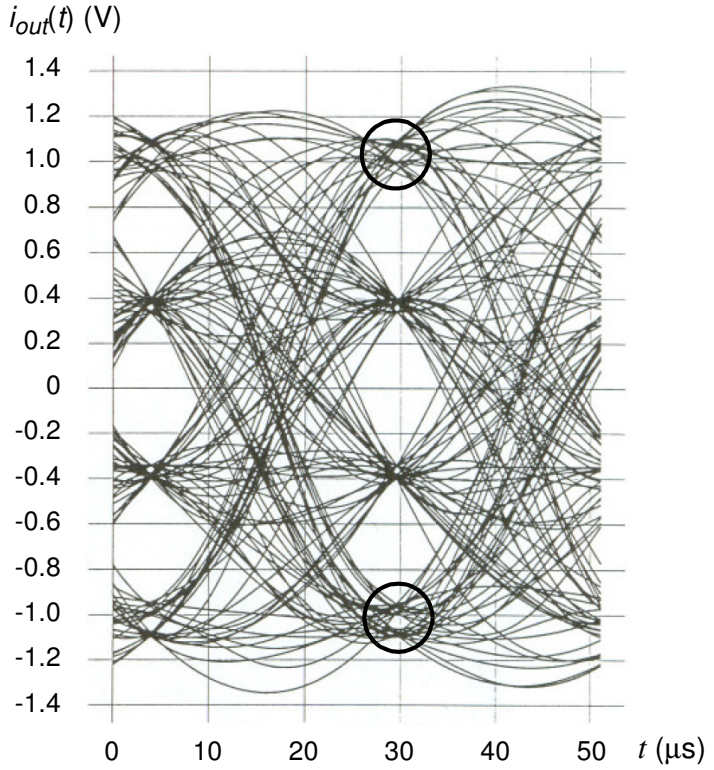

Figure 6-13. Measured eye diagram of $i(t)$ at the output of the amplifier (no predistortion) for maximum levels of $i(t)$ and $q(t)$ at the input, which result in 2 dB compression. Two distinct levels become apparent at the indicated outer levels (circles) and even at the inner levels.

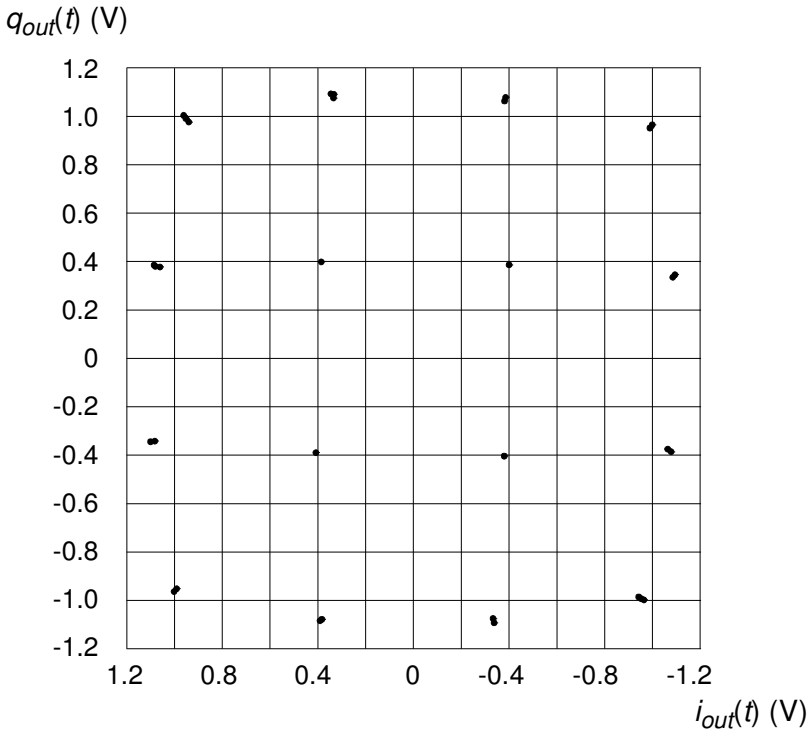

Figure 6-14. Measured constellation diagram at the output of the amplifier (no predistortion) for maximum levels of $i(t)$ and $q(t)$ at the input, which result in 2 dB compression.

resulting in an overall reduction of more than 16 dB. The corresponding $i(t)$ no longer shows the nonlinear effects which were clearly visible without predistortion.

