

ANRITSU
TECHNICAL
REVIEW

Sep.
2024
No. **32**

Anritsu Technical Review No.32
Issue on September 2024

© 2024 ALL RIGHTS RESERVED
by ANRITSU CORPORATION
5-1-1 Onna, Atsugi-shi, Kanagawa
243-8555, Japan
<https://www.anritsu.com>

Articles

Role of Anritsu 5G LAB in Resolving Industrial Digital Transformation (DX) Challenges by Satoshi Sasaki, Koichi Ryu, Genya Yats, Koji Arimochi, Yoshiaki Niwano, Tsutomu Konishi, Makoto Fujii	1
O-RAN Radio Unit Test System for Promoting Open RAN by Koichi Inoue, Takaaki Kamisawa, Koji Fukazawa, Hiroyuki Tanikage, Hiroki Azuma, DongGeun Park	8
Development of Wideband Linear Amp with New InP HBT by Tomoaki Ishii, Yukio Ohkubo, Masaya Suzuki, Takumi Nemoto, Yutaka Kobayashi, Takashi Kamizono, Moeko Jimba	17
Development of the Next-Generation Flagship Checkweigher Series by Hirotaka Anzai, Kenta Yamada, Tomohiro Okamura, Yusuke Noguchi, Tsubasa Susa, Jun Kubota, Toshiaki Yamazaki	24
Automating External Visual Inspection Using Collaborative Robot by Tomohiro Akiyama, Takashi Yoshida, Hideyuki Sakamoto, Tsunayoshi Hara, Ryo Kondo	32
Enhanced Raman scattering by zigzag edges of graphene and spectral analysis using the Bayesian method by Taisuke Ochi, Tomohiro Matsui	39
Overview of design method of millimeter-wave direct coupled resonator filters by Takahiko Mizuno	48
Development of Pulse Optical Spectrum Measurement Technology using Time Averaging by Soichiro Kai, Maki Ueno, Shinji Morimoto, Ryosuke Nakagawara, Shinichi Seto	58
Comparative Study of Antenna Array Calibration Methods Based on Symmetrical Mutual Coupling of 4-Element Linear Array by Takashi Mori	64
RF Switch MMIC Based on InP DHBT Technologies for Sub-THz by Mohammad Salah, Alexander Chenakin	70
Switching Speed Measurements of Barium Strontium Titanate (BST) Varactors by Tom Roberts, Jon Martens	75
Synchronized Vector Network Measurements Over Long Distances by Don Bradley, Paul Davis, Aaron Miller, Jon Martens, Navneet Kataria	81
Contribution to other Publications	86

Role of Anritsu 5G LAB in Resolving Industrial Digital Transformation (DX) Challenges

Satoshi Sasaki, Koichi Ryu, Genya Yatsu, Koji Arimochi, Yoshiaki Niwano, Tsutomu Konishi, Makoto Fujii

[Summary]

Progress in digital transformation (DX) using IoT and AI technologies in various industries is advancing against the social background of labor shortages. Collecting digital data to manage factory equipment, monitor cameras and control robots, etc., over a digital network faces key challenges such as not only analyzing network faults, but also predicting and preventing problems. This article introduces Anritsu 5G LAB leveraging our company's long experience in data communications and measurement technologies as a site to test network technologies forming the basis of industrial DX, solve issues, and co-create valuable information freely with customers.

1 Introduction

Progress in digital transformation (DX) is increasing expectations for local 5G as one technology for resolving challenges, such as labor shortages in various industries.

Local 5G is not a communications carrier's business but is instead a dedicated 5G network system installed in buildings and on sites owned by businesses and local governments. The first applications for a license to operate a wireless station were accepted in December 2019, and by late January 2024 there were 1,311 stations operating in the 4.8 GHz band (sub-6 GHz) and 123 stations in the 28 GHz band (mmWave).

Like commercial operators' 5G communications businesses, Local 5G not only features high speed and capacity communication, low-latency, multi-user connections but also has the advantages of high security and resilience to the impact of external traffic due to usage-based design of base stations and networks. Consequently, this technology is being developed widely by businesses and local governments on construction sites, in factories, at disaster-prevention locations, etc. (Figure 1).

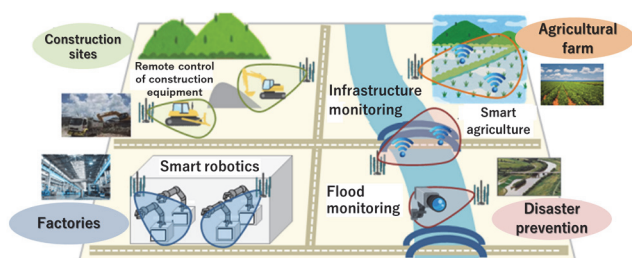
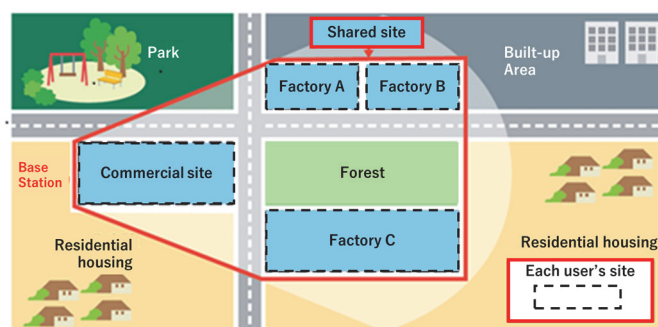


Figure 1 Quoted from FY2020 Ministry of Internal Affairs and Communications Report on "Validating Development of Local 5G, etc., for Solving Regional Issues"

Moreover, since 2023 a shared-use concept has started to develop under certain conditions where Local 5G access is shared between one owner's site and another site, supporting flexible introduction of the technology (Figure 2).



Source: FY2022 Report from the New Generation Mobile Communication System Committee of the Ministry of Internal Affairs.

Figure 2 Shared-Use Concept

However, introduction of Local 5G requires similar work procedures to those of communications operators, ranging from selection of equipment to area coverage design, acquiring wireless license, and post-installation operation and maintenance, all of which require high-level skills that can present major challenges for businesses, local governments, and system integrators (SIer). In addition, due to most of the regarding technologies are still in the demonstration stages, there are few practical environments where people can experience how local 5G can be used to solve local issues faced by business, local governments, and the other organizations yet.

As a result, Anritsu opened the Anritsu 5G LAB (5G LAB) in April 2021 to promote understanding of Local 5G and support technical validation by introducing Local 5G use cases and innovations. The 5G LAB is helping implement smart factories as part of Japan's industrial DX as well as solve

challenges in operating and maintaining Local 5G systems (Figure 3).



Figure 3 Local 5G Service Solutions

This article explains the role of the 5G LAB in supporting solutions for industrial DX challenges using Local 5G in different business fields.

2 Establishment Approach

A project team was established in July 2020, targeting the start of 5G LAB operations in April 2021, and investigation of millimeter wave (mmWave) base station designs was started.

The functions and performance of base stations from both Japanese and overseas makers were compared and a Nokia product was selected due to its excellent functions, performance, availability, and cost. To acquire knowledge about the procedures from obtaining a radio license to starting operations, each of these procedures were handled by ourselves. Figure 4 shows the required stages from obtaining the license to starting operations.



Figure 4 Local 5G Base Station Qualification Process

2.1 Area and Site Design

The area coverage and site designs to minimize interference are the most important procedures in obtaining a license.

A key principle in deploying Local 5G is use only within owned sites and structures and taking strong countermeasures to prevent radio waves leaking outside the above range and stop radio interference with other radio stations. On the other hand, the radio wave must be easily available within the usage range, requiring careful planning to satisfy both needs by simulating changing base station conditions, such as output power, and antenna height, gain, and directivity to

design the area and site and create the final distribution map in conformance with the license standards. The distribution map is created using the commercial coverage simulation KCAMP software from KOZO KEIKAKU ENGINEERING Inc. (Figure. 5).



Creating Heat Map by Processing Geographic Tiles

Figure 5 Local 5G Base Station Coverage Simulation

2.2 Interference Adjustment

The distribution map created by KCAMP shows the radio wave power range in the 5G LAB coverage area as well as areas where there may be an impact on other services (Adjustment target area). In addition to use as reference materials at inspection by the Ministry of Internal Affairs and Communications (MIC), the created map was also used for adjusting to prevent outside leakage and interference with nearby communications operators. An application was made to MIC after exchanging the agreement documents and the 5G LAB received a license in February 2021.

Licensing of the NEC Sub-6 GHz base station used the same procedure as the mmWave base station. The license was obtained, and operations started in February 2022.



Figure 6 28 GHz Band Local 5G Base Station



Figure 7 4.8 GHz Band Local 5G Base Station

3 Role of 5G LAB

The role of the 5G LAB includes radio wave visualization using instrumentation, measurement and verification services for brought-in systems and demonstration test of use cases.

3.1 Observing Waveforms with Test Instruments

Although there are expectations for using Local 5G as a high reliability wireless network technology for mission critical applications, sometimes the wireless environment is affected by small changes in the local environment, such as the movement of people and vehicles and changes in the layout. The 5G LAB uses resources, such as test instruments and metal shields, to analyze the impact of obstructions, reflections, multiple transmission paths, etc., on changes in the radio-wave power, transmission speed, and latency. The information obtained by this analysis can be used to test equipment layouts and troubleshoot problems to help optimize the site design and prevent faults.

Securing optimum area coverage as well as stable communications quality requires both simulation and actual survey evaluations. The 5G LAB uses measurements with test instruments for evaluating coverage to visualize the impact of people in the vicinity of the base station and UE, changes to structures, etc., on radio-wave propagation that simulations are unable to determine.



Figure 8 Visualizing 5G Waveform using Test Instrument

5G operators commonly want to continuously monitor changes in the radio environment from moment-to-moment. As a result, we developed the Local 5G performance monitor (EX1001A) to visualize radio waves by combining the results of simulated radio-wave propagation technology and actual measurements using multiple probe terminals capturing data simultaneously at multiple test points. This supports visualization of radio waves as a heat map as well as confirmation of the status before and after faults using time-series graphs (Figure 9).

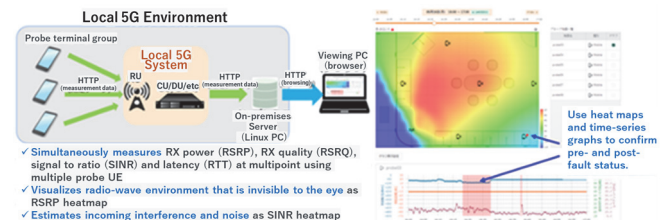


Figure 9 Performance Monitor Outline

The performance monitor is being tested in a 5G LAB environment to verify its usefulness and to improve its functions while discussing needs and issues with visiting customers.

3.2 Testing and Validating Visitors' Carry-in Systems

To stably use the system developed in the 5G environment, we have built an environment that can evaluate communications stability, improve dead zones, and evaluate communication performance.

(1) Evaluating Communications Stability

A customer system (real-time viewing system for in-vehicle camera images using local 5G) was experiencing noise in the video, and there was a consultation to solve the problem. As a result of comparing and verifying the measured values of each of the customer

system and 5G LAB, it was possible to analyze that the instantaneous transmission delay (spike delay) was the cause of the noise (Figure 10).

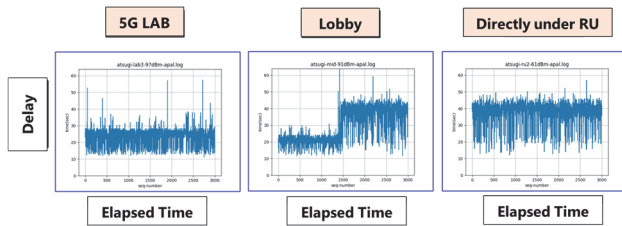


Figure 10 Comparison of Latency Measurements

(2) Improving Dead Zones

Reflectors are expected to guide radio waves in the highly linear 28 GHz band to areas that cannot be reached from the base station and were evaluated using the 5G LAB environment.

The 5G LAB evaluated an approximate 10 dB improvement in the strength of highly directional 28 GHz band radio waves by using a reflector plate to divert the signal from the base station into a dead-zone region created intentionally using a metal plate (Figure 11).

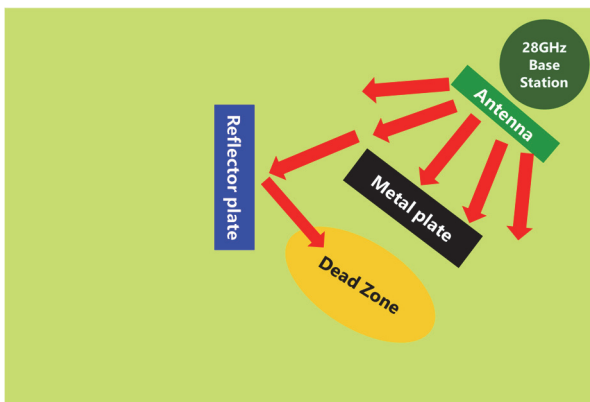


Figure 11 Reflector Evaluation Configuration

(3) Evaluating Communications Performance

5G LAB provides the required environment for measuring the communications performance, such as throughput and propagation latency of customers carried-in systems.

For customers testing introduction of Local 5G at their own factories, there can be issues with wireless propagation delays for industrial controller equipment, such as Programmable Logic Controllers (PLC). Consequently, 5G LAB examines the possibility of adopting Local 5G using the Monitoring Manager3

(NF7201A) and PureFlow (NF7501A) to monitor, collect, and analyze actual traffic data and measure propagation delay (Figure 12).

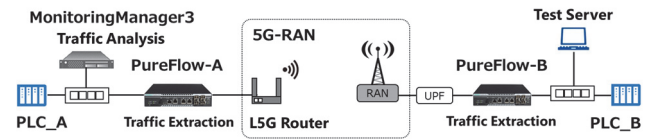


Figure 12 Typical Configuration for Measuring PLC Communications Delay

3.3 Proving Use Cases

5G LAB is a site for testing possible Local 5G use cases, such as control of aerial drones, a field with potential to change society.

When inspecting key social infrastructure, such as bridges and wind-power generators, drones must transmit video in real-time to a remote operator. Currently, video data is recorded to an onboard SD card, frequently requiring a return flyby if there is missing data.

By testing the usefulness of real-time video capture, 5G LAB helps clients obtain an experimental license to operate a drone with onboard Local 5G to advance examination of actual future operation.

4 5G LAB Usage Status

Since its 2021 opening, 5G LAB has been used by more than 300 companies from a wide range of fields, including Local 5G Sier, business handling service applications, and factories testing Local 5G introduction to promote understanding of potential solutions and building of collaborative relationship.

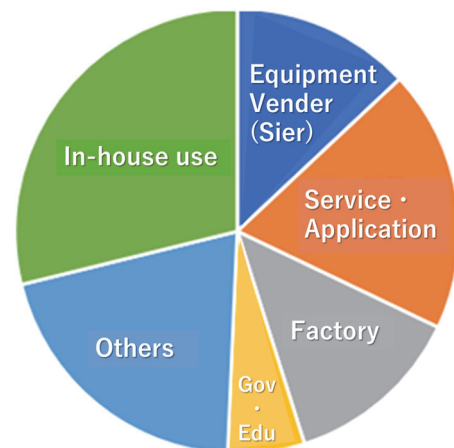


Figure 13 5G LAB Users

5G LAB is even used in-house by Anritsu. For example, it is used for various purposes, including development, and testing of Local 5G equipment; to prepare exhibition demonstrations; for in-house training and recruitment activities; for stockholder announcements; neighborhood school social studies visits, etc. In addition, 5G LAB helps understand and gather knowledge about the real challenges in configuring and operating Local 5G environments.

5 Collaborating with Partners

Due to worker shortages and work style reforms, how to reduce workloads and improve efficiency have become problems at construction sites and in factories.

As a result, there are high hopes for technologies to observe production status using high-definition video and low-latency video transmissions facilitating remote control of construction machinery as well as for AI technology to quickly ‘understand’ and analyze factory problems.

5.1 Partnering with Low-Latency Camera Technology

While 5G technology facilitates low-latency data transfer, real-time video streaming required for remote operation cannot be achieved using conventional cameras.

Konica Minolta Inc. (Konica) is researching the latest technologies for monitoring people’s safety and security using video and is in the process of developing low-latency cameras but faces challenges at the live-network evaluation stage.

To solve these issues, Anritsu’s 5G LAB in Kanagawa Prefecture and Konica’s 5G Open Laboratory in Osaka were network connected using Anritsu’s traffic analysis system to evaluate low-latency camera transmission performance. Moreover, mobile robots with built-in low-latency cameras have been remote-controlled over 5G to confirm the possibility of unmanned inspection work. Additionally, both Anritsu and Konica are working to understand how issues other than latency, such as jitter and packet loss degrade video quality, etc.

5.2 Partnering with AI Technology

Konica has AI technology for recognizing people and things from camera images and is exploring use cases to support on-site monitoring operations.

Food processing plants face many problems with labor shortages and worker monitoring to prevent accidents and contamination. To solve these issues with easy-to-

understand use cases, the 5G LAB is conducting a demonstration of detecting human falls and intrusions by skeletal detection that combines video surveillance and AI technology. Both Anritsu and Konica use the 5G LAB with these new technologies as a marketing test site.

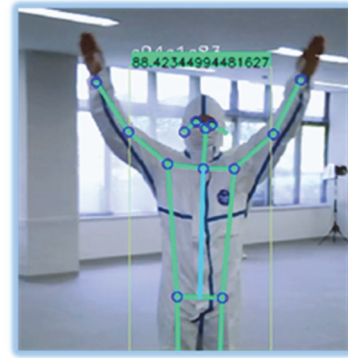


Figure 14 Frame Detection using AI Technology



Figure 15 Example of Fall Detection Warning

6 Conclusions

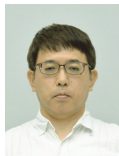
Opening the 5G LAB, obtaining a license, and building our own test environment taught us a lot about supporting businesses examining Local 5G introduction. Additionally, the 5G LAB proved its value in collaboration with partners with use-case demonstrations in a 5G environment using Konica’s leading technologies and as well as in partnering with customers aiming to create AI solutions.

Future plans are to create an environment not just for advanced-technology businesses but also to create value for other business types, such as power-related businesses working in carbon-neutral energy and local governments promoting infrastructure resilience during natural disasters.

Authors



Satoshi Sasaki
1st Sales Dept.
Sales Div.
Environmental Measurement Company



Koichi Ryu
2nd Business Development Dept.
Business Development Div.
Environmental Measurement Company



Genya Yatsu
2nd Business Development Dept.
Business Development Div.
Environmental Measurement Company



Koji Arimochi
3rd Business Development Dept.
Business Development Div.
Environmental Measurement Company



Makoto Fujii
3rd Business Development Dept.
Business Development Div.
Environmental Measurement Company



Yoshiaki Niwano
3rd Business Development Dept.
Business Development Div.
Environmental Measurement Company



Tsutomu Konishi
3rd Business Development Dept.
Business Development Div.
Environmental Measurement Company

Table 1 Local 5G Base Station Specifications

	4.8 GHz Base Station	28 GHz Base Stations	
System Type	5G SA	5G NSA	
Band	n79	B41	n257
Center Frequency	4849.98 MHz	2585 MHz	28249.92 MHz
Communication Technology	TDD (Synchronous)	TDD (Synchronous)	TDD (Synchronous)
Bandwidth	100 MHz	20 MHz	100 MHz
Antenna Power	4.58 W	0.312 W	0.126 W
TRx Systems	2	2	192 (V: 96 H: 96)
Antenna Type	Omni	Omni	Active Phased Array
Beamforming	Not supported	N/A	Supported
Modulation Method (Primary Modulation)	UL: QPSK 16QAM 64QAM DL: QPSK 16QAM 64QAM 256QAM	UL: QPSK 16QAM 64QAM DL: QPSK 16QAM 64QAM 256QAM	UL: QPSK 16QAM 64QAM DL: QPSK 16QAM 64QAM 256QAM
MIMO	UL: SISO DL: 2×2	UL: 2×2 DL: 2×2	UL: 2×2 DL: 2×2

Publicly available

O-RAN Radio Unit Test System for Promoting Open RAN

Koichi Inoue, Takaaki Kamisawa, Koji Fukazawa, Hiroyuki Tanikage, Hiroki Azuma, DongGeun Park

[Summary]

As the importance of Open Radio Access Networks (Open RAN) grows, and wireless access networks become more open, multi-vendor, and virtualized, interoperability issues due to multi-vendor equipment sourcing require test instruments to discover and resolve these issues. We developed a system to test the wireless performance of the O-RAN Radio Unit (O-RU) that transmits/receives wireless signals and to measure the digital data transmission/reception. This test system uses an intuitive Graphical User Interface (GUI) to flexibly set waveforms and offers automated test functions, including synchronization, which we hope will further support the spread of O-RAN.

1 Introduction

The O-RAN Alliance is an international organization promoting open standardization of the Radio Access Network (RAN) part of wireless networks. Open RAN (O-RAN) wireless networks are divided into three parts: the O-RAN Radio Unit (O-RU), O-RAN Distributed Unit (O-DU), and O-RAN Central Unit (O-CU), each with a published open standard interface that facilitates interoperability between equipment from different vendors. As well as resulting in reduced costs due to price competition between many different communications equipment makers, telecommunication companies can freely choose the equipment combination matching their requirements. On the other hand, non-compliance with the standard can cause various interoperability issues and identifying and solving the cause often takes time. Resolving interoperability issues requires tests to confirm that the equipment interface is standards-compliant.

This article describes development of an O-RAN test system with functions for confirming O-RAN operation and compliance with standards.

The system confirms the operation of each function at the C-Plane (control plane), U-Plane (user plane), and S-Plane (Synchronization plane) of the Fronthaul digital data communications section between the O-RU and O-DU and supports the O-RAN Alliance-defined Conformance Test for verifying protocol compliance between the O-DU and O-RU. Moreover, since the system complies with the 3rd Generation Partnership Project (3GPP) TS38.141 standard, it can also evaluate O-RU Radio Frequency (RF) characteristics.

The design concept, details, and features of this developed system are described below.

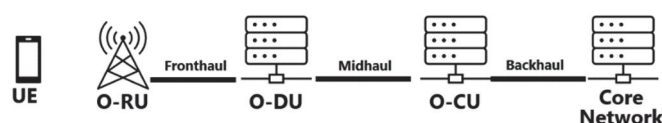


Figure 1 Radio Access Network Components

2 Development Concept

In developing the O-RAN Radio Unit test system, our aim was to build a system that would both test according to the O-RAN/3GPP standards but would also support easy changes test methods to meet the need for verification and non-standard testing. The design concepts are listed below.

(1) Easy Changes to Test Standard Parameters and Test Procedures

When a test result is Fail, the test parameters and procedure can be changed from the test standard to do troubleshooting and specify the issue. In particular, we developed software to support replacing the order, adding, and deleting operation sequence settings.

(2) Flexible Changes to Waveform Patterns

Part of system uses our company's IQproducer waveform generation software to create many different Downlink and Uplink waveforms.

(3) Intuitive GUI for Settings and Operation

GUI-based software was developed to support easy and intuitive operation, including test sequence editing. In particular, an easy-to-understand Tree View format display is used for M-Plane which can be difficult to set and edit without specialist skills. We aimed to create a GUI supporting easy editing by displaying parameter ranges determined by the standard.

(4) Supports Commercial and Existing Products

With respect to hardware, costs were reduced and development time was shortened by using commercial products and our company's existing testers.

3 Test System Configuration

We choose hardware and developed the software based on the above concepts. Figure 2 shows the O-RAN Radio Unit test system hardware configuration and Figure 3 shows the software packages.

(1) Supports MT8000A and MX800046A

The system uses the MX800046A software to measure the O-RU RF characteristics by operating the MT8000A as a base station using 3GPP TS38.141-compliant test item settings. IQproducer creates the Uplink waveform output from the MT8000A.

(2) O-DU Emulator Development

We developed an O-DU Emulator to connect with and test the Fronthaul, O-RU digital interface. Like an actual O-DU, we used a commercial PC server running the Linux operating system (OS) with an accelerator add-in-card (AIC) and developed a Field Programmable Gate Array (FPGA) for the accelerator card and System On a Chip (SoC) software to implement the C/U/S/M-Plane functions defined by the Fronthaul standard. The O-RU Fronthaul is connected via the SFP28 interface of the accelerator card.

(3) Adoption of MT1000A

The O-DU Emulator transmits U/C-Plane Downlink packets while synchronizing with the O-RU using S-Plane, and as a result, a Downlink waveform is output from the O-RU.

In order to perform RF measurement of the Downlink waveform with the MX800046A, clock synchronization between the O-DU Emulator and the MX800046A is required. Additionally, since accurate analysis of the Downlink waveform and output of the Uplink waveform requires a 1 pulse per second (PPS) signal, the MT1000A is used as the clock source. Combination with the MU100090B and rubidium oscillator in the MT1000A outputs the high accuracy 1PPS and 10 MHz synchronization signals required for RF measurement. Measurement is also possible using a clock synchronized to the Global Positioning System (GPS) signal.

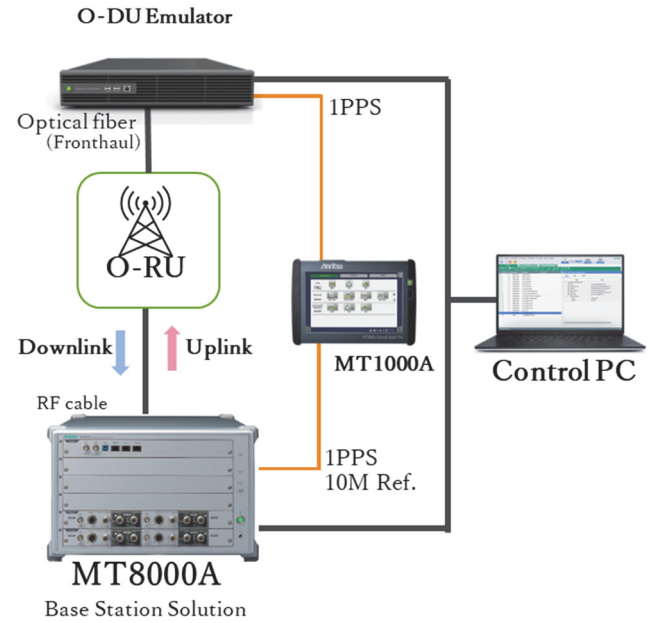


Figure 2 O-RAN Radio Unit Test System

(4) MX773000PC Development

The MX773000PC software running on a Windows PC controls the O-DU Emulator. The MX773000PC can operate and test the C/U/S/M-Plane on the O-DU Emulator with a single GUI. The Downlink waveform created by IQproducer is converted to packet data by the MX773000PC and Downlink tests are implemented by transmission to the O-RU. For Uplink testing, the MX773000PC has a function to capture packet data sent from the O-RU and display the analysis results.

An easy-to-understand Tree View style display makes it easy to edit M-Plane parameters.

(5) MX772000PC Development

To automate the O-RAN Alliance-defined Conformance Test and the 3GPP TS38.141 tests, we developed the MX772000PC software to control the MX773000PC and MX800046A automatically. By using the MX772000PC, the operator can execute tests by sequentially controlling the MX773000PC and MX800046A; the MX772000PC can also automatically done Pass/Fail evaluation at specified threshold values. Since sequences can be edited easily using the GUI, changes from the test standards as well as new tests are performed easily.

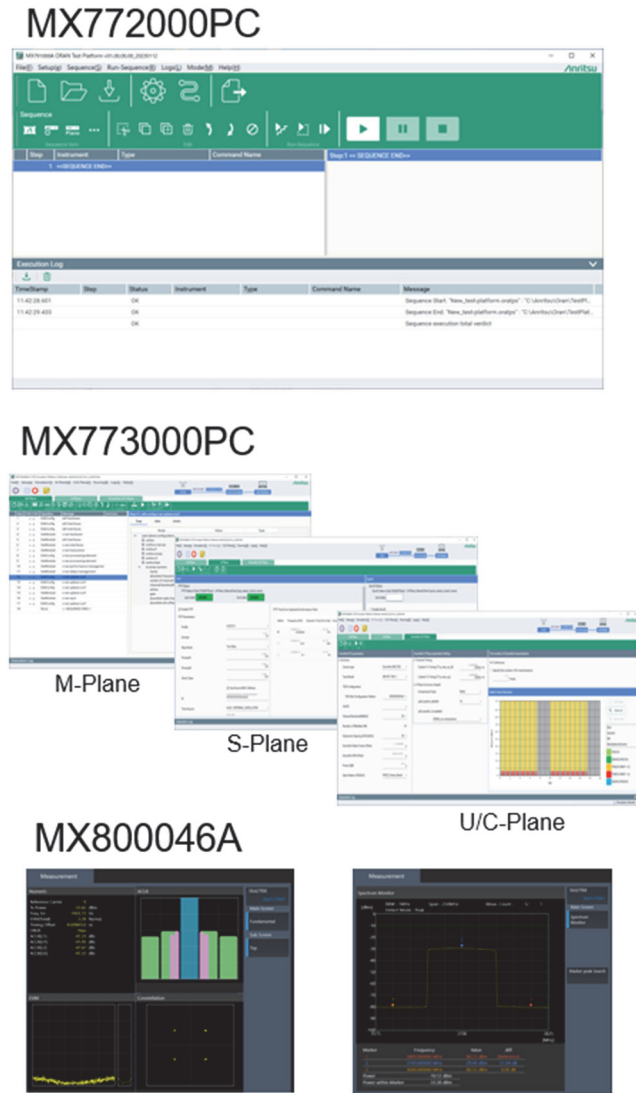


Figure 3 O-RAN Measurement Software

4 Design Features

This section explains the actual design requirements.

4.1 Hardware

4.1.1 Accelerator Card and Server PC

We adopted existing commercial products due to the advantages of lower hardware cost, shorter development time, expandability, and availability for the customers. To make use of these advantages, in choosing hardware, we focused on ease of globally obtaining compliant hardware with legally regulation as well as easy appropriation of design resources with general configurations.

The accelerator card required the following conditions.

- Support for both 25 and 10GbE network interface
- FPGA which can design internal circuits is installed in order to implement unique functions as an O-DU Emulator.
- Adequate FPGA resources supporting expandability

- Timing reference and accurate clock source, such as external 1PPS to implement high clock precision

The AMD (Xilinx) Xilinx T1 telecom accelerator card (T1 card hereafter) was used because it meets the above requirements.

A Dell PowerEdge R750 is used as the Server PC because it supports the two PCIe Gen3 x8 interface slots required by the T1 card, and also has a 25 GbE network card for M-Plane communications with the O-RU via the T1 card.

Other network interface than T1 card on the Server PC can also use to connect with the M-Plane to O-RU, so the T1 card is unnecessary when testing only the M-Plane. Also, if the O-RU uses a network interface different from the C/U/S-Plane for the M-Plane, it is possible to use the NETCONF (Network Configuration Protocol) Client other than the O-DU Emulator for the M-Plane.

4.1.2 FPGA

This section explains the FPGA functions.

The FPGA performs network communications with the O-RU. Figure 4 shows the FPGA internal data flow. The FPGA is composed of Processing System (PS) and Programmable Logic (PL). The PS performs S-Plane processing. When the M-Plane is communicating using the same network interface as the C/U/S-Plane, M-Plane packets are relayed with the Server PC. On the U/C-Plane packets transmission, it transmits the C-Plane and U-Plane data created by the control software with the evolved Common Public Radio Interface (eCPRI) protocol framing repeatedly.

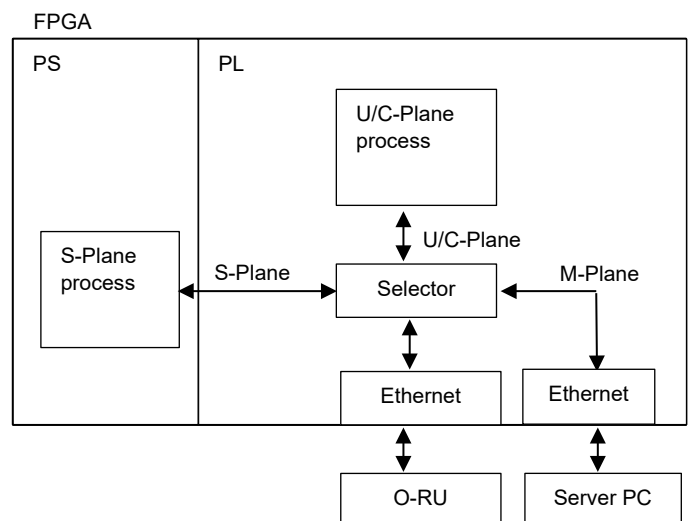


Figure 4 FPGA Data Flow

The 25 and 10GbE interfaces with the O-RU are switched

without changing the FPGA object. Use of Reed Solomon Forward Error Correction (RS-FEC) by the 25GbE interface can be enabled/disabled.

The FPGA PL generates the clock used by the Precision Time Protocol (PTP) as the timing management function. The packet send and receive times used by PTP are acquired and reported to the PS. At PTP One-Step operation, as shown in Figure 5, PL acquires the packet send time and insert it to the Sync message. At Two-Step operation, as shown in Figure 6, PS inserts the Sync message send time acquired by PL to the Follow_Up message. Acquiring the send time using PL increases the time accuracy.

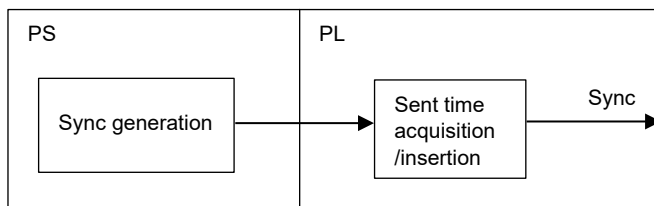


Figure 5 One-Step Operation Sync Message Tx Time Insertion

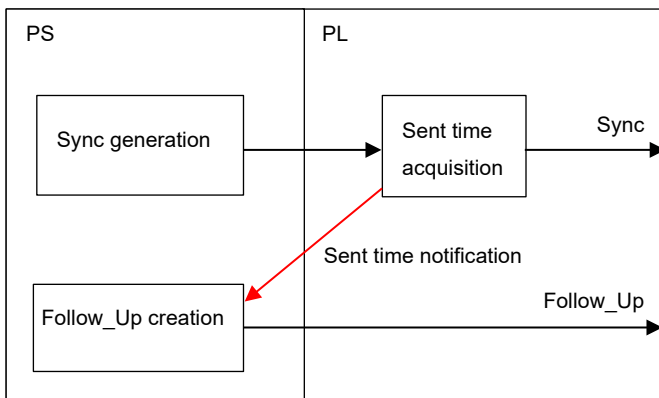


Figure 6 Two-Step Operation Follow_Up Message Tx Time Insertion

4.2 Software

4.2.1 SoC/Linux application

This section shows the SoC software and Linux application software blocks that are controlled from the MX773000PC software.

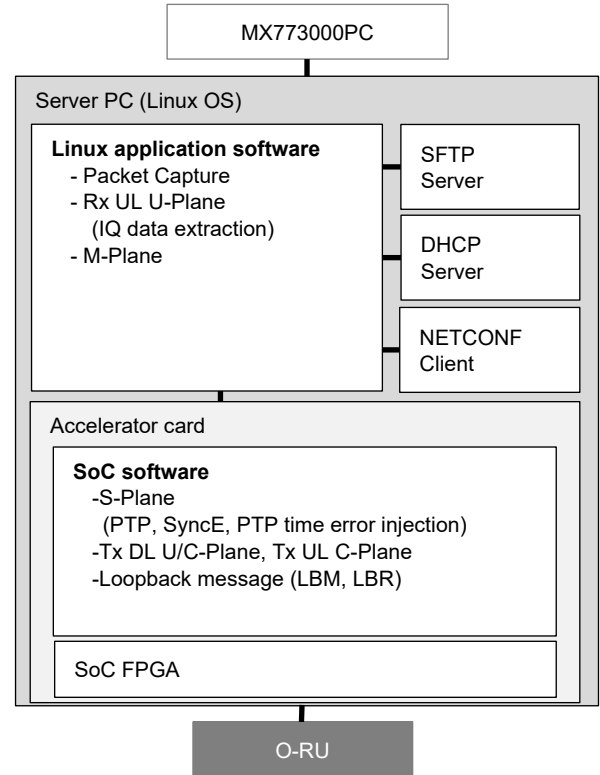


Figure 7 SoC and Linux Application

(1) SoC software

The SoC software is built into the T1 card to communicate with the C/U/S-Plane and test cases.

- S-Plane (PTP, SyncE (Synchronous Ethernet))

For the S-Plane, we developed a PTP Master function to manage the PTP time in synchronization with the external PPS signal input. To support the O-RAN Conformance test ITU-T G.8275.1 Profile described in section 3.3.3 O-RU Performance Test, we developed a function for adding errors to PTP. In addition, ESMC packet transmission is supported as SyncE. The Quality Level (QL) is included in the ESMC packets as an indicator of the SyncE quality. S-Plane status transitions are tested in accordance with the O-RAN Conformance Test specified in the ITU-T G.8275.1 profile in section 3.3.2 O-RU Functional Test. QL can be set manually to support this test.

- U/C-Plane Tx

The U/C-Plane transmission processing block controls the U/C-Plane packet transmission timing. To support the Delay Management test to verify the O-RU behavior when the packet arrive time is delayed relative to the current time, we implemented manual setting of the transmission timing.

- Loopback Message TRx

We have developed a function to transmit and receive a Loopback message for IEEE 802.1Q-2018 to support the O-RAN Conformance Test specified in section 3.1.13.1 Ethernet Connectivity Monitoring.

- (2) Linux application software

The Linux application software handles communication with the T1 card and M-Plane communications.

- Packet Capture

This function captures and reads packets via the T1 card memory. It outputs the contents of communications with the O-RU to a pcap file that can be read by other applications for efficient debugging.

- Uplink U-Plane IQ Data Extraction

This function has been developed to extract IQ data from Uplink U-Plane packets. Compressed IQ data is extracted automatically, and the extracted IQ data is transferred to the MX773000PC for demodulation analysis.

- NETCONF Client

The NETCONF Client processing block has been developed to supports the M-Plane communication with O-RU. To assure stable connections, NETCONF Client was implemented using the available Open-source Software (OSS) in widespread use. Combined use with the MX773000PC supports data communications based on Yet Another Next Generation (YANG) modules prescribed in the O-RAN specifications. As well as connecting using a static Internet Protocol (IP) address at the O-RU, it also supports connections using NETCONF Call Home used in combination with a Dynamic Host Configuration Protocol (DHCP) server.

4.2.2 MX773000PC

When connecting and testing a combination of O-RU and O-DU from different vendors, issues such as failed connections and different characteristics can occur, making it necessary to investigate the cause at both the O-RU and O-DU, which requires a lot of time to debug and isolate. To simplify this troubleshooting, we developed the MX773000PC O-DU Emulator Platform Software for testing the O-RU in accordance with O-RAN Alliance specifications.

Typical MX773000PC functions are listed below.

- (1) M-Plane Function

- Generating XML (Extensible Markup Language) from Tree

The M-Plane is a management function for the O-RU software, settings, etc., which are implemented using the NETCONF communications protocol. To create the NETCONF XML it is necessary to understand the YANG file format for each YANG module in the O-RAN Alliance specification and to describe the XML in accordance with the YANG file definition, but understanding and describing are very labor intensive. Consequently, the MX773000PC has a built-in function that converts the YANG file format to a Tree View where parameters on the Tree are set to easily create the XML. To improve the readability of NETCONF data expressed in XML format, the Tree also indicates the original parameters hierarchically.

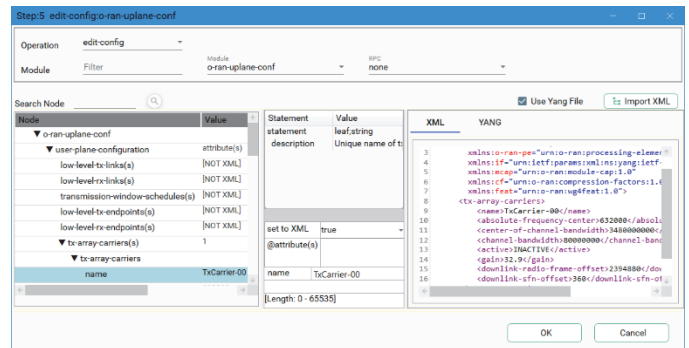


Figure 8 Generating XML from M-Plane Tree

The relationship between the YANG file, Tree and XML is shown below.