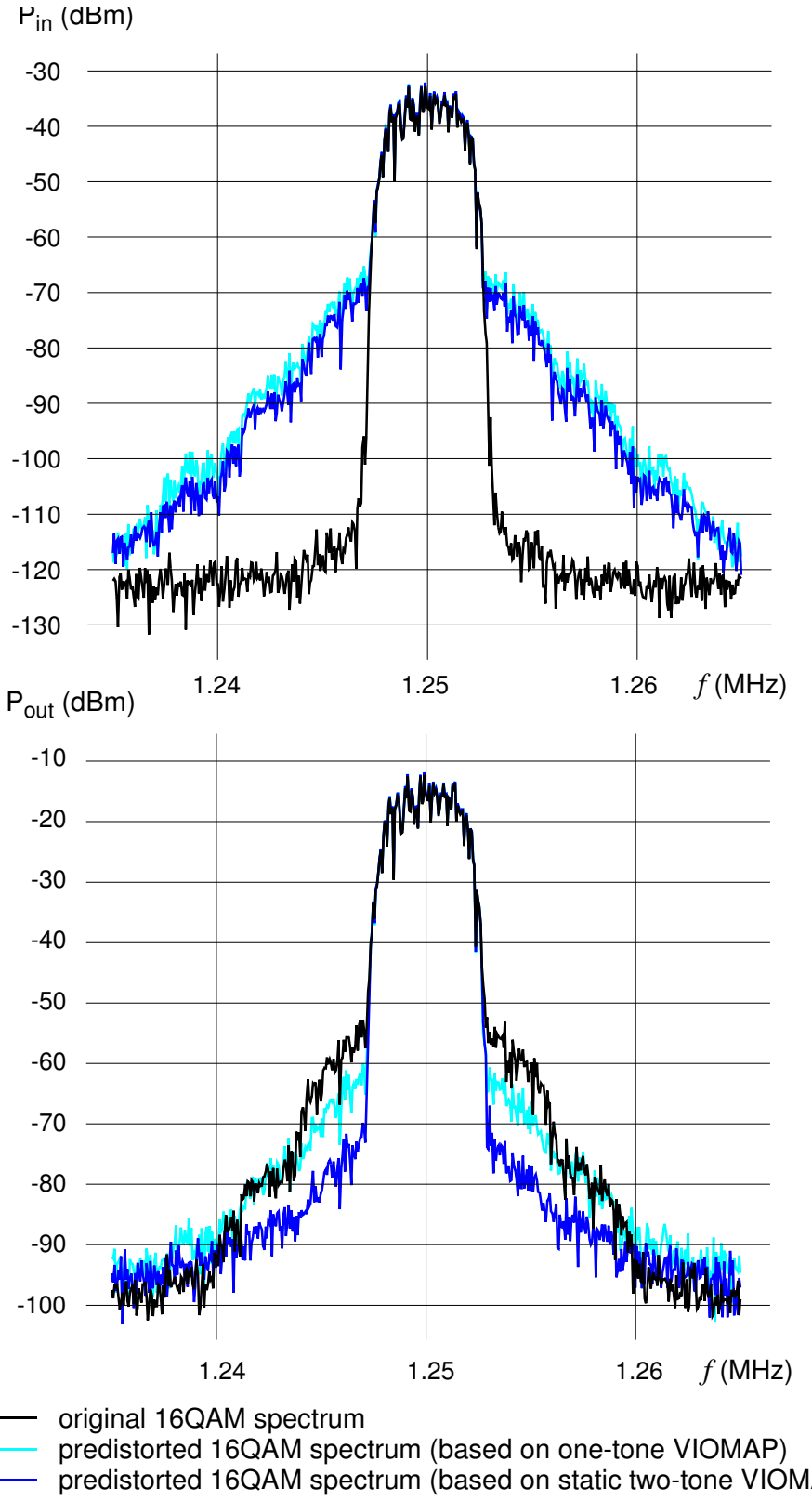

Figure 6-15. Original and predistorted 16QAM input and output spectrum.