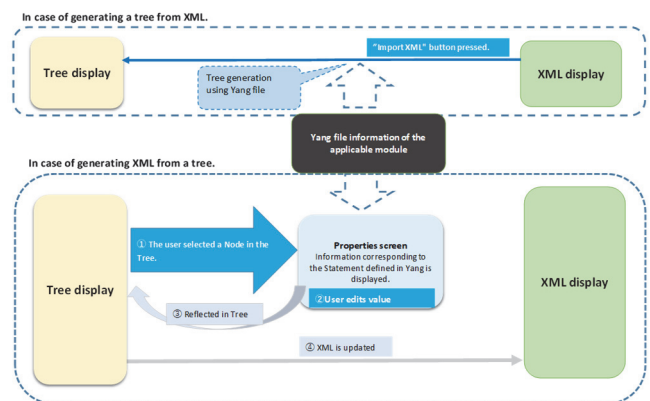


Figure 9 Relationship between YANG File, Tree, and XML

- XML Analysis Function

This built-in function reflects XML acquired from the O-RU as well as from the user's existing XML in the Tree. Acquiring the XML by converting from logs when using the O-DU and O-RU in a live environment supports M-Plane reproducibility tests.

Additionally, the locations of XML description (syntax)

errors and XML parameter settings that do not match the YANG for the specified module are displayed to discover setting mistakes and trouble spots.

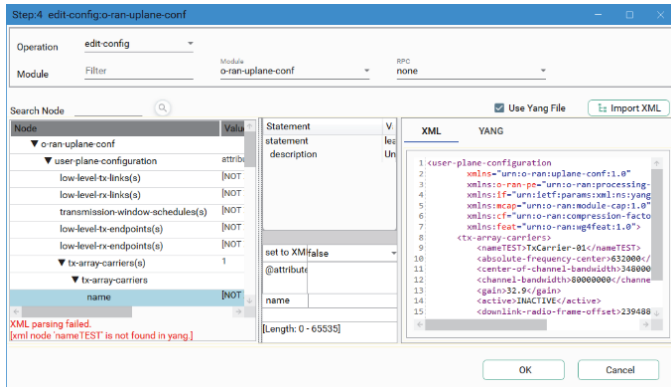


Figure 10 Checking YANG File at Tree Generation from XML

• M-Plane Message Sequential Execution

In addition to the XML addition and editing functions, the MX773000PC has other functions to set the wait time at each step, display user message dialogs, record execution logs, and add/edit comments.

These functions can be executed sequentially at each step.

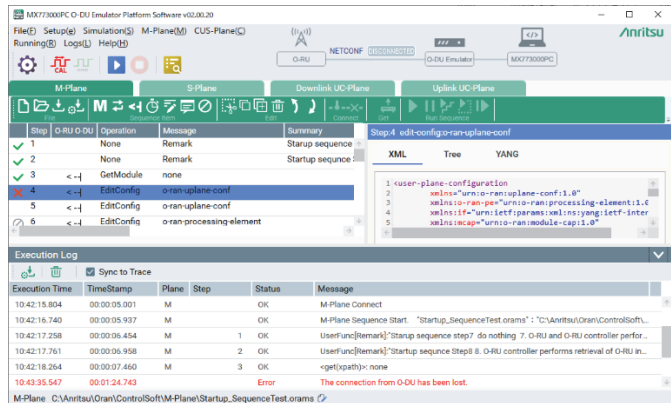


Figure 11 M-Plane NETCONF xml Sequential Execution

(2) U/C-Plane Function

3GPP TS38.141-compliant waveforms must be created to test the O-RAN C- and U-Planes. The MX773000PC has functions cooperating with IQproducer of existing product to create flexible waveform patterns and then for Downlink/Uplink U/C-Plane transmission/reception.

• Downlink Function

Using this function, the MX773000PC can transmit any Downlink U/C-Plane packet to the O-RU. The waveform is generated by using IQproducer and the generated waveform data is installed in the Server PC. Installed waveforms are listed in the GUI to switch easily to the waveform to output by selecting the waveform data from the list.

Additionally, the selected waveform is also displayed as a graph on the GUI to set and confirm the waveform Radio Frame Structure.

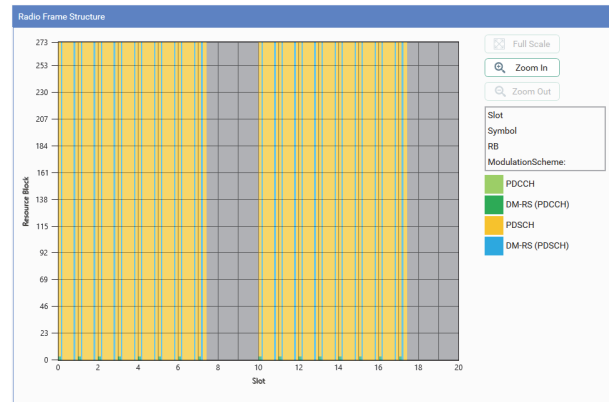


Figure 12 Downlink Waveform Display

• Uplink Function

Using this function, the MX773000PC can transmit any Uplink C-Plane packet to the O-RU and receive the Uplink U-Plane from the O-RU. Additionally, there is a function for analyzing the received U-Plane and displaying results as a graph.

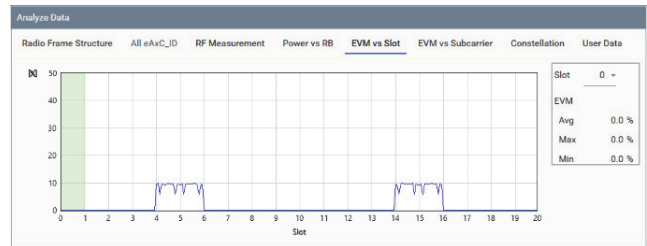


Figure 13 Uplink Analysis Results Display

(3) S-Plane Function

While mobile networks require precision time synchronization, because multi-stage switches and other transmission equipment in the transmission path cause delay. Using this function, the MX773000PC simulates the latency state in the transmission path by adding Time Error to the S-Plane and has a built-in function for measuring the C-Plane and U-Plane.

PTP Time Error Injection(Conformance Test)			
Select	Frequency [Hz]	Dynamic Time Error [ns]	Constant Time Error [ns]
☒	0.003362 0.1 0.003362	24.710 710	710.1396 710
☐	0.003362 0.1 0.01	24.710 239	710.1396 1181
☐	0.003362 0.1 0.1	24.710 24	710.1396 1396

Figure 14 Example of Time Error Input Parameters

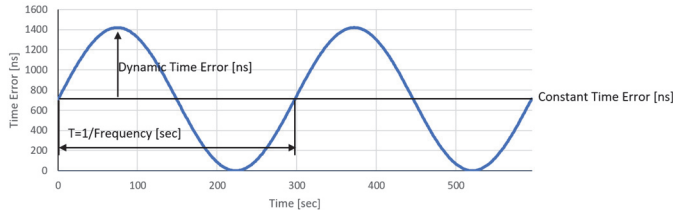


Figure 15 Time Error Injection Graph

4.2.3 MX772000PC

The O-RAN Alliance Working Group 4 (WG4) defines the O-RAN Conformance Test for the evaluating the O-RU, while 3GPP defines the TS36.141 and TS38.141 standards.

We developed the MX772000PC ORAN Test Platform to automate these tests and support continuous testing and pass/fail evaluations.

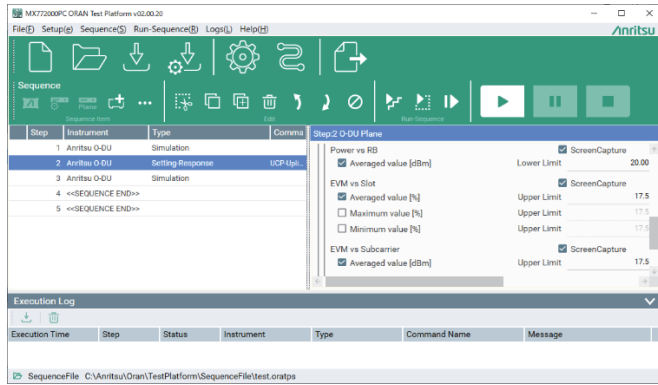


Figure 16 Sequential Execution of Pass/Fail Evaluation for Each Tester Settings and Results

The MX772000PC software controls multiple software, such as the MX773000PC and MX800046A, as well as the hardware to execute step settings line-by-line and the M/S/U/C-Plane tests sequentially. As well as controlling MX773000PC and MX800046A, it also has functions to set the wait time at each step, display user dialogs, record execution logs, add step comments, and repeat any step, and add/edit settings. These functions can be combined flexibly into an automated sequence supporting near-to-live environment testing.

The pass/fail sequence execution result for each step is displayed. In addition, pass/fail evaluation results are displayed for the entire sequence after execution of the whole sequence is finished.

- M-Plane Test

This function controls the MX773000PC and sets the O-RU M-Plane. It has a built-in pass/fail function to evaluate whether the setting result is the expected value.

- Downlink U/C-Plane Test

This function controls the MX773000PC and transmits U/C-Plane packets to the O-RU. It also controls the MX800046A, and acquires and evaluates analysis results for the Downlink signal output from the O-RU. Additionally, as well as evaluating one measurement result, it also has a function to evaluate multiple measurement results during a specified period.

- Uplink U/C-Plane Test

This function controls the MX800046A and transmits the Uplink signal to the O-RU. Additionally, it also controls the MX773000PC and transmits the Uplink C-Plane to the O-RU. It also has a built-in function to evaluate acquired analysis results after receiving Uplink U-Plane packets from the O-RU and performing signal analysis at the MX773000PC.

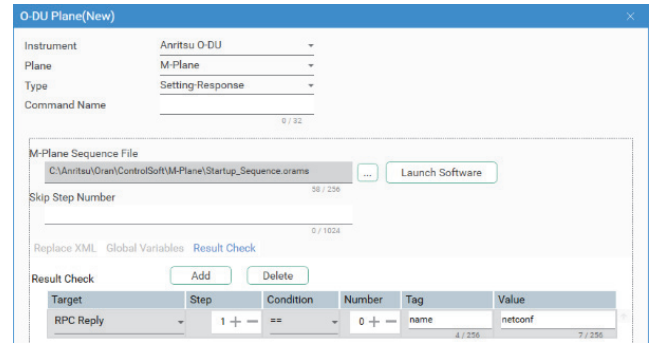


Figure 17 MX773000PC Control (M-Plane Pass/Fail Condition Settings)

- S-Plane Function

This function controls the MX773000PC and performs O-DU and O-RU PTP and SyncE synchronization functions as well as adds Time Error to PTP for testing.

4.2.4 MX800046A/IQproducer

This software creates the waveform data using IQproducer at Downlink/Uplink testing and transmits and receives RF signals. Although there have been base-station manufacturing solutions prior to this development, the Downlink/Uplink test in the O-RAN Conformance Test measures bit errors (mostly for PN23 patterns) on the Physical Downlink Shared Channel (PDSCH)/Physical Uplink Shared Channel (PUSCH), so this measurement item was to be added to the functionality.

When receiving the Downlink and executing RF analysis at the MX800046A, the PDSCH data is analyzed simultaneously after De-Scrambling. Moreover, analysis settings and expected measurement values can be generated

automatically from Downlink waveform information created by IQproducer when creating the waveform.

The expected PDSCH data values generated when the waveform is created as well as the MX800046A analysis results are acquired by the MX772000PC to calculate and evaluate the bit error rate and count.

5 Conclusion

We developed the O-RAN Radio Unit Test System for the growing Open RAN market.

This automated test system uses an intuitive GUI to confirm that O-RAN Radio Unit operation complies with the 3GPP TS38.141 and O-RAN Conformance Test items.

Future development is focused on test solutions to resolve equipment interoperability issues between Open RAN equipment, which we hope will help spread and grow Open RAN.

References

- 1) 3GPP TS38.104 NR Base Station (BS) radio transmission and reception
- 2) 3GPP TS38.141-1 NR Base Station (BS) conformance testing Part 1: Conducted conformance testing
- 3) 3GPP TS38.141-2 NR Base Station (BS) conformance testing Part 2: Radiated conformance testing
- 4) O-RAN Alliance e.V. O-RAN Working Group 4 (Open Fronthaul Interfaces WG) Control, User and Synchronization Plane Specification Management Plane Specification
- 5) O-RAN Alliance e.V. O-RAN Working Group 4 (Fronthaul Working Group) Conformance Test Specification
- 6) O-RAN Alliance e.V. O-RAN Working Group 4 (Open Fronthaul Interfaces WG) Management Plane Specification
- 7) Anritsu Corporation. O-RAN Radio Unit Test Solution product introduction (web), Nov 28 2023. <https://www.anritsu.com/en-us/test-measurement/products/mx772000pc>
- 8) Xilinx Inc. T1 Telco Accelerator Card Data Sheet (DS999)
- 9) Xilinx Inc. T1 Telco Accelerator Card User Guide (UG1495)

AMD and Xilinx are trademarks of Advanced Micro Devices, Inc.

PCIe is trademark of PCI-SIG.

Dell and PowerEdge are trademarks of Dell Inc. or its subsidiary.

Authors



Koichi Inoue
2nd Team
Product Development Dept.
IoT Test Solutions Div.
Test & Measurement Company



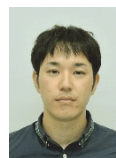
Takaaki Kamisawa
2nd Team
Product Development Dept.
IoT Test Solutions Div.
Test & Measurement Company



Koji Fukazawa
2nd Team
Product Development Dept.
IoT Test Solutions Div.
Test & Measurement Company



Hiroyuki Tanikage
2nd Team
Product Development Dept.
IoT Test Solutions Div.
Test & Measurement Company



Hiroki Azuma
2nd Team
Product Development Dept.
IoT Test Solutions Div.
Test & Measurement Company



DongGeun Park
Application Engineering Dept.
Engineering Div.

Table 1 Key MX773000PC Specifications

Communication Functions	U/C-Plane	Frequency Range	FR1 (Frequency Range 1)
		Bandwidth	20, 40, 60, 80, 100 MHz
		Subcarrier Spacing (SCS)	15, 30 kHz (15 kHz only at 20 MHz bandwidth)
		No. of Layers	1
		No. of Component Carriers (CC)	1
		No. of Controlled RF Ports	1
		Supported Protocols	VLAN (Virtual Local Area Network) *Multi-tags not supported O-RAN Split Option 7-2x C-Plane for Downlink Scheduling U-Plane for Downlink IQ U-Plane for Uplink IQ (Decoding not supported) *IP/UDP (User Datagram Protocol) not supported *Supports eCPRI as Transport Header but RoE (Radio over Ethernet) not supported
		Compression Algorithm	Block Floating Point (BFP) 9/12/14 bit
	M-Plane	Supported Protocols	VLAN *Multi-tags not supported Netconf Client IPv4, IPv6 DHCPv4, DHCPv6, Static IP SSHv2 (Secure Shell 2.0), TLS (Transport Layer Security) 1.2
	S-Plane	Supported Protocols	PTPv2, SyncE (ESMC) PTP Profile: G.8257.1
Measurement Functions	U/C-Plane	—	Uplink: TX Power, EVM (Error Vector Magnitude), Frequency Error, Origin Offset, Timing Offset, Descramble results display, EVM/Power/Constellation graph display Downlink: Tx function for IQproducer-created waveform data, Function for regenerating U/C Plane packets saved in pcap
	M-Plane	—	Edit any packet sequence and set any parameters Displays O-RU response to executed sequence Displays M-Plane connection status
	S-Plane	—	Functions for changing PTP and SyncE parameters Function for confirming PTP and SyncE sync status via M-Plane Changeable PTP parameters: Domain, Step Mode, Priority#1/#2, Clock Class, ID, Time Source, Clock Accuracy, Announce Interval, UTC Offset, Sync Interval, Minimum Delay Request Interval Changeable SyncE parameters: QL, Source MAC Address, Event Flag
		Time Error Injection Function	Function for adding time error defined in item 3.3.3 of O-RAN Conformance Test (O-RAN.WG4.CONF.0) to PTP time
		PTP Correction Function	Function for correcting PTP update timing time error in input 1PPS signal and at O-DU Emulator
		Frequency Correction Function	Function for correcting input PPS signal and Ethernet frequency error
	Others		Functions for capturing Ethernet packets and outputting to pcap file

Publicly available

Development of Wideband Linear Amp with New InP HBT

Tomoaki Ishii, Yukio Ohkubo, Masaya Suzuki, Takumi Nemoto, Yutaka Kobayashi, Takashi Kamizono, Moeko Jimba

[Summary]

Faster data transmission requires higher performance devices. Until now, Anritsu has used $f_T/f_{\max}=220/350$ GHz Indium Phosphide (InP) Heterojunction Bipolar Transistor (HBT) designs in its test instruments. This article reports the development of a new high-performance ($f_T/f_{\max}=323/510$ GHz) HBT by optimizing the structure of the IC layers and processing technology. This newly fabricated HBT was used in a high-output, high-speed wideband linear amp with good 120 Gbaud pulse response characteristics. It was possible to amplify input signal of 0.3 V_{pp} up to 2.0 V_{pp}. Future work will focus on designing even faster ICs using the developed HBT.

1 Introduction

Required network transmission capacity is increasing rapidly due to the advancements of the information society as represented by the growth of cloud computing. Device technologies to achieve high transmission speeds exceeding 100 Gbps both between and within data centers are becoming common, and standardization discussions are shifting from 400 to 800 Gbps. Research and development is advancing into implementing high-speed communications exceeding 100 Gbps even in the 6G wireless communications world, and there is also a strong demand for technologies supporting increased data speeds, such as Peripheral Component Interconnect Express (PCIe). Innovative applications, such as Artificial Intelligence (AI) and Machine Learning (ML), are becoming commonplace, increasing the demand for faster transmissions. As a result, test equipment must also support more standards. In particular, devices used in test equipment must meet the latest high-speed transmission communication standards, and always exposed to global competitive environment.

Anritsu Devices has Monolithic Microwave Integrated Circuit (MMIC) technology that uses in-house InP based heterojunction bipolar transistors (HBT) and has been working hard to improve device technologies with the aim of creating a 58 Gbaud Pulse Amplitude Modulation (PAM4) solution and develop measurement instrument supporting the latest PCIe standards.

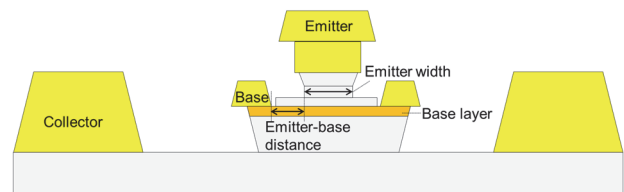
This article explains both the layer structure of transistors forming the heart of the improved performance of second generation (Gen2) InP HBT and the fabrication process technology. In addition, it also explains the MMIC design and

packaging technologies for the implementing bandwidth exceeding 100 GHz in our 120 Gbaud (equivalent to 240 Gbps PAM4) AH15190A Linear Amplifier using Gen2 HBT.

2 High-Performance InP HBT and MMIC

(1) Structure and Performance of High-Speed HBT

(a)



(b)

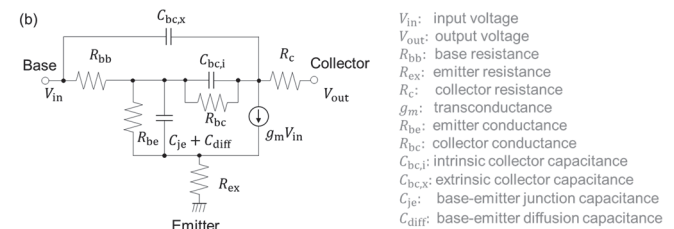


Figure 1 (a) HBT Cross Section (b) HBT Equivalent Circuit

The typical indices of HBT high-frequency performance are the cutoff frequency f_T , and the maximum oscillation frequency $f_{\max}$. In high-frequency amplifier ICs, it is said that assuring amplification performance requires an $f_{\max}$ of two or three times the IC modulation frequency and is a key index for increased speed. Figure 1 (a) and (b) show the HBT structure and equivalent circuit, respectively. The values of f_T and $f_{\max}$ are defined¹⁾ by the following equations (1) and (2), respectively.

$$f_T = \frac{1}{2\pi} \left(\tau_B + \tau_C + \frac{n_c kT}{qI_C} (C_{je} + C_{bc}) + (R_{ex} + R_c) C_{bc} \right)^{-1} \quad (1)$$

$$f_{\max} = \sqrt{\frac{f_T}{8\pi R_{bb} C_{bc,i}}} \quad (2)$$

where, τ_B is the base transit time, τ_C is the collector transit time, C_{je} is the base emitter joint capacitance, n_c is the ideal collector current coefficient, k is the Boltzmann constant, T is temperature, q is the elementary charge, I_C is the collector current, C_{bc} is the base collector capacitance, R_{ex} is the external emitter resistance, R_c is the collector resistance, R_{bb} is the base resistance, and $C_{bc,i}$ is the base collector internal capacitance contributing to $f_{\max}$. In addition, $C_{bc,x}$ (Figure 1 (b)) is the base collector external capacitance not contributing to $f_{\max}$.

Thinning the base and collector layers to shorten the electron transit times τ_B and τ_C is an effective method for increasing f_T . On the other hand, $f_{\max}$ is decreased because thinning the base layer increases the base resistance. The base resistance is expressed as Eq. (3) as the sum of the contact resistance between the metal base and semiconductor, the semiconductor resistance between the base and emitter, and the diffusion resistance immediately under the emitter layer.

$$R_{bb} \cong \frac{\rho_{Bs}}{2 L_E} \left(L_t + W_{Bg} + \frac{W_E}{6} \right) \quad (3)$$

where, ρ_{Bs} is the base sheet resistance, L_E is the emitter length, L_t is the transfer length, W_{Bg} is the distance between the emitter and base, and W_E is the emitter width. Effective methods for reducing the base resistance are increasing the doping density of the base layer to lower the base sheet resistance (ρ_{Bs}) and shortening the distance between the base and emitter (W_{Bg}).

Since the emitter current time (item 3 in Eq. (1)) becomes smaller as the collector current increases, f_T increases. However, when many electrons are injected, the potential gradient felt by the electrons traveling through the collector layer decreases due to the Kirk effect, which makes it difficult for the current to flow, causing f_T to decrease. Therefore, f_T has a peak at a certain collector current.

Since reducing the emitter thickness reduces the internal capacitance of the base collector as well as the contact capacitance with the base emitter, it is effective in increasing f_T . However, only reducing the emitter thickness lengthens the distance between the emitter and base, which increases the base resistance. As shown in Eq. (2), increasing the base resistance reduces $f_{\max}$. Increasing $f_{\max}$ not only requires

increasing f_T but also requires decreasing the base resistance and the base collector internal capacitance.

(2) MMIC Technology and Development Aims

Since 2003, Anritsu has developed MMIC gallium arsenide (GaAs) and InP HBT to use in its test equipment^{2), 3)}. The three-layer wiring process enables high integration. Additionally, ground (GND) is strengthened, and high-frequency IC operation is stabilized by forming via through holes. So far, we have been making prototypes of DA converters using first generation (Gen1) InP HBT devices, but limitations in the operating speed have become apparent with just making integration higher. Improvement in transistor performance is becoming essential not only in digital circuits but also in analog circuits such as linear amplifiers. For this reason, we started developing high-performance HBT elements that would be useful for both digital and analog applications. IC operating speed is limited to about 100 Gbps when using Gen1 HBT (Figure. 2). 6G communications require a minimum speed of 100 Gbps, so test instruments must operate at speeds higher than 100 Gbps. To implement operating speeds of at least 160 Gbps, well beyond the 100 Gbps limit, we developed a Gen2 HBT process targeting $f_T > 300$ GHz and $f_{\max} > 450$ GHz. The Gen2 HBT targets are summarized below.

- High speed ($f_T > 300$ GHz, $f_{\max} > 450$ GHz) and high durability (>5 V)
- High-speed operation at practical current density ($< \text{several mA/cm}^2$)
- Practical long-term reliability ($\geq \text{several decades}$)

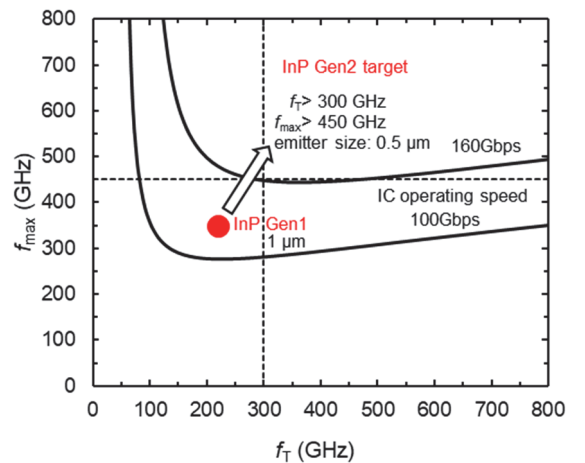


Figure 2 Relationship Between IC Operation Speed and $f_T/f_{\max}$

3 Layer Structural Design and Fabrication Process

To improve f_T , we aimed to shorten the electron transit

times (τ_B and τ_C) by optimizing the structure of the base and collector layers. We designed the collector structure to suppress the potential modulation of the band structure as the current density doubles when the decrease in f_T due to the Kirk effect occurs. To reduce base resistance, the p^+ -InGaAs base-layer doping density was increased from $4 \times 10^{19} \text{ cm}^{-3}$ to $6 \times 10^{19} \text{ cm}^{-3}$ and the thickness was reduced from 50 nm to 30 nm. From the viewpoint of changes in electron mobility⁴⁾, we expected a 28% decrease in the base sheet resistance. Additionally, we planned to reduce base resistance between the base and emitter by minimizing the emitter-to-base separation distance (minimizing W_{Bg}).

We used the emitter-mesa-formation dry-etching process for processing emitter tungsten silicide (WSi) electrodes, and the wet etching process for the semiconductor layer. Minimizing the emitter-to-base separation distance was important to improve performance, so it was necessary to form a narrow mesa structure for the emitter etching process. We formed $\sim 0.5 \mu\text{m}$ width emitter. On the other hand, large side etching (undercut) can lead to an increase in the emitter-to-base separation distance. We reduced the undercut amount and optimized the etching conditions. For dry etching, we adopted Inductively Coupled Plasma (ICP)-Reactive Ion Etching (RIE) offering new high anisotropy. For the wet etching process, we optimized the etching time to form a $\sim 0.5 \mu\text{m}$ wide emitter. From the Scanning Electron Microscope (SEM) image (Figure 3) of the fabricated Gen2 HBT, we can clearly see a transistor with our target emitter width of $\sim 0.5 \mu\text{m}$ and the emitter-to-base separation distance of $0.25 \mu\text{m}$.

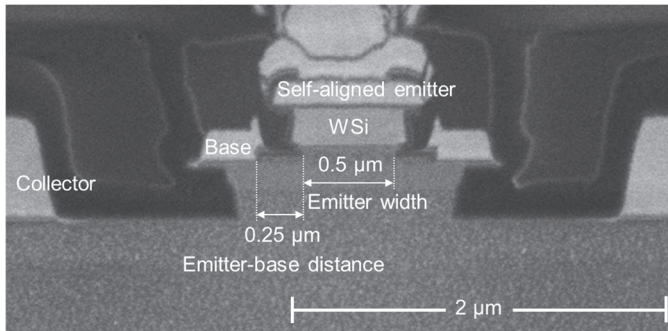


Figure 3 Gen2 HBT Cross-Section SEM

Assuring reliability requires controlling recombination current at the surface of the high-density base layer. To assure long-term reliability for the Gen2 HBT, we used the common guard ring structure (ledge layer). To prevent exposure of the base layer surface, we used n-InP emitter layer

with a depleted doping density of $3 \times 10^{17} \text{ cm}^{-3}$ on the upper part of the base-layer as the ledge layer.

4 HBT Features

The direct current (DC) and frequency characteristics of the prototype Gen2 HBT element with an emitter size of $0.5 \times 5 \mu\text{m}^2$ were measured and the results were used to extract the equivalent circuit parameters for circuit design. Figure 4 shows the measured static characteristics. The DC current amplification factor averaged about 57 in the base current (I_b) range of 50 to 350 μA . Although an HBT current amplification factor of about 30 has been reported⁵⁾, it is suitable for high-output amplifier applications since this fabricated IC has a relatively high current amplification factor. Additionally, it also has a high collector-to-emitter withstand voltage satisfying the minimum target value of 6 V.

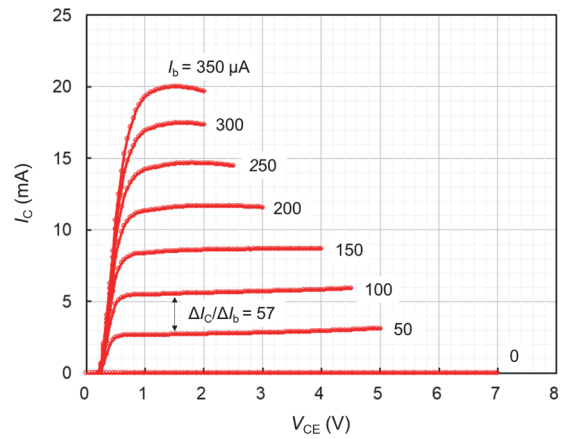


Figure 4 Static Characteristics

The S-parameter was measured using Anritsu's MS4647B Vector Star to evaluate high-frequency characteristics. The S-parameter measured at each bias was used to derive H_{21} and the Unilateral power Gain (UG). Figure 5 shows the H_{21} and UG frequency dependency at the f_T peak bias (1.4 V collector voltage, and 10 mA collector current). The f_T/f_{max} values of 323/510 GHz were extracted by the least-squares method. These values satisfy the target f_T/f_{max} values of $>300/450$ GHz. In addition, high-accuracy measurement is performed using the iterative de-embedding method⁶⁾.

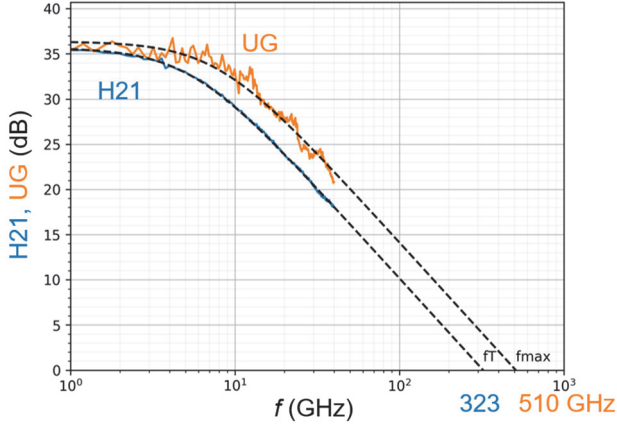
Figure 5 H₂₁ and UG Frequency Response at 1.4 V, 10 mA

Figure 6 shows measured f_T and f_{max} collector current dependency (○ plot) at a collector voltage of 1.4 V. The f_T value peaks (323 MHz) at a collector current of 10 mA. The current density at an emitter size of $0.5 \times 5 \mu\text{m}^2$ is 400 kA/cm², which satisfies the target value of less than several mA/cm². The capacitance and resistance were derived⁷⁾ from $-\text{imag}(Y_{12})/2\pi f$ and from real ($Z_{11}-Z_{12}$) for the base-to-collector capacitance C_{bc} (green) and base impedance Z_b (brown), respectively. In the created model (solid line), parameters were adjusted so that the base-collector capacitance C_{bc} (brown) and base impedance Z_b (green) were close to the measured values, making it a model suitable for high-frequency circuit design. The base resistance R_{bb} in the Gen2 HBT model using the extracted parameters was 15Ω , which was about 35% less than that in the Gen1 HBT model. From Eq. (3), the base resistance is found as a product of the base sheet resistance and size effect. As previously described, base sheet resistance changes according to the layer structure and is 28% lower in Gen2 HBT than in Gen1 HBT. To examine the size effect, we substituted $L_t=0.2 \mu\text{m}$, with $W_{Bg}=0.35 \mu\text{m}$ and $W_E=1 \mu\text{m}$ into Eq. (3) for Gen1 HBT, and $L_t=0.2 \mu\text{m}$, with $W_{Bg}=0.25 \mu\text{m}$ and $W_E=0.5 \mu\text{m}$ for Gen2 HBT, and found a 11% drop in base resistance. Therefore, we expect a total drop in base resistance of 36%. This is close to the approximate 35% resistance drop obtained by the model. Based on the above, we think that shortening the emitter-to-base distance and increasing the base-layer doping density explains the lower base resistance.

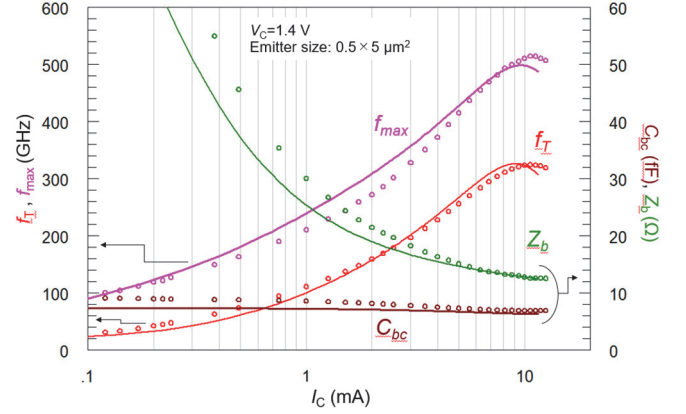
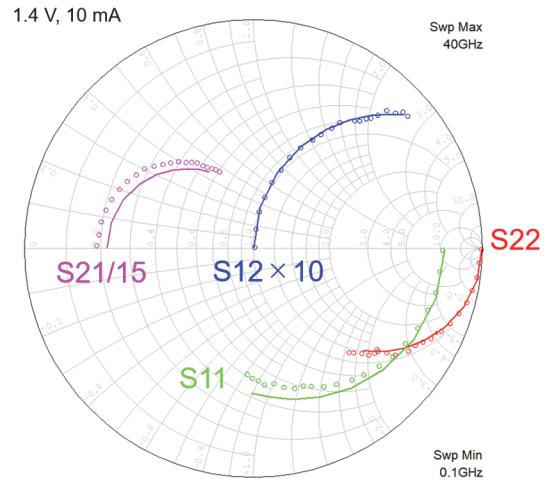
Figure 6 f_T , f_{max} , C_{bc} , Z_b Collector Current Dependency Test Results (Plot) and Model (Solid Line)

Figure 7 shows the measured S-parameters (plots) at a collector voltage of 1.4 V and collector current of 10 mA. For modeling, we adjusted the high-frequency characteristics parameters based on parameters obtained from the DC characteristics to obtain a model (solid line) closely approximating the actual measurements (plot).

With respect to long-term reliability, using a ledge structure confirmed a life of several decades (≈ 70 years) more than the design target. Based on the above, we have developed a Gen2 HBT meeting our targets.

Figure 7 S-parameter Test Results (○ Plot) and Model (Solid Line) S_{12} plotted at 10 times and S_{21} plotted at 1/15 times

5 Wideband Linear Amp Design

(1) Linear Amp Development Target

With the recent growth of cloud and mobile applications, IP network traffic continues to grow, and IEEE is examining introduction of a new Ethernet 800GbE/1.6TbE standard to support this growth. As a result, PAM4 signaling at speeds above 100 Gbaud will become mainstream. A 2.0 Vpp (single-

end) 112 Gbaud PAM4 signal is required by the small Transmitter Optical Sub-Assembly (TOSA) with a 200 Gbps per lane Electro-absorption Modulator integrated Laser diode (EML) using Intensity Modulation/Direct Detection (IM/DD) for 800GbE/1.6TbE signaling; evaluation of the Coherent Optical Sub-Assembly (COSA) optical module for digital coherent communications longer than 80 km using the 800ZR requires a 2.0 Vpp (single-end) 118 Gbaud or faster PAM4 signal. To meet these needs, we developed a wideband linear amplifier using the developed Gen2 HBT supporting 120 Gbaud PAM4 signaling.

(2) Linear Amp Design

In designing the 120 Gbaud wideband linear amplifier, as well as targeting a wideband of at least 100 Gbaud we also targeted (Table 1) output of a 2.0 Vpp high-amplitude signal.

Table 1 Linear Amplifier Target Specifications

	Target Specification
Max. Operating Baud Rate	120 Gbaud typ.
Output Amplitude	2.0 Vpp typ.
Min. Signal Gain	16 dB typ.
-3 dB band	>100 GHz

We used the Gen2 HBT development process. To support pulse responses to the high 120 Gbaud rate, the IC design used a distributed amplifier circuit for wideband flat gain. Moreover, when improving the accuracy of simulation beyond 100 GHz, it helps minimize degraded characteristics due to packaging.

Figure 8 shows the small signal evaluation results for the linear amplifier standalone IC measured using Anritsu's ME7838D VectorStar. The small-signal gain was 17 dB, and the -3 dB gain bandwidth was 113 GHz.

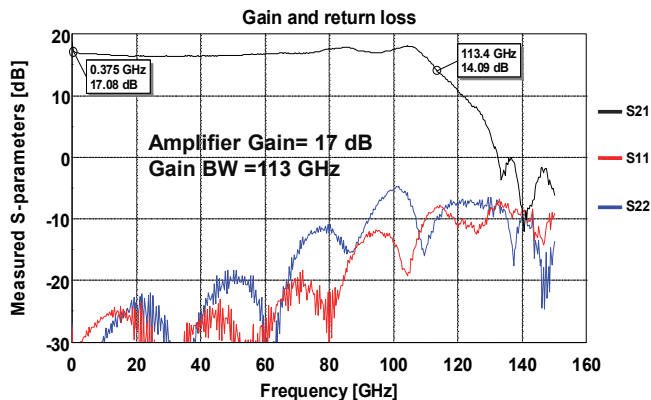


Figure 8 Linear Amp IC Minimum Signal Characteristics Evaluation Results

We then designed the packaged linear amp IC module (Figure 9).

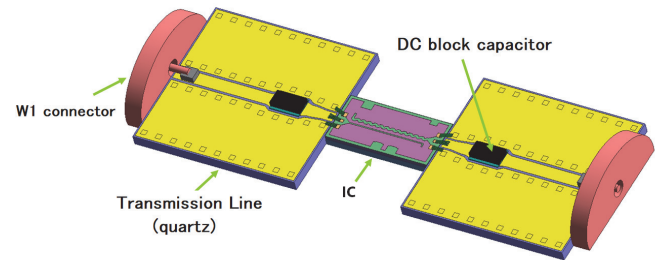


Figure 9 Designed Linear Amp Module Package

The linear amp module I/O uses W-connectors. Quartz board with capacitor in the transmission line filter DC components. Wiring connects the linear amp IC and quartz boards. The results for the first prototype showed module implementation narrowed the -3 dB bandwidth to 60 GHz. Subsequently, we simulated the module transmission characteristics. Based on the simulation results, we made the following improvements; 1) strengthened the GND of the transmission line; 2) used a common GND between separated connector and the transmission line; and 3) improved impedance mismatch.

Figure 10 shows the minimum signal characteristics of the linear amp module measured using the ME7838D. Flat, good pass characteristics of better than 15 dB (16 dB typ.) above 100 GHz were obtained even with the packaged module. Moreover, with a -3 dB bandwidth covering 200 kHz to 105 GHz (typ.) the linear amp module supports the 100 GHz band. These values satisfy the minimum signal gain and band specifications (Table 1).