Conclusions

It is shown that using an inverse VIOMAP, it is possible to enhance the linearity of digital communication channels. It is also demonstrated that one should be very careful when extrapolating a nonlinear model that has been extracted at the carrier frequency, to predict the response of the system to narrowband signals. Based on swept two-tone measurements one was able to explain the overcompensation which showed up during the 16QAM experiment when predistorting the baseband signals using an inverse VIOMAP based on measurements at the carrier frequency only. At the same time these measurements allowed the extraction of a better model resulting in an overall reduction of the spectral regrowth of more than 16 dB.

References

[1] M. Schetzen, “The Volterra and Wiener Theories of Nonlinear Systems”, Robert E. Krieger Publishing Company, 1989
[2] W.J. Rugh, “Nonlinear System Theory. The Volterra/Wiener Approach”, John Hopkins University Press, 1981
[3] M. Vanden Bossche, “Measuring Nonlinear Systems. A Black Box Approach for Instrument Implementation”, Doctoral Dissertation, Vrije Universiteit Brussel, May 1990
[4] F. Verbeyst and M. Vanden Bossche, “The Volterra Input-Output Map of a High Frequency Amplifier as a Practical Alternative to Load-Pull Measurements”, published in the *Conference Proceedings of IMTC/94* and published in the *Special Issue on Selected Papers IMTC/94 of the IEEE Transactions on Instrumentation and Measurement*, vol. 44, no. 3, June 1995, pp 662 - 665.
[5] F. Verbeyst and M. Vanden Bossche, “VIOMAP, the S-parameter equivalent for weakly nonlinear RF and microwave devices”, published in the *Microwave Symposium Digest of IEEE 1994 MTT-S International* and published in the *1994 Special Symposium Issue of the MTT Transactions*, vol. 42, no. 12, pp. 2531 - 2535.
[6] F. Verbeyst, J. Verspecht and M. Vanden Bossche, “VIOMAP, a Way to Predict the Distortion of a Constellation Diagram due to Amplifier Nonlinearities”, published in the *Digest of IEEE MTT-S European Topical Congress, Technologies for Wireless Applications*, Turin, November 1994.
[7] F. Verbeyst and M. Vanden Bossche, “VIOMAP, 16QAM and Spectral Regrowth: Enhanced Prediction and Predistortion based on Two-Tone Black-Box Model Extraction”, published in the *Proceedings of the 45th ARFTG Conference*, Orlando, June 1995 and winner of the “Best Conference Paper Award”.
[8] K. Feher, “Digital Communication. Satellite/Earth Station Engineering”, Prentice-Hall, 1983.

Conclusions and ideas for further research

"Large-Signal Network Analysis" is still in its infancy. The work presented here are humble contributions. Nevertheless it is believed that they represent a meaningful contribution to a growing community in a world where nonlinearities can no longer be ignored.

The LSNA hardware abstraction layer was conceived more than 10 years ago and survived the replacement of several major hardware components without causing any frustration neither to the conceiver nor to its users. It is still used today at different locations worldwide. A few years ago, Microsoft launched its .NET[1] initiative. The promise is that each of us can select the programming language we feel most comfortable with, as long as it targets the CLR[2]. It resulted in a new language called C#[3]. It is an attempt to take the best of C++ and Java. I believe languages like C, C++ and even Fortran remain key to the scientific community. Nevertheless, researchers should keep an eye on this new evolution, especially when they have the ambition to allow as many people as possible to use their work "as is". I'm convinced that new versions of the LSNA hardware abstraction layer should be ".NET aware", either by directly targeting the .NET platform or by using the interoperability provided by Microsoft between .NET managed code and native unmanaged code. Another evolution is driven by the multiple-core processors which were recently released by companies like Intel. Microsoft is working on concepts which should alleviate the pain for those who want to take advantage of multi-threaded programming.