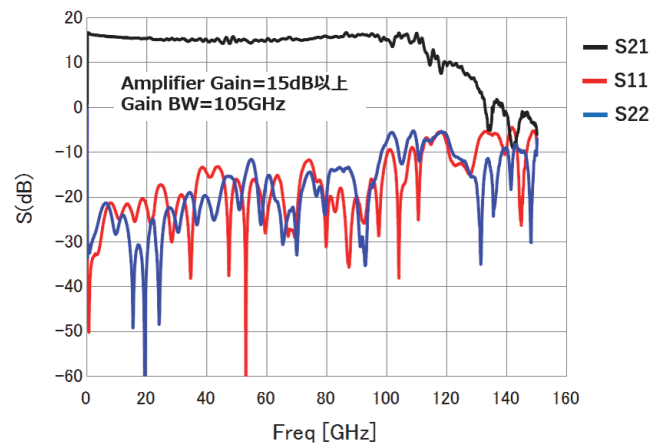


Figure 10 Linear Amp Module Minimum Signal Characteristics Evaluation Results

(3) Linear Amp Product Performance

Figure 11 shows the external appearance of the AH15199A 120 Gbaud Wideband Linear Amp.



Figure 11 External View of AH15199A

Figure 12 shows the AH15199A 120 Gbaud pulse response characteristics (single-end). Compared to the PAM4 input waveform (left), the AH15199A output waveform Eye is open, indicating linear amplification of the input waveform. Consequently, the maximum operating baud rate satisfies the target specification (120 Gbaud typ.). Moreover, the output is amplified to 2.0 Vpp compared to the 0.3 Vpp single-end input amplitude and the voltage gain is 16 dB. As a result, the output amplitude satisfies the target specification (2.0 Vpp typ.).

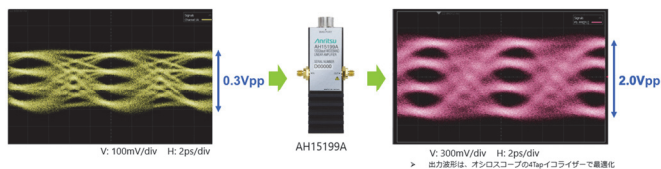


Figure 12 AH15199A 120 Gbaud Pulse Response Properties (Single-End)

As an additional application, combining two AH15199A amps supports amplification of differential output waveforms. However, common mode noise cannot be filtered because this is not fully differential. Figure 13 shows the AH15199A 120 Gbaud pulse response characteristics (differential equivalency) where the input 120 Gbaud 0.6 Vpp input amplitude can be amplified to an output amplitude of 4.0 Vpp.

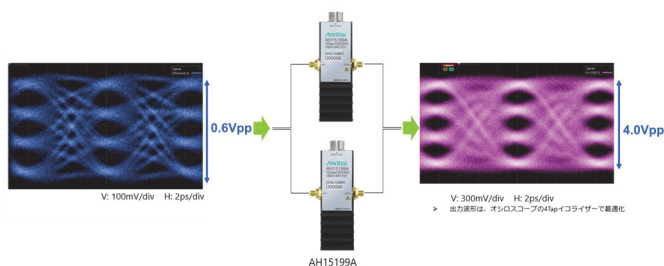


Figure13 AH15199A 120 Gbaud Pulse Response Characteristics (Differential Equivalency)

6 Conclusions

We developed a high-performance ($f_T/f_{\max}=323/510$ GHz) InP HBT by optimizing the structure of the transistor layer and processing technology. We designed a wideband linear amp with good 120 Gbaud pulse response characteristics using the fabricated Gen2 HBT. It was possible to amplify input signal of 0.3 Vpp up to 2.0 Vpp.

Future work will focus on designing even faster ICs using the developed HBT.

References

- 1) M. Urteaga, J. Hacker, Z. Griffith, A. Young, R. Pierson, P. Rowell, M. Seo, and M.J.W. Rodwell, IEDM16-711, 29.2.1 (2016).
- 2) A. Takagi, Y.Matsuoka, M.Ejima, Y.Amano, Y.Ohkubo, T.Koji, K.Kashiwagi, Y.Arashiki, "Development of Functional Modules using High-Performance InP DHBT for Measuring High-Speed Communications Systems", Anritsu Technical Review No. 82, Mar. 2006. (In Japanese)
- 3) Y.Arashiki, Y.Ohkubo, T.Kamizono, Y.Amano, N.Sasaki, "Development of Functional Modules using High-Performance InP DHBT for Measuring High-Speed Communications Systems", Anritsu Technical Review No. 92, Mar. 2017. (In Japanese)
- 4) M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh, Journal of Applied Physics **87**, 2890 (2000).
- 5) Z. Griffith, E. Lind, and M.J.W. Rodwell, IEEE 19th International Conference on Indium Phosphide & Related Materials (2007).
- 6) C. R. Bolognesi, R. Flückiger, M. Alexandrova, W. Quan, R. Lövlom and O. Ostinelli, IEDM16-723, 29.5.1 (2016).
- 7) Z.J. Zhou, K. Ren, J. Liu, W. Cheng, H.Y. Lu, and L.L. Sun., J. Semicond. **36(2)**, 024006 (2015).

Authors



Tomoaki Ishii
2nd Development Dept.
Development Div.
Sensing & Devices Company



Yukio Ohkubo
2nd Development Dept.
Development Div.
Sensing & Devices Company



Masaya Suzuki
2nd Development Dept.
Development Div.
Sensing & Devices Company



Takumi Nemoto
2nd Development Dept.
Development Div.
Sensing & Devices Company



Yutaka Kobayashi
2nd Development Dept.
Development Div.
Sensing & Devices Company



Takashi Kamizono
2nd Development Dept.
Development Div.
Sensing & Devices Company



Moeko Jimba
2nd Development Dept.
Development Div.
Sensing & Devices Company

Publicly available

Development of the Next-Generation Flagship Checkweigher Series

Hiroataka Anzai, Kenta Yamada, Tomohiro Okamura, Yusuke Noguchi, Tsubasa Susa, Jun Kubota, Toshiaki Yamazaki

[Summary]

Anritsu developed a new flagship checkweigher, with an industry-leading sorting performance of up to 1000 pieces/minute and a maximum measurement accuracy of $3\sigma \pm 0.006$ g. The biggest challenge to achieving these ultra-fast weighing speeds is improving responsivity. This was resolved by designing a force-balance type weighcell with a new built-in filter processing and digital feedback control, and by optimizing the weighing conveyor's center of gravity. The diameter of conveyor rollers was reduced to 5 mm, improving product transfer stability. The weighcell pedestal was separated from the checkweigher frame and was made as a truss structure, achieving 1.75 times higher natural frequency when compared to the previous checkweigher series, and as a result, reducing the impact of vibration from external sources. The rejector unit was equipped with a built-in rejector timing correction function, improving line yields by decreasing sorting errors caused by conveyance of disordered products.

1 Introduction

Checkweighers (CW) are used mostly on food and pharmaceutical production lines to measure the mass of every product. They are coupled with a rejector unit that then removes defective products, those that failed to meet the specified mass criteria from the line. There are several purposes for introducing a CW in a production line, from sorting of underweight/overweight/missing products, to minimizing filling losses by providing feedback control to the filler machine. Anritsu entered the CW market in 1964, and since then has been producing models that meet the needs of the changing times, maintaining a strong presence in Japan's domestic market.

More recently, there is a demand for increased production line productivity, a result of rising labor costs due to worker shortages and shorter working hours due to social reforms. As an example, the latest packing machines on the production lines for individually packed candies and pharmaceuticals can handle more than 800 pieces per minute. This requires the surrounding machinery to match the maximum performance of the packing machines. Production line operating rates are also increasing, and more lines are running for 24 hours a day. Implementing high operating rates requires machines that can run for long periods without failure, and that can undergo maintenance for a minimal downtime. There is also an increasing need for high accuracy weighing: to improve yields to offset the rising costs of raw materials, and to cut losses due to rejection of faulty products.

To meet these market needs, Anritsu has developed its 9th generation flagship CW, the AW9 series, using its dynamic weighing technologies accumulated from over 50 years of experience (Figure 1). Achieving a maximum sorting performance of 1000 pcs/minute and a maximum measurement accuracy of $3\sigma \pm 0.006$ g, it is a machine designed to support ultra-fast production lines. Furthermore, it has an easy-to-clean and high-reliability design that features tool-free assembly/disassembly, reducing downtimes which leads to improved productivity. This article introduces the technical approach to developing this series.



Figure 1 External View of the AW9 Series Checkweigher

2 Basic Structure and Operation Principle of the AW9 Series Checkweigher

2.1 Basic Structure

As shown in Figure 2, the AW9 series is configured with a weighcell and three belt conveyors of the infeed, the weighing, and the rejector units. The weighcell is located under the belt conveyor of the weighing unit.

2.2 Weighing Unit (Weighcell)

The weighcell converts the load due to the mass of the weighed product to a readable measurement value. A force-balance type weighcell supporting high-speed weight measurement is designed for the AW9 series.

2.3 Conveyor

Products are conveyed by the three belt conveyors of the infeed, the weighing, and the rejector units. During measurement, since there must only be one product on the weighing unit to get a stable reading, the infeed unit conveyor adjusts the fed product interval so that products to be weighed are carried onto the weighing unit conveyor at a constant interval.

2.4 Rejector Unit

The rejector unit separates under/overweight products from correct-weight products based on the measurement results. It uses an air jet mechanism, and both two and three-direction rejection are supported.

2.5 Operation Unit

The operation unit is for displaying the measurement values, evaluation results, statistical data, etc., as well as for setting operation parameters. It uses a high-visibility 15-inch color LCD touch panel. It can also be rotated and tilted; its position can be adjusted depending on the usage conditions.

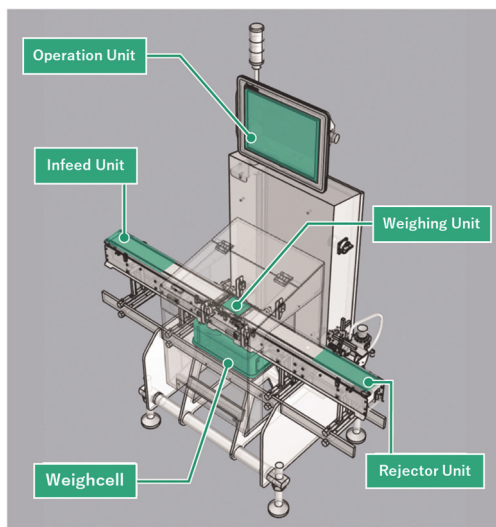


Figure 2 Structure of the AW9 Series Checkweigher

2.6 Basic Operation Principle¹⁾

Figure 3 shows the basic CW operation principle.

An optoelectronic sensor located at the gap between the infeed unit conveyor and weighing unit conveyor detects the passage of the product to be weighed, and measurement is made after the weight signal stabilizes at a certain time interval. These measurement values are evaluated as having passed or failed based on the set limits, and under/over-weight products are removed by the rejector unit.

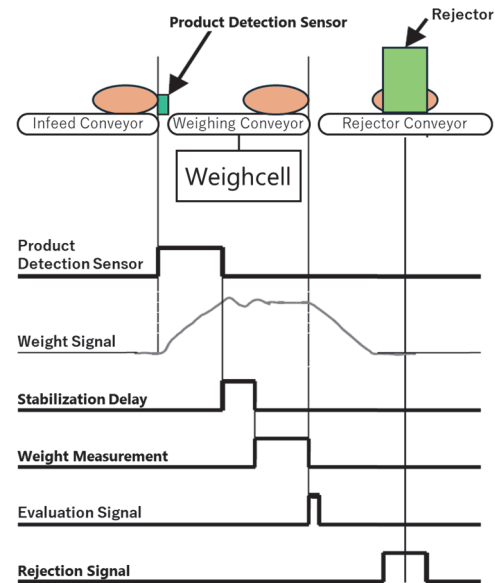


Figure 3 Basic Operation Principle

3 Development Concept

The development of the AW9 series focused on the following key issues to meet market requirements.

- (1) Ultra-Fast, High-Accuracy Weighing
 - Upgrade the maximum sorting performance to 1000 pcs/minute from 600 pcs/minute of the previous CW series
 - Achieve $3\sigma \pm 0.006$ g maximum measurement accuracy
 - Ensure accurate sorting performance that matches fast weighing
 - Improve resistance to external vibrations that cause reduced measurement accuracy
- (2) Reduced Downtime
 - Implement a design with better cleanability and IP54 water resistance
 - Improve sanitary design by using solid materials and fully welded parts
 - Use long-life mechanical parts

- Support tool-free assembly/disassembly
- (3) Improved Operability
- Use a large, high-visibility display for better operability

4 Development Points and Implementation

4.1 Ultra-Fast, High-Accuracy Weighing

Figure 4 shows the weight signal waveforms when a product is weighed multiple times. In comparison to the previous CW series that has a measurement range of about 1.5 g, the AW9 series has a lower range of 0.4 g, achieving higher measurement accuracy. This section introduces the design of a new force balance weighcell and conveyors, technologies that made ultra-fast and high-accuracy weighing possible.

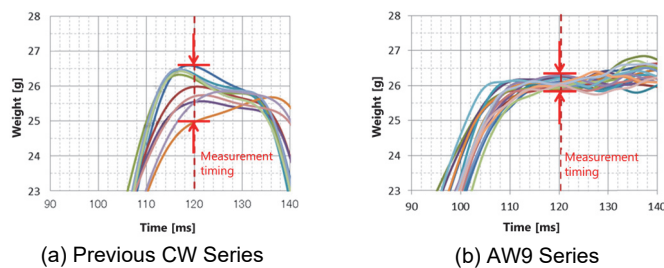


Figure 4 Comparison of Measurement Range

4.1.1 Improved Balance Responsivity

With the filter function of the previous CW series, when sorting performance exceeds 600 pcs/minute, the delay time increases to the point where obtaining correct measurements becomes difficult. In the AW9 series, the number of stages in the analog filter built into the weighcell is reduced by one and the cutoff frequency is optimized for ultra-fast weighing. As a result, the delay time is reduced by 24% when compared to the previous CW series, making ultra-fast weighing possible (Figure 5). The window functions of the digital filter are also increased, enabling automatic selection to achieve optimal weighing depending on product mass and belt speed, making high-accuracy weighing possible.

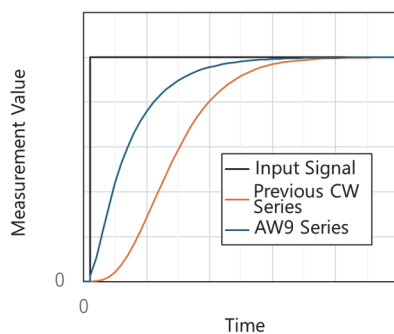


Figure 5 Comparison of Step Input vs Response

In addition, the weighcell feedback control contributes to high-accuracy weighing by changing parameters according to the product type, the conveying conditions, and the installed weighing conveyor. In the previous CW series, fine adjustment of parameters is not supported because analog circuits are used, which prevents optimization according to varying conditions. In contrast, the AW9 series uses digital circuits, so while the circuit board area remains the same size as the previous CW series, the parameter adjustment range is about 300 times bigger, and the resolution is about 9000 times better. As a result, settings can now be adjusted to fit with various conditions, making optimized control possible (Figure 6).

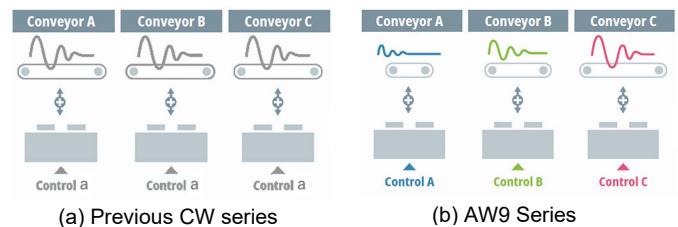


Figure 6 Optimum Feedback Control

4.1.2 Improved Responsivity Through Center of Gravity Optimization and Mass Reduction of Conveyor

Although the previously mentioned feedback control parameters can be optimized to fit every specific conveyor, in order to further increase responsivity, the center of gravity (CoG) of the weighing conveyor should be positioned near the center of the weighcell to minimize any unbalanced load. To change the position of the CoG, the restrictions in the positioning of the weighing conveyor must be solved and the design principle of the previous CW series needs to be changed. In the AW9 series, a new design principle is adopted, where it is standardized to have both an infeed conveyor and a rejector conveyor, with the weighing conveyor positioned in between. This configuration shifts the CoG of the conveyor closer to the center of the weighcell (Figure 7). In addition to shifting the CoG, reducing the mass of the conveyor is also effective in improving responsivity, thus the weighing conveyor is made lighter by 22% in comparison to the previous CW series. A model with compact conveyor of belt width 70 mm and length 100 mm, targeted for smaller products, is also added to the lineup of the AW9 series.

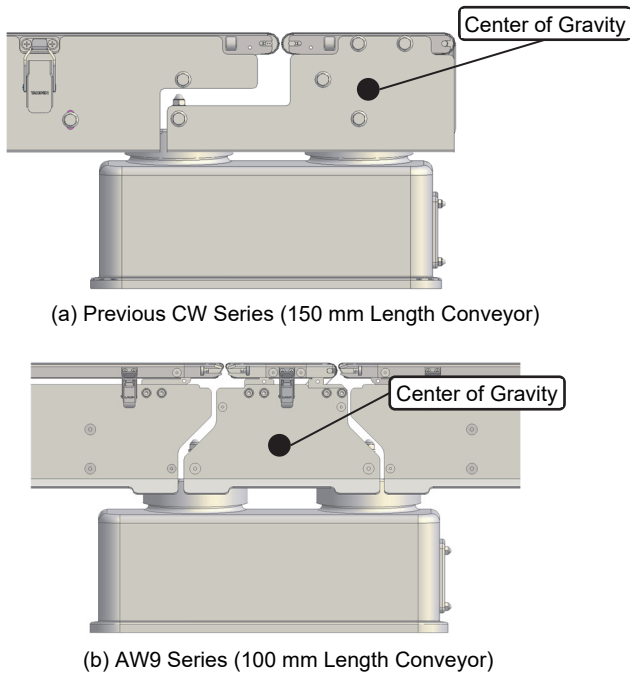


Figure 7 Weighing Conveyor

4.1.3 Improved Product Transport Stability and Conveyor Vibration Reduction Through Smaller Rollers

When products to be weighed are transported from the infeed conveyor to the weighing conveyor, stable weighing can be prevented by discrepancies in the measurement timing due to the orientation of the product becoming disordered, or by vibrations caused when products bump as they pass through the gap between the conveyors. These effects become more pronounced as conveying speeds become faster. To reduce these issues, the gap between the conveyor rollers should be minimized. In the AW9 series, the diameter of rollers at the gap between the infeed conveyor and weighing conveyor is reduced from the previous CW series' 18-mm to 5-mm (Figure 8). As a result, the orientation of products remains stable as they pass through the gap between the conveyors, even at ultra-fast sorting performance of 1000 pcs/minute, making high-accuracy weighing possible. Moreover, high-accuracy weighing of tall products with relatively poor stability is also possible.

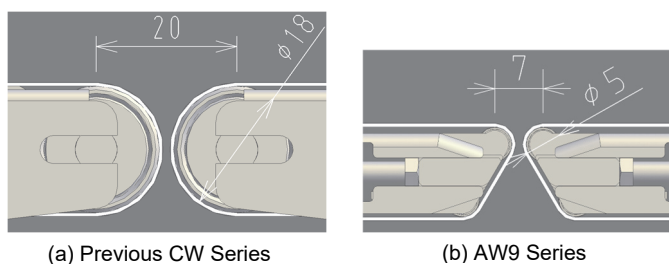


Figure 8 Gap Between the Infeed and the Weighing Conveyors

There is also the innate vibration generated from the rollers and the belt of the weighing conveyor itself, which contributes to the dispersion of measurement values. By reducing the diameter of the rollers, the frequency of the generated vibration increases at the same conveyor speed. Since the attenuation effect of the weighcell filter function is greater on higher frequencies, the dispersion of measurement values due to conveyor vibration is reduced by as much as 45% when compared to the previous CW series (Figure 9).

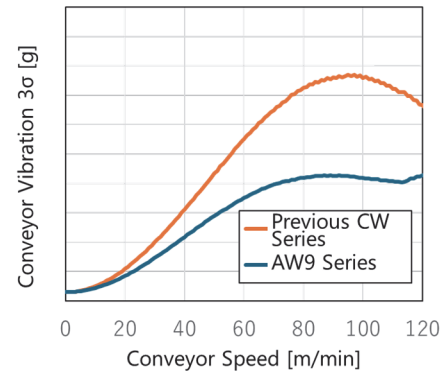


Figure 9 Dispersion of Measurement Values due to Conveyor Vibration

4.2 Air Jet and Rejector Timing Correction Function

4.2.1 Effect of Air Jet on the Measurement Values

When using an air jet mechanism on a checkweigher, the effect on the measurement values due to airflow and pressure change while air is being blown becomes an important consideration. This effect is influenced by the orientation of the air nozzle and its distance from the weighing unit. And by finding the optimum balance between these influencing factors, high-speed sorting performance coupled with high-accuracy weighing is achieved (Figure 10).

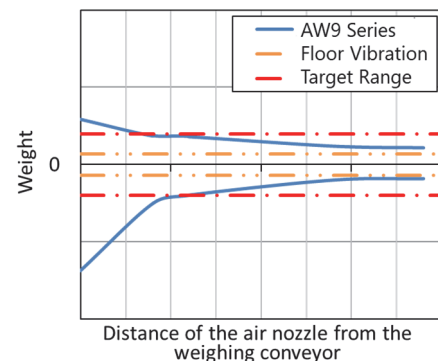


Figure 10 Dispersion of Measurement Values While Air is Being Blown

4.2.2 Rejector Timing Correction Function

As conveying speeds increase, the orientation and spacing between products can also become more disordered. When this happens, there will be variation in the time between when the product passes through the product detection sensor and when it reaches the rejector, leading to sorting errors. In the previous CW series, a sorting confirmation function is equipped to detect sorting errors, removing any product affected regardless of its measurement value. However, this approach reduces yields in the production line.

In the AW9 series, in addition to improved product transfer stability due to the conveyor design, there is now also a built-in rejector timing correction function (Figure 11).

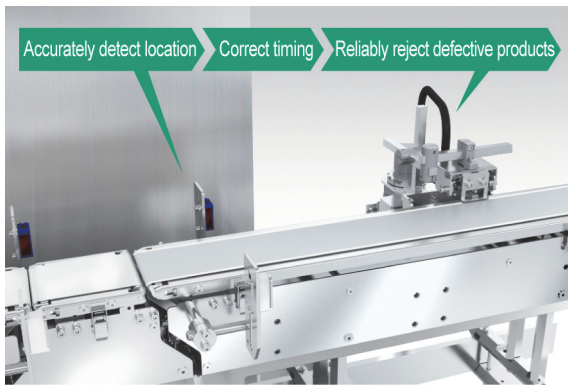


Figure 11 Rejector Timing Correction Sensor

An optoelectronic sensor mounted at the rejector unit detects the passage of weighed products, and the rejector timing correction function calculates the remaining time until the products reach the rejector. By using this calculated time to adjust the rejector timing, the rejector activates at a definite time interval after products pass through the optoelectronic sensor (Figure 12). With this function, yield is improved by reducing sorting errors caused by conveyance of disordered products.

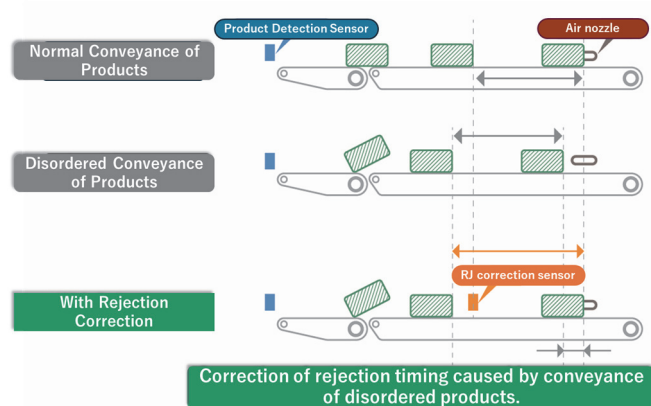
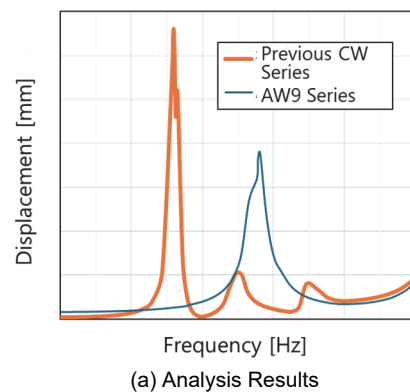


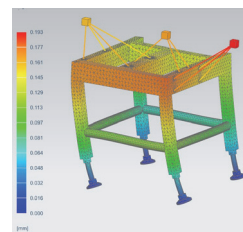
Figure 12 Rejector Timing Correction Operation

4.3 Countermeasures for External Vibration

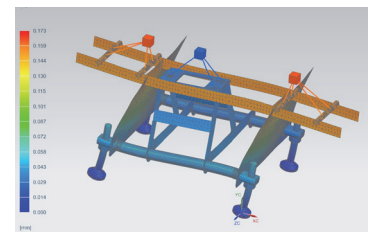
Vibration has a large negative impact on the measurement accuracy of a weighcell. In a production line, surrounding equipment transmits vibration through the floor, and components from the CW itself such as conveyors transmit vibration through the frame, making it impossible to completely eliminate the effects of vibration. In the AW9 series, the weighcell pedestal is separated from the CW frame, reducing the amount of vibration being transmitted from vibration sources to the weighcell. Additionally, the weighcell pedestal is made as a truss structure, increasing the rigidity in all the X, Y, and Z directions, and as a result improving the weighcell responsivity and filter attenuation effect. Figure 13 shows the dynamic response analysis results. The natural frequency of the frame is 1.75 times higher when compared to the previous CW series and the displacement of the weighcell pedestal is reduced by 42%.



(a) Analysis Results



(b) Previous CW Series



(c) AW9 Series

Figure 13 Dynamic Response Analysis Results and Contour Diagram of the Frame

4.4 Reducing Downtime

To reduce downtime, the development of the AW9 series is focused on the following: shortening of cleaning time by improving on sanitary design, and reducing parts replacement frequency by using highly durable parts, primarily on the moving components of the conveyors. Particularly for the cleanability, special attention is paid to the sanitary requirements listed in Table 1. This section explains the details.

Table 1 List of Sanitary Requirements

Requirement
Minimize microbial growth.
Prevent accumulation of water on surfaces.
Ensure the ease of cleaning.
Use only materials specified in regulations for parts in contact with foods.

4.4.1 Cleanability

Considering regular washdowns and exposure to water, the frame is designed to prevent the growth of microbes. This is accomplished by using solid materials, by reducing flat surfaces as much as possible, and by incorporating fully welded joints. In addition, the surface contact between the parts of the conveyors is minimized and the chassis and the windshield covers are made to have inclined surfaces, preventing water and dust accumulation. The AW9 series has an IP54 rating, guaranteeing protection against dust and water splashes.

4.4.2 Usage of Long-Life Mechanical Parts

To meet the demand for increased productivity, production lines running for 24 hours a day have been increasing. In response, motors with higher capacity than the ones in the previous CW series are used. By increasing the load margin for conveyance, the heat generated by the motor during operation is reduced, leading to increased lifespan. Lastly, two choices for motor capacity are added to the lineup of the AW9 series, allowing optimized selection depending on product mass and specific conveying conditions.

4.4.3 Tool-free Assembly/Disassembly

The newly designed high-speed air jet and product guides are fixed in place using clamp levers and thumb screws, allowing for their easy removal. The conveyor body is attached to the conveyor frame using catch clips, and the tension of the conveyor belt is adjusted using thumb screws. Everything can be removed and disassembled completely without tools, facilitating the quick cleaning of the conveyors (Figure 14).

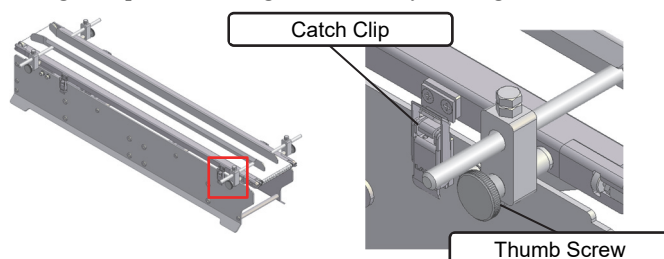


Figure 14 Tool-Free Assembly/disassembly

4.5 Improved Operability

A high-visibility 15-inch color LCD touch panel is used to easily view and quickly understand the operating status of the machine, the production conditions, and the presence of errors/alarms. This section introduces some of the operation screens.

(1) Statistics Screen

This screen is used to confirm various types of information, such as the overall and the passed products statistics. It can also be used to confirm the production start time, total mass measured, average mass, standard deviation, etc., allowing the detailed checking of production conditions (Figure 15).



Figure 15 Statistics Screen

(2) Magnified Measurement Display Screen

This screen is configured in such a way that the measurement values and the evaluation results (pass/underweight/overweight) can be easily viewed by the user, even at a distance. Information regarding sorting performance can also be confirmed here (Figure 16).



Figure 16 Magnified Measurement Display Screen

(3) IO Timing Adjustment Screen

This screen is used to check signals from the sensors and the rejectors. Rejector timing can also be confirmed here, making it easy to make timing adjustments for each individual product. Sampling cycle can be selected from 1 ms/5 ms/10 ms (Figure17).

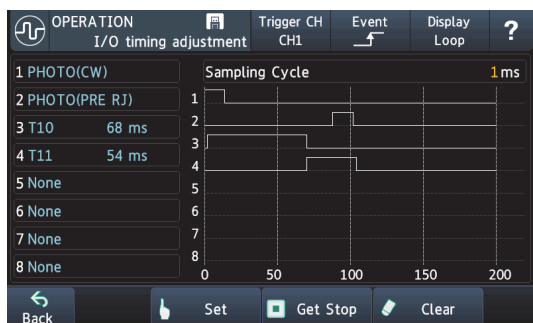


Figure 17 IO Timing Adjustment Screen

(4) Rejector Timing Correction Adjustment Screen

This screen is used to automatically adjust rejector timing correction related parameters by feeding the target product multiple times. This simplifies the usage of the RJ timing correction function (Figure 18).

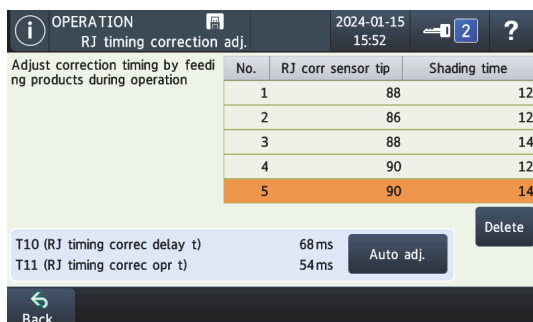


Figure 18 Rejector Timing Correction Adjustment Screen

5 Main Specifications

Figure 19 shows the AW9 series' external dimensions diagram, and Table 2 lists the main specifications.

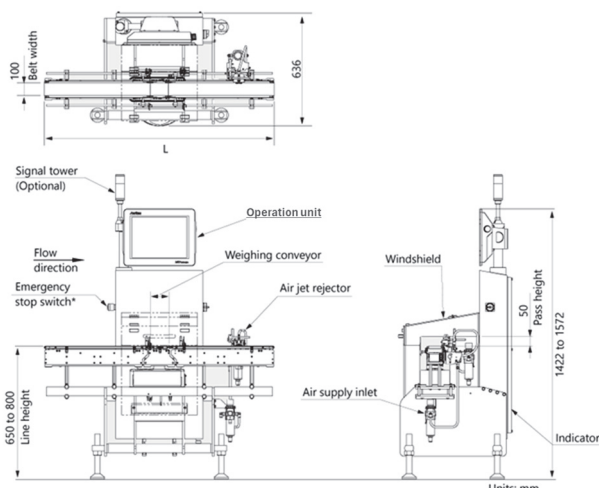


Figure 19 KWA7241APANN External Dimensions Diagram

6 Conclusion

Anritsu has successfully developed its flagship AW9 checkweigher series with a weighcell and conveyors

optimized for ultra-fast weighing. It boasts an industry-leading sorting performance of up to 1000 pcs/minute and a maximum measurement accuracy of $3\sigma \pm 0.006$ g. This flagship model is the ideal solution for product quality assurance and productivity improvement on food and pharmaceutical production lines. With its aim to continue providing new value to its customers, Anritsu will focus on developing new models and expanding the lineup of the AW9 series.

References

- 1) A.Iida, J.Tamura, K.Yamada, K.Suzuki, H.Sato, T.Ito, H.Anzai, M.Bajo: "Development of SSV Series Multilane Checkweighers" Anritsu Technical Review No.27, pp.17-24 (Sept. 2019)

Authors



Hirotaka Anzai
Product Development Dept.
Development Div.
Infivis Company



Kenta Yamada
Product Development Dept.
Development Div.
Infivis Company



Tomohiro Okamura
Product Development Dept.
Development Div.
Infivis Company



Yusuke Noguchi
Product Development Dept.
Development Div.
Infivis Company



Tsubasa Susa
Product Development Dept.
Development Div.
Infivis Company



Jun Kubota
Product Development Dept.
Development Div.
Infivis Company



Toshiaki Yamazaki
Product Development Dept.
Development Div.
Infivis Company

Table 2 Main Specifications

Model		KWA7241AP■N△	KWA7242AP■N△	KWA7243AP■N△	KWA7244AP■N△	KWA7245AP■N△
		■: A (High-speed airjet)/N (no rejector) △: C (CE marking compliance)/N (No optional standards compliance)				
Weighing Range		1.2 to 660 g				
Display Scale		0.002 g				
Max. Sorting Performance		1000 pcs/minute	900 pcs/minute	720 pcs/minute	540 pcs/minute	400 pcs/minute
Max. Measurement Accuracy (3σ)		±0.006 g				
Display		15-inch touch panel color TFT LCD				
Operation Method		Touch panel				
Max. Displayed Mass		660.9 g				
Supported Product Variety		200 product types				
Rejector Stages		2 (3 as option)				
Weighing Conveyor Length		100 mm	120 mm	150 mm	200 mm	270 mm
Supported Product Dimensions	Width	20 to 70 mm	20 to 100 mm			
	Length	24 to 90 mm	24 to 100 mm	24 to 130 mm	24 to 180 mm	24 to 240 mm
	Height	5 to 90 mm	5 to 100 mm	5 to 130 mm	5 to 180 mm	5 to 240 mm
Belt Speed		110 m/minute	132 m/minute	165 m/minute	180 m/minute	
Power Supply		AC100 to 120 V +10%/–15% or AC200 to 240 V +10%/–15%, 50/60 Hz single phase				
Power Consumption		140 VA				
Air Supply		0.5 to 0.9 MPa				
Mass		155 kg				
Operating Environment		0° to 40°C (≤5°/h change) Relative humidity 30% to 85% with no condensation				
Protection Class		IP54				
External Finish		Stainless Steel (SUS304)				
Data Output		USB2.0/Ethernet (10BASE-T, 100BASE-TX)				

Publicly available

Automating External Visual Inspection Using Collaborative Robot

Tomohiro Akiyama, Takashi Yoshida, Hideyuki Sakamoto, Tsunayoshi Hara, Ryo Kondo

[Summary]

Efforts are in progress to automate production at our main Anritsu Tohoku plant. One plan is to investigate automation of finished instrument external inspection now done visually by quality inspectors. This article introduces development of a prototype system combining a collaborative robot and AI image processing for this inspection and presents some examples of automated external inspection results.

1 Introduction

Anritsu main product is measuring instruments that are manufactured mostly in Japan at its Tohoku Anritsu plant. To assure the accuracy of manufactured measuring instruments, various adjustments and inspections are made. Quality assurance (QA) requires experienced skilled workers but an ageing workforce and falling population will probably be securing sufficient employees difficult in future. To support recent changing market demands, it is necessary to abandon from labor-based production systems, and promotion of DX (digital transformation) to save labor and introduce automation is becoming a key approach.

2 Automating External Visual Inspection

2.1 Target Process

Figure 1 shows the measuring instrument production process at Tohoku Anritsu.

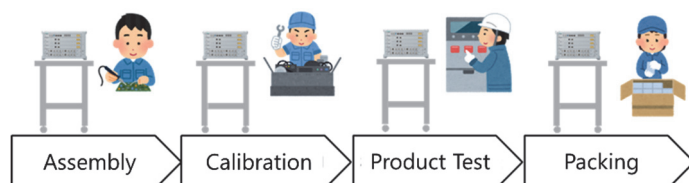


Figure 1 Measuring Instrument Production Process

Each instrument is assembled after picking the required parts. The next calibration stage makes each adjustment necessary to assure the measuring instrument function. Following calibration, the instruments are tested and inspected to confirm compliance with each specification, heat tested, and evaluated for performance. After passing all tests, they are packed and shipped.

This article explains the process of external visual inspection of finished products at the shipping inspection prior to shipment.

Since many types of measuring instruments are produced in small lots, Anritsu Tohoku uses the cell production method where each instrument moves on a cart between each process. Even the final external visual inspection of a finished product is performed on this cart (Figure 2).



Figure 2 Measuring Instrument on Cart

2.2 Inspection Contents

Table 1 lists the external visual inspection items of finished measuring instruments. Figure 3 shows an external view of a typical measuring instrument.

Table 1 Inspection Items

Item	Inspection Contents
Text Matching	Correctly printed panel and label text
Barcode Reading	Readability of barcode label
Label Position	Presence of correct label at correct position
Connector Shapes	Correct connector shapes
Damage and Dirt	Absence of damage and dirt



Figure 3 Typical Instrument External Appearance

To prevent shipment of faulty products, the same inspection is performed twice by two inspectors.

2.3 Challenges at Finished Product External Inspection

External visual inspection by workers has the following challenges. Automating this inspection would not only solve labor shortages but would also solve these challenges.

- (1) Shipment of faulty instrument due to inspection oversight

Varied small-lot production of measuring instruments make it difficult for inspectors to fully understand the reference inspection standards since there are many specifications according to the shipping destination.

Moreover, text matching requires repeated checking of text that inspectors may not understand, which increases the work burden, causing fatigue and making avoidance of careless mistakes difficult.

- (2) Visual inconsistencies between inspectors

Different testers have different personal standards regarding damage and dirt. Although the damage standard is defined using a sample, if the damage is less than the sample, the inspector may still reject the product, possibly leading to excessive quality standards.

- (3) Provision of shipping inspection evidence (images)

When responding to post-shipment warranty claims, despite having inspection evidence, this is only inspection items and results and does not prove that there was no problem at shipment. The amount of time required for inspectors to take photographic images of every item for all products manually is unrealistic.

2.4 Automation Concepts

Automating finished product external inspection was based on the following concepts as the first step in targeting future fully automated inspection.

- Rather than create a large-scale automation system, assume instruments are conveyed on the product cart.
- Focus on obtaining evidence without manual intervention. Inspection time (tact time) is not important.
- The system must be expandable for future production of various types of instruments.
- First, we aim to implement the current inspection system, which is conducted by two inspectors, with an automated system plus one inspector.

3 Prototype Development

3.1 Required Specifications

The required specifications for the prototype system performing automated external inspection of finished products are listed below.

- Must inspect all five surfaces excluding base while instrument on cart
- Must execute inspection from start to finish without manual intervention
- Must save images captured during inspection

3.2 System Configuration

Figure 4 shows the external diagram of the developed prototype and Table 2 lists the main parts.

A camera unit mounted on the end of the robot arm captures images from every angle of an instrument on the cart.

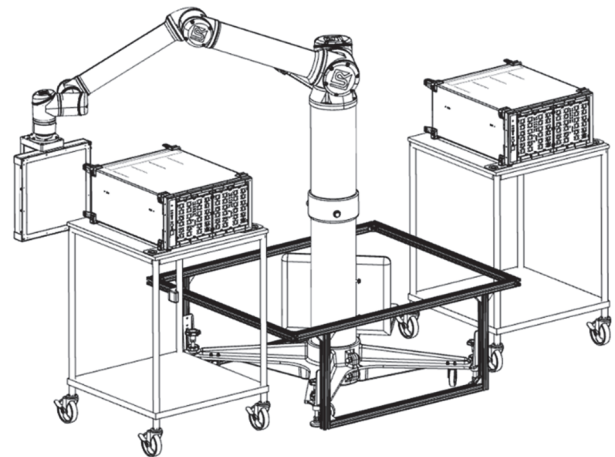


Figure 4 Prototype External Inspection Diagram

Table 2 Main Structural Parts

No	Component	Qty	Remark
1	Camera Unit	1	Flat dome lighting + camera
2	Robot System	1	Robot arm + control box
3	Robot trestle	1	Cast aluminum frame + positioner
4	PC Controller	1	Windows PC running inspection software
5	Operation Panel	1	Touch panel
6	Monitor	1	Inspection screen display

3.3 Details of Each Structural Element

The details of the prototype system parts are listed below.