The streamlined and enhanced implementation of the nose-to-nose calibration technique and its application as a part of the calibration of the Large-Signal Network Analyzer and the Lightwave Component Analyzer have served the community well and still does. The involvement of institutes like NIST was and remains essential to show to the community that good calibration techniques are key components to obtain good measurements. Furthermore, it fuelled the research at NIST related to the electro-optic sampling system to become an alternative to the nose-to-nose calibration technique. It also pushes the upper frequency limit at which calibrated LSNA measurements are possible. I'm convinced that it remains crucial to have independent techniques and as such I welcome further efforts, both with respect to nose-to-nose and EOS-based calibration techniques. Work has been performed at NIST (Technical Note 1528 by K. Remley) explaining the impact of the internal sampling circuitry on the phase error of the nose-to-nose calibration. This work and the work reported by J. Scott, as referenced in the "Comparison of the nose-to-nose and EOS-based calibration technique" chapter can serve as a starting point to improve the actual nose-to-nose calibration further. To convince the sceptics amongst

1. pronounced as "dot NET"
2. Common Language Runtime
3. pronounced as "C sharp"

us, it may help to set up a comparison between the results obtained at NIST and those based on other electro-optic sampling systems. Clever researchers may come up with a third method that allows broad-band calibration. One challenge is to push the upper frequency limit, increasing the frequency resolution of the phase calibration is another, as it is a key component to accurate large-signal measurements under modulated excitation. Therefore, I warmly welcome the recent efforts at the department and look forward to the fruits of this work.

Many high-frequency sampling oscilloscopes still suffer from time base imperfections, namely distortion, drift and jitter. To my knowledge, the system identification techniques which are applied to enhance both jitter and drift estimation embody the first work which tackles the effects of both jitter and drift simultaneously. Although the results can stand the comparison with other state-of-the-art techniques, it is expected that even better results can be obtained by combining the results of the chapter on jitter and drift estimation. Ideally, all time base effects should be considered simultaneously and correctly dealt with. Researchers, who feel challenged, may learn from a method[1] which was recently reported by people at NIST and the implementation is made available to others. A drawback of this approach is that it requires the measurement of two quadrature sinusoids performed simultaneously with the waveform of interest.

Simulations and designs rely on good models. The potential of Volterra-based models has been demonstrated through several applications. It is essential for practising engineers to understand that omnipotent models which truthfully describe the nonlinear behaviour of their components under all possible large-signal conditions simply don't exist. One of the major goals of the early work on predistortion is to show the potential pitfalls when using a model which is extracted using one class of excitation signals to predict the output of the system when applying another class of excitation signals. Predistortion based on such a model has the effect of a magnifying-glass. There is a strong need for additional research on measurement-based behavioural models. Researchers should "sell" new contributions by demonstrating their potential, without being tempted to "oversell" their models. I strongly believe that the use of random multi-sines will provide new insights in the analysis and modelling of large-signal behaviour. The major challenge there is to build a strong case to defend it in front of a potentially biased jury.

I believe that the research community has the difficult but challenging mission to push "Large-Signal Network Analysis" forward, without forgetting the poor souls out there, doing their best based on their present knowledge and experience. A balanced mix of fundamental research, education and carefully chosen applications which "speak the language" of the practising community is key to make "Large-Signal Network Analysis" a success for all of us.

1. "Compensation of Random and Systematic Timing Errors in Sampling Oscilloscopes", submitted for publication in the *IEEE Transactions on Instrumentation and Measurement*.
A copy of the revised version which was submitted to IEEE and the software can be downloaded from http://www.boulder.nist.gov/div815/HSM_Project/Software.htm

Publications

International periodicals

[1] Frans Verbeyst, M. Vanden Bossche, “Speed up power amplifier design by fast source-pull, real-time load-pull and accurate measurement-based behavioural models”, *Microwave Engineering Europe - Editorial*, November 2005, pp. 24-30 and Maury Microwave Application Note 5C-077

[2] J. Scott, J. Verspecht, B. Behnia, M. Vanden Bossche, A. Cognata, F. Verbeyst, M. Thorn, D. Scherrer, “Enhanced on-wafer time-domain waveform measurement through removal of interconnect dispersion and measurement instrument jitter”, *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 12, pp. 3022-28, December 2002

[3] F. Verbeyst, M. Vanden Bossche, “The Volterra Input-Output Map of a High-Frequency Amplifier as a Practical Alternative to Load-Pull Measurements”, *IEEE Transactions on Instrumentation and Measurement*, Vol. 44, No. 3, pp. 662-65, June 1995

[4] F. Verbeyst, M. Vanden Bossche, “VIOMAP, the S-parameter equivalent for weakly nonlinear RF and microwave devices”, *Special Symposium Issue of IEEE Transactions on Microwave Theory and Techniques*, Vol. 42, No. 12, pp. 2531-35, December 1994

[5] R. Pintelon, P. Guillaume, Y. Rolain, F. Verbeyst, “Identification of linear systems captured in a feedback loop”, *IEEE Transactions on Instrumentation and Measurement*, Vol. 41, No. 6, pp. 747-54, December 1992

National periodicals

[6] F. Verbeyst, E. Vandamme, “Large-signal network analysis. Overview of the measurement capabilities of a large-signal network analyzer”, *Revue HF*, No. 4, pp. 57-66, 2002

[7] F. Verbeyst, M. Vanden Bossche, “VIOMAP, 16QAM and spectral regrowth: enhanced prediction and predistortion based on two-tone black-box model extraction”, *Revue HF*, No. 2 : pp. 25-34, 1996

Conference papers

[8] L. Gommé, A. Barel, Y. Rolain, F. Verbeyst, “Fine frequency grid phase calibration setup for the Large Signal Network Analyzer”, *Proceedings of the IEEE MTT-S International Microwave Symposium*, MTT ‘06, June 2006, San Francisco, CA, USA

[9] F. Verbeyst, Y. Rolain, J. Schoukens, R. Pintelon, “System Identification Approach Applied to Jitter Estimation”, *IMTC Conference Proceedings*, pp. 1752-57, winner of a "Honorable mention recognized by the Award Commission of Agilent Technologies", IMTC ‘06, April 2006, Sorrento, Italy

[10] F. Verbeyst, M. Vanden Bossche, “Measurement-based Behavioral Model under Mismatched Conditions, a new and easy approach for an accurate model”, *Proceedings of the 35th European Microwave Conference*, October 2005, Paris, France

[11] S. Myoung, X. Cui, D. Chaillot, P. Roblin, F. Verbeyst, M. Vanden Bossche, S. Doo and W. Dai, “Large-signal network analyzer with trigger for baseband & RF system characterization with application to K-modeling & output baseband modulation linearization”, *Proceedings of the 64th ARFTG Conference*, pp. 189-95, December 2004, Orlando, USA

[12] F. Verbeyst, M. Vanden Bossche, “Real-time and optimal PA characterization speeds up PA design”, *Proceedings of the 34th European Microwave Conference*, pp. 431-34, October 2004, Amsterdam, The Netherlands

[13] S. Vandenplas, J. Verspecht, F. Verbeyst, E. Vandamme, M. Vanden Bossche, “Calibration Issues for the Large-Signal Network Analyzer”, *Proceedings of the 60th ARFTG Conference*, pp. 99-106, December 2002, Washington DC, USA

[14] G. Vandersteen, F. Verbeyst, P. Wambacq, S. Donnay, “High-frequency nonlinear amplifier model for the efficient evaluation of inband distortion under nonlinear load-pull conditions”, *Proceedings of the 2002 Design, Automation and Test in Europe Conference*, pp. 586-90, March 2002, Paris, France

[15] J. Scott, B. Behnia, M. Vanden Bossche, A. Cognata, J. Verspecht, F. Verbeyst, M. Thorn, D. Scherrer, “Removal of Cable and Connector Dispersion in Time-Domain Waveform Measurements on 40Gb Integrated Circuits”, *Proceedings of the IEEE MTT-S International Microwave Symposium*, MTT ‘02, Vol. 3, pp. 1669-72, June 2002, Seattle, USA

[16] J. Verspecht, F. Verbeyst, M. Vanden Bossche, “Network Analysis Beyond S-parameters: Characterizing and Modeling Component Behaviour under Modulated Large-Signal Operating Conditions”, *Proceedings of the 56th ARFTG Conference*, December 2000, Boulder, USA

[17] J. Verspecht, F. Verbeyst, M. Vanden Bossche, “Measurement based behavioral modeling of components under modulated large-signal operating conditions”, *Proceedings of the 30th European Conference on Wireless Technologies*, pp. 167-70, October 2000, Paris, France

[18] J. Verspecht, F. Verbeyst, M. Vanden Bossche, “Network analysis beyond S-parameters: Characterizing and modeling component behaviour under modulated large-signal operating conditions”, *Proceedings of the 30th European Microwave Conference*, Vol.2, pp. 373-76, October 2000, Paris, France