- (1) Camera Unit

The structure of the camera unit is shown in Figure 5 and Table 3 lists the parts. The system is composed of lighting, camera, and lens for shooting the required

images at inspection. There are two different optical cameras to shoot the images with pulling and close up.

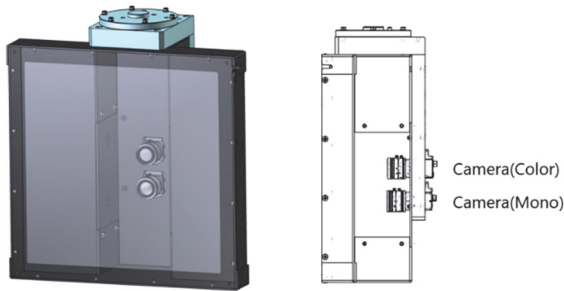


Figure 5 Camera Unit

Table 3 Camera Unit Parts

No	Part	Qty	Main Specification
1	Camera (color)	1	5 million pixels CCD camera + 12 mm lens For wide-angle images
2	Camera (monochrome)	1	5 million pixels CCD camera + 25 mm lens For damage/shape inspection
3	Flat Dome Light	1	Light size 30 cm × 30 cm
4	Camera Unit Handler	1	Camera and light base w/ attachment to robot arm

(2) Robot System

Table 4 lists the robot system specifications.

When choosing the robot system, key points were the ability to handle the camera unit weight and an arm length permitting imaging of all five surfaces of the inspected instrument.

We selected the collaborative robot from Universal Robot (hereafter UR). Since general-purpose production robots require work space and inspectors, it was necessary to secure a wide space. In a production plant mainly using cell-style production, it was thought a collaborative robot was advantageous from the layout efficiency perspective.

Table 4 Robot System Specifications

Item	Specification	Explanation
Carriage Weight	10 kg	Camera unit weight approx. 5 kg
Arm Length	1300 mm	—

(3) Inspection Software

Figure 6 shows the prototype system software structure.

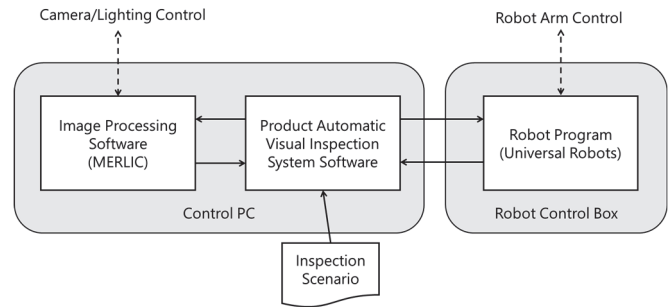


Figure 6 Prototype Software Blocks

Inspection using this system uses operation software running according to inspection scenarios on the PC controller.

The camera unit is moved via the robot program to required inspection locations according to inspection scenarios for each instrument type. The cameras and lighting are controlled to capture images that are then image-processed by software for inspection.

The captured image processing uses commercial software (MERLIC) which has rich image processing libraries such as OCR and barcode reading.

3.4 Inspection Items and Flow

The prototype system inspection flow is described below.

- (1) Positioning on Product Cart
 - Inspector positions instrument at specified position on cart using guide.
- (2) Inspection Start
 - Inspector starts inspection using operation panel.
 - Monitor displays inspection progress status.

Table 5 lists the inspection items executed by this system.

Table 5 Prototype Inspection Items

Inspection Item	Inspection Target
Text Matching	Front panel printed text
	Front panel label text
	Back panel label text
Barcode Reading	Back panel S/N
Label Position	Top panel Eco label
Connector Shape	Front panel N-connector center axis opening
Damage and Dirt	Top panel damage and dirt

(3) Inspection End

- System generates inspection results report (OK/NG evaluation result and photo images)
- Inspector moves cart with inspected instrument

3.5 Design Points

The following section describes the important design points when developing the prototype.

(1) Robot System Choice

When choosing the robot system, we used 3D simulation (Figure 7) to confirm whether candidate systems could capture images of all five instrument surfaces without problems.

3D data from each robot system vendor was used to evaluate the positional relationship between the robot and instrument and whether the instrument surface furthest from the robot could be imaged by the camera unit.

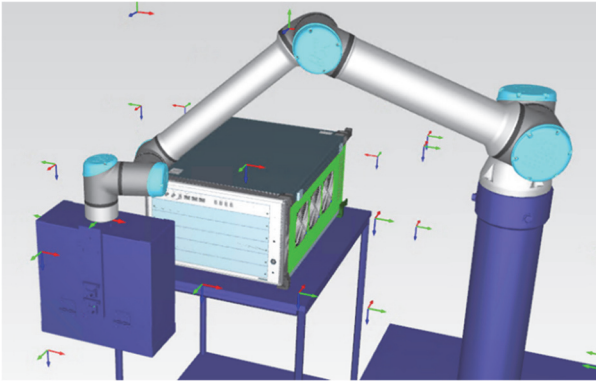


Figure 7 3D Simulation

(2) Lighting Choice

The lighting choice for external visual inspection of the instrument was most difficult. The main product panels are vertical metal sheets with satin-like finish that reflects light when illuminated by light from a ring or bar lamp using for ordinal visual inspection and prevents image processing at some angles. The best solution to this reflection issue was a dome light that illuminates the target equally from the 90° to surface direction. However, a large dome light could be hit by the robot arm and the camera opening in the dome tended to be a little darker, which was unsuitable for wide-area inspection.

Comparing the results of inspections using various light sources led to use of a flat dome light source, which gives the same equal illumination as a dome light but does not require a camera opening. Additionally, the LCD panel shape can be carried by the robot arm (Figure 8).

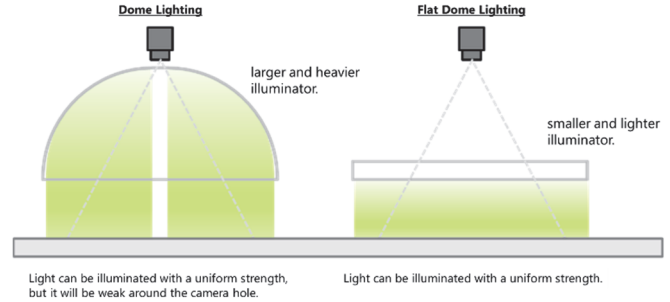


Figure 8 Dome and Flat Dome Lighting

(3) Instrument Position Detection

Although the instrument cart used for inspection is put in position using a guide, there can be somehow differences in the position of the instrument on the cart. Performing inspection requires an accurate understanding of the instrument position.

Generally, specific markers are attached to a target and the distance and angles between these markers are calculated by image processing. However, the target in this inspection is a product to which markers cannot be attached.

Therefore, we devised a method to detect the instrument corners by image processing (Figure 9).

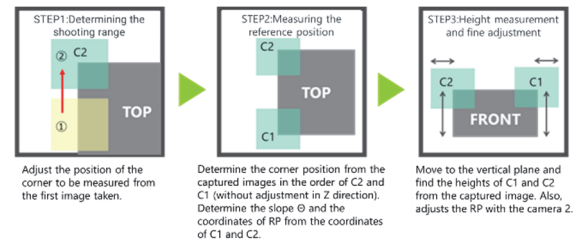


Figure 9 Instrument Position Detection

The procedure is as follows:

1. Detect corner (C2) of front panel by taking picture of product top surface
2. Detect corner (C1) of front panel by taking picture of product top surface
→ Top panel surface angle adjustment
3. Detect corner (C2) of front panel by taking picture of product top surface
4. Detect corner (C1) of front panel by taking picture of product top surface
→ Top panel height adjustment

The product position is calculated from the detected cabinet angle coordinates.

(4) Using AI (Deep Learning)

When detecting labels (presence/absence and position) by image processing, conventional pattern matching processing tends to be poor at the peripheral edges although the image center detection is good. This is thought to be due to breaks in the shape caused by lens distortion. Consequently, we used AI (Deep Learning) for the image processing software.

We used a learning model with 100 images and data on label types and annotations were appended to the evaluation to improve label recognition rates (Figures 10 and 11).



Figure 10 Deep Learning Function (Learning Screen Examples)



Figure 11 Using Deep Learning Function (Label Detection)

(5) Expandability

Assuming the range of target instrument types will expand in future, we constructed a software platform using scenario control.

Every software scenario can select the instrument inspection surface and move the camera unit, supporting many instrument types without requiring corrections to the robot program.

Additionally, image processing of each captured image is handled by loading the required module from the scenario, so just the image processing can be adjusted without changing the scenario.

4 Evaluations

4.1 Possible Inspections

We evaluated the prototype to discover issues about the accuracy of inspection results using image processing and

installing the automated system into the manufacturing line. Table 6 lists current inspection contents and Figure 12 shows a possible prototype inspection setup.

Table 6 Possible Prototype Inspection Contents

Inspection Item	Normal Test Sample Count	Abnormal Test Sample Count
Text Matching	20	20
Barcode Reading	20	20
Label Position	20	20
Connector Shape	20	1
Damage and Dirt	20	1



Figure 12 Possible Prototype Inspection Setup

4.2 Evaluating Inspection Accuracy

The inspection accuracy evaluation is shown below.

(1) Text matching Inspection

Table 7 lists the inspection accuracy evaluation results and Figure 13 shows an example of an inspection result image.

There are no problems with reading the text itself and an error in the product serial number (SN) on the seal attached to the instrument could be detected, which was the target purpose. The text matching accuracy achieved a satisfactory usage level.

Table 7 Inspection Accuracy Evaluation Results (Text Matching)

Text Matching Confusion Matrix		Evaluation with Prototype	
		OK	NG
Actual Result	OK	20	0
	NG	0	20

Verdict	OK	NG
Criteria	1123456781	
Detected	1123456781	1123456782
Image		

Figure 13 Example of Inspection Results Image (Text Matching)

(2) Barcode Reading Inspection

Table 8 lists the inspection accuracy evaluation results. Figure 14 shows an example of an inspection result image.

Similar to text matching, there was no problem reading the barcode itself and it was possible to detect a mis-attached barcode label. The barcode reading accuracy achieved a satisfactory usage level.

Since multiple barcodes can be verified as a batch, this function is expected to reduce inspectors' workloads.

Table 8 Inspection Accuracy Evaluation Results (Barcode Reading)

Barcode Reading Inspection Confusion Matrix		Evaluation with Prototype	
		OK	NG
Actual Result	OK	20	0
	NG	0	20

Verdict	OK	NG
Criteria	Printed characters and barcode must match	
Detected	#1 OK : TXT 9876543216 CODE 9876543216 #2 OK : TXT 9876543217 CODE 9876543217 #3 OK : TXT 9876543218 CODE 9876543218 #4 OK : TXT 9876543219 CODE 9876543219 #5 OK : TXT 9876543220 CODE 9876543220 #6 OK : TXT 9876543221 CODE 9876543221	#1 OK : TXT 9876543216 CODE 9876543216 #2 OK : TXT 9876543217 CODE 9876543217 #3 OK : TXT 9876543218 CODE 9876543218 #4 OK : TXT 9876543219 CODE 9876543219 #5 OK : TXT 9876543220 CODE 9876543220 #6 NG : TXT 9876543221 CODE 9876543222
Image		

Figure 14 Example of Inspection Results Image (Barcode Reading)

(3) Label Inspection

Table 9 lists the label accuracy evaluation results and Figure 15 shows an example of the evaluation results image.

There was no problem with label position detection and it was possible to detect whether the attachment

angle was within the reference standard. Label detection reading accuracy achieved a satisfactory usage accuracy level.

Table 9 Inspection Accuracy Evaluation Results (Label)

Label Detection Confusion Matrix		Evaluation with Prototype	
		OK	NG
Actual Result	OK	20	0
	NG	0	20

Verdict	OK	NG
Criteria	X direction: within ± 2 mm, Y direction: within ± 2 mm Angle: within $\pm 2^\circ$	
Detected	X direction: -0.61mm Y direction: -1.28mm Angle: 0.16°	X direction: -0.1mm Y direction: 0.6mm Angle: 2.1°
Image		

Figure 15 Example of Inspection Results Image (Label)

(4) Connector Shape Inspection

Table 10 lists the inspection accuracy evaluation results and Figure 16 shows the inspection results image.

The connector shape inspection accuracy achieved a satisfactory usage level. Moreover, this inspection has no criteria about opening angle of the central axis and the OK/NG verdict at conventional inspection has dependency to a vary inspector, therefore this robot inspection system must be much effective to improve such fluctuation as well.

Table 10 Inspection Accuracy Evaluation Results (Connector Shape)

Connector Shape Inspection Confusion Matrix		Evaluation with Prototype	
		OK	NG
Actual Result	OK	20	0
	NG	0	1

Verdict	OK	NG
Criteria	$r=0.8\pm0.2$ mm, quality ≥ 0.7	
Detected	$r=0.839$ mm Quality=0.860	$r=N/A$ Quality=N/A
Image		

Figure 16 Example of Inspection Results Image (Connector Shape)

(5) Damage and Dirt Inspection

Table 11 lists the inspection accuracy evaluation results and Figure 17 shows an example of an inspection results image.

There was no problem detecting deliberately inflicted damage but there some NG evaluations for OK instruments. The main cause was that dust and dirt that was present when the image was captured was evaluated as damage. We think the solution will require adjusting the image-capture environment and image processing.

Table 11 Inspection Accuracy Evaluation Results
(Damage and Dirt)

Damage and Dirt Confusion Matrix		Evaluation with Prototype	
		OK	NG
Actual Result	OK	12	8
	NG	0	1

Verdict	OK	NG
Criteria	Scratch : 0, Stain : 0	
Detected	Scratch : 0, Stain : 0	Scratch : 0, Stain : 1
Image		

Figure 17 Example of Inspection Results Image (Damage and Dirt)

4.3 Discovered Automated System Challenges

(1) Improving Inspection System Accuracy

A satisfactory accuracy was achieved for practical inspection of text matching, barcode reading, and label and connector shape inspection.

On the other hand, currently, the damage and dirt inspection misreported the presence of faults. We shall try to improve the image-processing detection accuracy necessary for an automated system assuming operation in collaboration with inspection personnel.

(2) Automation Inspection Work

This prototype system was able to completely automate inspection without manual intervention by inspection staff from after the candidate instrument (on its cart) had been correctly positioned. In particular, automated capture and storage of images as inspection evidence linked to inspection OK/NG evaluation results will probably have a large impact on production quality management.

Implementing future fully automated inspection without staff intervention still faces challenges in designing equipment for automated transport, positioning, and handling of inspected instruments.

5 Conclusions

We combined a collaborative robot and AI image processing to automate external visual inspections of finished instruments previously performed by quality inspectors.

In addition to solving remaining challenge, we are targeting rapid introduction based on this system to our production lines to fully automate manual inspections now using other human senses.

Authors



Tomohiro Akiyama
Fundamental Technology Dept.
Engineering Div.
/DX Promotion Dept.
Tohoku Anritsu Co.,Ltd,



Takashi Yoshida
DX Promotion Dept.
Tohoku Anritsu Co.,Ltd,



Hideyuki Sakamoto
Fundamental Technology Dept.
Engineering Div.



Tsunayoshi Hara
Fundamental Technology Dept.
Engineering Div.



Ryo Kondo
Fundamental Technology Dept.
Engineering Div.

Publicly available

Enhanced Raman scattering by zigzag edges of graphene and spectral analysis using the Bayesian method

Taisuke Ochi, Tomohiro Matsui

[Summary]

Surface enhanced Raman scattering (SERS) is a phenomenon, in which the Raman scattering of adsorbed molecules is enhanced by a factor of 10^3 – 10^6 through the interaction between the molecules and the substrate. It is attracting attentions as a method to identify and evaluate material from a minute amount. In this study, we focused on the electrons localized at the zigzag (zz) edge of graphene and verified if the Raman scattering is enhanced through the localized electrons. By investigating the Raman scattering of rhodamine 6G (R6G) molecule adsorbed on the zz graphene nano-mesh (zGNM) as a substrate, it is found that zGNM shows stronger enhancement of the Raman scattering than pristine graphene, despite the smaller graphene area. This suggests that the Raman scattering is enhanced on a zz edge. Furthermore, the study about the enhancement dependent on the R6G concentration shows that the R6G molecule is selectively adsorbed on the zz edge. Since the Raman spectrum contains many peaks, the analysis of the Raman spectrum is not straightforward. In this study, the Bayesian method was used to decompose the spectrum into peaks.

1 Introduction

When light is irradiated on a crystal or a molecule, a part of the incident light excites vibrations in the material, and the wavelength of scattered light is shifted from original one by the energy of the excited vibrations. This process is called the Raman scattering. Since the energy of the excited vibration is intrinsic to the molecule, its constituting atom, and the structure of the material, one can identify the material and evaluate its structure from the wavelength difference (Raman shift) between the incident and the Raman-scattered lights. This technique is known as the Raman spectroscopy. It is expected to be applicable to various fields such as chemical process and drug managements, pathological diagnosis, and atmospheric/environmental measurement, because it is a label-free, non-destructive, and non-invasive measurement^{1)–3)}. However, this measurement does not work for a minute amount of molecule, since the intensity of the Raman-scattered light is extremely weak. It is about 10^{-6} order of magnitude smaller than the Rayleigh-scattered light that does not involve vibrational excitations. Surface enhanced Raman scattering (SERS)^{4), 5)} is one of the methods to overcome this problem. It is a phenomenon that the Raman scattering is enhanced by a factor of 10^3 – 10^6 through the interaction between the substrate and the adsorbed molecule, and it enables the Raman spectroscopy even for a minute amount of molecule.

Although the curve fitting by the least-squares method is commonly used for analysis of spectroscopic spectrum such as Raman spectroscopy, this method sometimes overestimates peaks and mis-identifies noise as a peak when the spectrum is complicated. To overcome such problems, an analysis based on the Bayes' theorem (Bayesian method)⁶⁾, which estimates the result based on the measured data under pre-defined conditions, is efficient.

In this study, enhancement of the Raman scattering through the electronic states localized around the zigzag (zz) edge of graphene^{7)–9)} was investigated. The Raman spectra of rhodamine 6G (R6G) molecule, one of the general dyes, adsorbed on a zz graphene nano-mesh (zGNM)^{12), 14)} fabricated by hydrogen plasma etching (HPE)^{9)–13)}, were measured, and it was found that the Raman scattering of R6G molecule is enhanced around the zz edge, and that R6G molecule is selectively adsorbed on the zz edge. Here, the Bayesian method was used to analyze the Raman spectra.

This article introduces the enhancement of the Raman scattering on zGNM as well as the spectral analysis using the Bayesian method. The structure of this article is as follows. Chapter 1 described the Raman scattering and its potential applications. Chapter 2 outlines the mechanism of SERS, and chapter 3 discusses the enhancement of the Raman scattering on zGNM. Finally, in chapter 4, the spectral analysis by the Bayesian method will be introduced.

2 Surface enhanced Raman scattering (SERS)

2.1 SERS

Raman scattering of molecule becomes strong on a specific substrate through the interaction between the molecule and the substrate^{4), 5)}. This effect is called SERS, and the substrate that causes SERS is referred to as a SERS substrate. The following two mechanisms have been proposed as the enhancement mechanism.

1. Electro-magnetic Mechanism (EM)¹⁵⁾
2. Chemical Mechanism (CM)¹⁶⁾

1. EM is the enhancement of Raman scattering by the surface plasmon on the substrate, while 2. CM is that through the charge transfer between the substrate and the adsorbed molecule. Each of these mechanisms is roughly explained in the following sections.