[19] D. DeGroot, P. Hale, M. Vanden Bossche, F. Verbeyst, J. Verspecht, “Analysis of Interconnection Networks and Mismatch in the Nose-to-Nose Calibration”, *Proceedings of the 55th ARFTG Conference*, pp. 116-21, winner of the "Best Poster Paper Award", June 2000, Boston, USA

[20] J. Verspecht, F. Verbeyst, M. Vanden Bossche, P. Van Esch, “System Level Simulation Benefits from Frequency Domain Behavioral Models of Mixers and Amplifiers”, *Proceedings of the 29th European Microwave Conference*, Vol. 2, pp. 29-32, October 1999, Munich, Germany

[21] M. Vanden Bossche, F. Verbeyst, J. Verspecht, “The Three Musketeers of Large Signal RF and Microwave Design - Measurement, Modeling and CAE”, *Proceedings of the 53rd ARFTG Conference*, June 1999, Anaheim, USA

[22] J. Verspecht, M. Vanden Bossche, F. Verbeyst, “Characterizing Components Under Large Signal Excitation: Defining Sensible `Large Signal S-Parameters'?!”, *Proceedings of the 49th ARFTG Conference*, pp. 109-17, June 1997, Denver, USA

[23] F. Verbeyst, “Using Orthogonal Polynomials as Alternative for VIOMAP to Model Hardly Nonlinear Devices”, *Proceedings of the 47th ARFTG Conference*, June 1996, San Francisco, USA

[24] F. Verbeyst, M. Vanden Bossche, “VIOMAP, 16QAM and Spectral Regrowth: Enhanced Prediction and Predistortion based on Two-Tone Black-Box Model Extraction”, *Proceedings of the 45th ARFTG Conference*, and winner of the “Best Conference Paper Award”, June 1995, Orlando, USA

[25] F. Verbeyst, M. Vanden Bossche, “Enhancing the Linearity of Digital Communication Channels Using IQ Predistortion Based Upon an Inverse VIOMAP”, *IEEE MTT-S International Microwave Symposium*, MTT ‘95, May 1995, Orlando, USA

[26] F. Verbeyst, J. Verspecht, M. Vanden Bossche, “VIOMAP, a Way to Predict the Distortion of a Constellation Diagram due to Amplifier Nonlinearities”, *Proceedings of the IEEE MTT-S European Topical Congress*, Technologies for Wireless Applications, pp. 81-85, November 1994, Turin, Italy

[27] F. Verbeyst, M. Vanden Bossche, “VIOMAP, the S-parameter equivalent for weakly nonlinear RF and microwave devices”, *Proceedings of the IEEE MTT-S International Microwave Symposium*, Vol. 3, pp. 1369-72, MTT ‘94, May 1994, San Diego, USA

[28] F. Verbeyst, M. Vanden Bossche, “The Volterra input-output map of a high frequency amplifier as a practical alternative to load-pull measurements”, *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, Vol. 1, pp. 283-86, IMTC '94, May 1994, Hamamatsu, Japan

[29] R. Pintelon, P. Guillaume, Y. Rolain, F. Verbeyst, “Identification of linear systems captured in a feedback loop”, *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, pp. 14-20, IMTC '92, New York, USA

Patents

"Method of and an arrangement for characterizing non-linear behavior of RF and microwave devices in a near matched environment."

F. Verbeyst, J. Verspecht
US Patent Application Publication, No. US 2003/0057963 A1
USA, March 2003

Awards

[1] Best Conference Paper Award at the 45th ARFTG Conference for the paper: “VIOMAP, 16QAM and Spectral Regrowth: Enhanced Prediction and Predistortion based on Two-Tone Black-Box Model Extraction” by F. Verbeyst and M. Vanden Bossche, Orlando, USA, June 1995.

[2] ARFTG Technology Award[1] for the “Development of Large-Signal Measurement Technology”, Washington DC, USA, December 2002.

[3] IMTC 2006 Honorable Mention recognized by the Agilent Technologies Award Commission for the paper: “System Identification Approach Applied to Jitter Estimation” by F. Verbeyst, Y. Rolain, J. Schoukens and R. Pintelon, Sorrento, Italy, April 2006.

1. Recipients of the same award: M. Vanden Bossche and J. Verspecht.