2.2 Electro-magnetic Mechanism (EM)

EM is an enhancement mechanism of Raman scattering by the localized surface plasmon (LSP). Figure 1 shows a schematic diagram of this mechanism. Plasmon is a collective vibration of electrons in a solid, and a plasmon that appears on a solid surface is particularly called surface plasmon. This surface plasmon usually appears locally at nanostructures on surfaces. When light irradiates metal surfaces (particularly particles and rough surfaces), LSP is excited and resonates with the incident light that results in the enhancement of the optical field 10^2 – 10^5 times larger than the incident light. This enhanced optical field increases the frequency of Raman scattering. In addition, the Raman-scattered light is also enhanced by resonating with LSP. As a result, the Raman scattering is enhanced in both incident and scattered processes. The enhancement is particularly strong in a few nm gaps between metal nanoparticles (hot spots), at where the enhancement factor of the Raman scattering reaches 10^9 ⁵⁾.

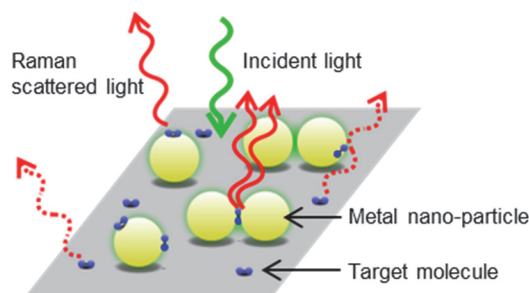


Figure 1 Schematic view of the enhancement of the Raman scattering of molecules by LSP at metal nanoparticles. The green arrow represents the incident light, while the red arrows represent the Raman-scattered light. The enhancement is stronger at gaps between nanoparticles than on nanoparticles.

Surfaces covered with silver or gold nanoparticles are often used as an EM-based SERS substrate, with which enhancement factor of 10^5 – 10^6 is achieved in average. Since this mechanism is applicable to all molecules in principle, the enhancement can be expected regardless of the molecular species. However, the enhancement factor depends on the hot spot, and the resonance with LSP may generate a large amount of heat, which may cause oxidation of the substrate and alteration of the sample molecules. Furthermore, since the substrate is not uniform, it is unclear how and where the molecule adsorbs on the substrate. Therefore, EM-based SERS has disadvantages in terms of reproducibility, uniformity, biocompatibility, etc¹⁷⁾.

2.3 Chemical Mechanism (CM)

CM is the enhancement mechanism by the charge transfer resonance between the substrate and the molecule. In the conventional Raman scattering of molecule, as shown in figure 2(a), electrons in the highest occupied molecular orbital (HOMO) are excited to a virtual state by the incident light (process ①), and then the electrons transit to excited state of molecular vibration by emitting the Raman-scattered light (process ②). When the energy of the incident light is the same as the energy difference between HOMO and lowest unoccupied molecular orbital (LUMO), the Raman scattering becomes strong because electrons are excited to an actual electronic state rather than to a virtual state (figure 2(b)). It is called resonance Raman scattering. On the other hand, in the case of the Raman scattering through the charge transfer resonance, i.e., the CM of SERS, the electrons in the substrate are excited to LUMO by the incident light, and then transit to the vibrational excited state by emitting the

Raman-scattered light (figure 2(c)). The change of the polarization of molecule caused by adsorption to the substrate is also considered to be a cause of CM²⁾.

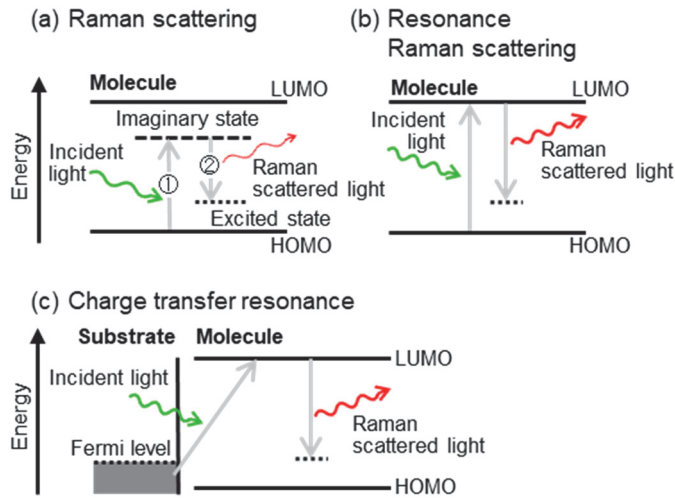


Figure 2 Schematic diagrams of (a) conventional Raman scattering process, (b) resonance Raman scattering process and (c) Raman scattering process by the charge transfer resonance.

2.4 Graphene as a SERS substrate

Although both EM and CM coexist for SERS in general, study about CM has been limited. This is because, in addition to the fact that the enhancement by EM is dominant, it is difficult to prepare a SERS substrate based only on CM, i.e., a flat substrate in nanometer scale, on which LSP is not excited. However, the SERS based only on CM is re-evaluating since the discovery of graphene, an atomically flat two-dimensional material. It has been confirmed that the Raman scattering of molecules such as rhodamine 6G (R6G)^{18), 19)}, phthalocyanine¹⁸⁾, and protoporphyrin IX^{18), 20)} is enhanced by using graphene as a substrate, and that their enhancement factor dependent on the electronic state and the energy of the incident light can be explained by the CM²¹⁾. Although the enhancement factor of the Raman scattering using graphene as a substrate is $\sim 10^3$, which is smaller than that of SERS based on EM, there is no heat generation because of the LSP. Moreover, molecules can be adsorbed on the substrate homogeneously, since graphene is flat in atomic scale. In such senses, graphene can overcome the disadvantages of SERS using metal nanoparticles as a substrate.

3 Raman scattering enhanced on zigzag edges

3.1 Zigzag edge state of graphene

Graphene consists of a honeycomb lattice with hexagons,

in which a carbon atom sp^2 -bonds to other three carbon atoms, as shown schematically in figure 3(a). Reflecting this crystal structure, two atomic structures, i.e., zz and armchair, can be considered for the edge of graphene. The electronic state of graphene is characterized by linear energy dispersion crossing at one point (Dirac point) (black line in figure 3(b)), while a dispersion-free (flat) energy band appears at the Dirac point around the zz edge (red line in figure 3(b)). This is due to the breaking of the honeycomb symmetry of graphene at the zz edge, and the band structure indicates that a localized electronic state appears around the zz edge (zz edge state)⁷⁾⁻⁹⁾. As mentioned in the previous chapter, the electronic state of the substrate plays an important role in the CM-based SERS. Therefore, the enhancement of the Raman scattering through these localized electrons around the zz edge can be expected.

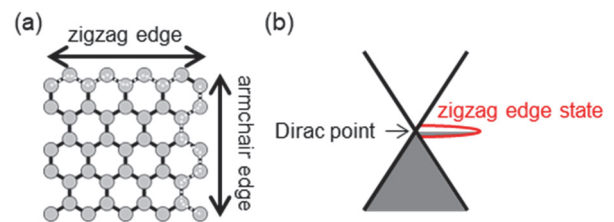


Figure 3 Schematic diagrams of (a) crystal structure of graphene and its edges, (b) the electronic state of graphene. The red line indicates the zigzag edge state.

3.2 Experimental

We have established a technique to create hexagonal nanopits terminated by zz edges on graphene at desired positions^{12), 14)} by hydrogen plasma etching (HPE)⁹⁾⁻¹³⁾. Figure 4 shows an optical microscope image (figure 4(a)) and an atomic force microscope (AFM) image (figure 4(b)) of a graphene device, as an example. In this paper, tri-layer graphene flakes obtained by exfoliating graphite on a SiO_2/Si substrate were used. In this case, high-quality graphene with less defects can be obtained, although the area is limited to $10 \mu m^2$ at most. Both pristine and zGNM regions were fabricated in one flake as shown in figure 4, and the Raman scattering from both regions was obtained in one spectral mapping. This makes it possible to ignore uncertainties such as variation of laser intensity in each measurement, surface contamination, and contact condition between graphene and the substrate. To investigate the enhancement of the Raman scattering by the zz edge state, as long zz edge as possible is

desired within a Raman laser spot ($\sim\phi 500$ nm). Therefore, hexagonal nanopits of ~ 100 nm diagonal length triangularly arranged with 200 nm interval were prepared as a zGNM. Hexagonal holes in the zGNM region shown in figure 4(b) are those fabricated by HPE, and the substrate SiO_2 is exposed inside the nanopits. The edges of the nanopit are the zz edges of graphene.

A laser Raman microscope²²⁾ with a 532 nm laser was used for the Raman scattering measurement. To compare the spectra obtained at different measurements, all Raman spectra are normalized by the intensity of the Si-originated signal that appears at ~ 950 cm^{-1} . Accordingly, the spectral intensity shown in this paper can be considered as the intensity per unit area of a device. The R6G molecule, which is commonly used in the study of SERS, was adopted as a probe molecule to evaluate the enhancement. R6G molecule was adsorbed onto graphene devices by immersing the devices in an R6G solution with a concentration of 0.1, 1, 10, 100 μM ²³⁾ (which corresponds to approximately 0.05, 0.5, 5, 50 ppm, respectively).

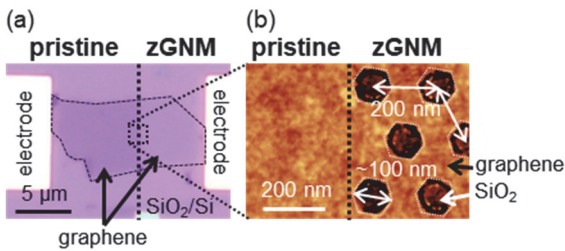


Figure 4 (a) An optical microscope image and (b) an atomic force microscope image of a graphene device.

3.3 Results and discussion

Figure 5 shows the Raman spectra obtained on (a) pristine graphene and on (b) zGNM modified by R6G molecule of 10 μM concentration. The gray dots are the measured data, the colored solid lines are the spectral peaks decomposed by the Bayesian method, which will be introduced in the next chapter, and the black solid line is the sum of the decomposed peaks. In all cases, the spectra are decomposed into peaks named a–m originated from R6G molecule together with the G and D bands of graphene. Since G band is related to the honeycomb structure of graphene, its intensity is proportional to the graphene area. Therefore, G-band is smaller for the mesh structure of zGNM than for pristine graphene. On the other hand, D-band is related to the intervalley scattering inside the crystal, and is proportional to the amount of

defects in the honeycomb structure. As a result, D-band is negligible on pristine graphene, while it appears on zGNM which is micro-fabricated.

Raman scattering of R6G molecule is hardly enhanced on the SiO_2 surface, and any peaks originated from R6G molecule are not observed on the SiO_2 substrate. Therefore, the fact that the peaks originated from R6G molecule are observed on pristine graphene and zGNM indicates that both surfaces have an effect to enhance the Raman scattering of R6G molecule. Among several peaks originated from R6G molecule, a peak named “g” (g-peak) at ~ 1500 cm^{-1} will be focused on in this paper. Figure 6(a) shows the R6G concentration dependence of the g-peak intensity. The peak intensity on pristine graphene (P_g) is shown by an open circle, while that on zGNM (M_g) is shown by a filled circle. Each point is slightly shifted along the horizontal axis to make the graph easier to see. It shows that both P_g and M_g increase with increasing the R6G concentration, and that M_g is larger than P_g at any concentration, namely, the Raman scattering is enhanced stronger on zGNM than on pristine graphene. Considering that no peaks originated from R6G molecule are obtained on the SiO_2 substrate, it is noteworthy that zGNM shows stronger spectral intensity per device area than a substrate fully covered by graphene, even though zGNM has smaller area that contributes to enhance the Raman scattering of R6G molecule. It suggests that there is a strong Raman scattering enhancement at the zz edge around the nanopits, which zGNM has but pristine graphene doesn't.

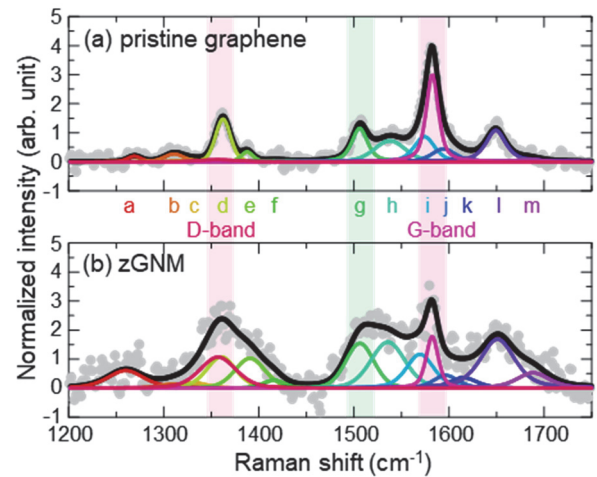


Figure 5 Raman spectra of R6G molecule of 10 μM concentration measured on (a) pristine graphene and on (b) zGNM. The G and D bands originated from graphene and the g-peak among the a–m peaks originated from R6G molecule are marked by red and green, respectively.

In the discussion below, it is assumed that the peak intensity obtained on zGNM is the sum of the contribution of the graphene area and that of the zz edge. By denoting the peak intensity per unit area of pristine graphene and zGNM as P and M , respectively, the peak intensity due to the zz edge per unit edge length as E_0 , the ratio of graphene area in the unit area as r , and the total zz edge length in the unit area as L , above assumption can be expressed as follows.

$$M = r \cdot P + L \cdot E_0 \quad (1)$$

From this equation, contribution of the zz edge to the g-peak (E_{0g}) can be calculated using the g-peak intensities (P_g, M_g) obtained from Raman spectra and r, L measured from AFM images. The resultant E_{0g} depending on the R6G concentration is shown in figure 6(b). It shows that E_{0g} is almost constant at concentrations below 10 μM , and then increases at between 10 and 100 μM . This suggests that the amount of R6G molecule adsorbed on the zz edge is already saturated at a concentration of 0.1 μM . It can be concluded that R6G molecule is adsorbed on basal plane of graphene after selectively adsorbing on the zz edge, and that the second and higher layer promotion starts at a concentration between 10 and 100 μM .

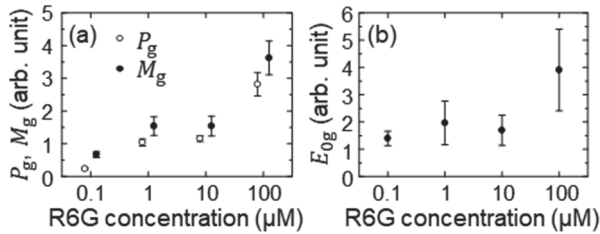


Figure 6 R6G concentration dependence of (a) g-peak intensity on pristine graphene and zGNM (P_g, M_g), and (b) contribution of the zz edge to the g-peak (E_{0g}).

4 Raman spectrum analysis by the Bayesian method

4.1 The Bayesian method

The analysis of spectroscopic data is a trial to reproduce the measured data $y = (y_1, y_2, \dots, y_n)$ for variables $x = (x_1, x_2, \dots, x_n)$ using an estimated value e obtained by a specific function $f(x; \theta)$ under conditions of a certain parameter set θ . Considering that the error d is the difference between the estimated value e and the data y , e and y can be expressed as follows.

$$e = f(x; \theta) \quad (2)$$

$$y = e + d \quad (3)$$

Three methods are generally discussed to find the appropriate parameter set θ . Those are the least-squares method, the maximum-likelihood method, and the Bayesian method.

In the least-squares method, θ is evaluated so that the error d (square-root of d , precisely speaking) is minimized. Here, estimated value e obtained from the parameter set θ is unique, and the difference between e and the measured data y is evaluated. On the other hand, the maximum-likelihood and the Bayesian methods introduce a concept of probability, that is to say, “the probability to take *this* value is the highest”. In the maximum-likelihood method, the error d is thought to be a random variable that follows a specific probability distribution, and it considers the probability $p(y|\theta)$ to obtain measured data y as $f(x; \theta) + d$ for given parameter set θ . Namely, it is a method to find the parameter set θ that maximizes $p(y|\theta)$ (“maximum likelihood”). Since $p(y|\theta)$ is the probability to take the value y under the condition θ , it is called “conditional probability”. It is also called “likelihood” or “likelihood function” depending on the field of research. The least-squares and the maximum-likelihood methods are similar in the sense that both of them evaluate the error d . In fact, their solutions are essentially the same if the error d follows a Gaussian distribution²⁴⁾.

In contrast, the Bayesian method considers the parameter set θ as a random variable. Here, an assumption that “ θ follows the Gaussian distribution centered at $\hat{\theta}$ with a standard deviation of σ ” is given in advance, and then the parameter set θ is re-evaluated using the probability to reproduce the measured value from the estimated value based on the assumption. In other words, compared to the least-squares and maximum-likelihood methods that estimate the result from pre-conditions, the Bayesian method estimates the next result using the experiences. Since this process is essentially the same as a learning process and is easy to handle mathematically, this method is often used in the field of machine learning.

As can be found from this concept, the Bayesian method requires a starting probability distribution $p(\theta)$ for the parameter set θ . This $p(\theta)$ is called “prior (probability) distribution”. By reflecting the measured data on this $p(\theta)$, “posterior (probability) distribution” of the parameter set θ is obtained. Since the posterior distribution is the probability distribution of the parameter set θ based on the measured

data y , it is denoted as the conditional probability $p(\theta|y)$ and expressed as follows (the Bayes' theorem).

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (4)$$

Here, the denominator $p(y)$ is the probability that measured data y occurs, and it is obtained by integrating the numerator over all θ .

$$p(y) = \int p(y|\theta)p(\theta)d\theta \quad (5)$$

Equation (4) means that, " $p(\theta|y)$, which is the probability distribution of parameter set θ based on the measured data y , is the probability that the measured data y occurs when a certain parameter set θ is realized ($= p(y|\theta)p(\theta)$) under a condition that y occurs ($= p(y)$)". Incidentally, the maximum-likelihood method is the way to estimate parameters only from the likelihood by considering that the prior distribution is uniform²⁵⁾.

4.2 The Markov chain Monte Carlo (MCMC) method

As mentioned in the previous section, the Bayesian method calculates the posterior distribution $p(\theta|y)$ using equation (4). In this calculation, although the prior distribution $p(\theta)$ and the likelihood $p(y|\theta)$ can be handled relatively easily, the evaluation of the integral of their multiple (equation (5)) is not straightforward. To overcome this difficulty, some methods are known to simulate the posterior distribution using random numbers, and one of the methods is the Markov chain Monte-Carlo (MCMC) method. This is a method to generate random numbers from an arbitrary probability distribution, and it is based on two ideas. One is the concept of Markov chain that the i -th generated random number θ_i is influenced only by the previous one θ_{i-1} . And the other is the Monte-Carlo method that uses another random number for the random number generation²⁴⁾. The value and the overall feature of $p(y|\theta)p(\theta)$ for each θ can be obtained by substituting N parameter sets $\theta_1, \theta_2, \dots, \theta_N$ generated by MCMC method into $p(y|\theta)p(\theta)$. Furthermore, the integral of $p(y|\theta)p(\theta)$ can also be calculated using a method called Monte-Carlo integration. As a result, the posterior distribution $p(\theta|y)$ can be obtained from equation (4). The simulation can be performed efficiently by using this MCMC method, since parameter sets of higher posterior probability distribution are generated more intensively²⁴⁾.

4.3 Raman spectrum analysis by the Bayesian method

In general, a spectroscopic spectrum can be expressed as a convolution of the Lorentzian distribution, that is inherent in the electro-magnetic waves, and the Gaussian distribution such as the random thermal vibration of crystals and molecules. In this study, a peak in a Raman spectrum is considered as a pseudo-Voigt function $V(x; \theta)$ ²⁶⁾, which is the sum of the Lorentzian and the Gaussian functions. Therefore, the intensity of the Raman spectrum (measured data) y at Raman shift x (cm^{-1}) is expressed as the sum of M peaks, that is, the M pseudo-Voigt functions $\sum V(x; \theta_i)$ and the error d .

$$y = \sum_{i=1}^M V(x; \theta_i) + d = f(x; \theta) + d \quad (6)$$

Here, the parameter set θ_i for the i -th peak includes the spectrum peak intensity I_i , peak position μ_i , full width at half maximum (FWHM) Γ_i , and Gaussian ratio a_i . The Gaussian ratio a is the mixing ratio of the Gaussian and the Lorentzian functions, and $a = 0$ (1) makes the pseudo-Voigt function the Lorentzian (Gaussian) function. The error d is assumed to be a Gaussian distribution with a mean of 0 and a standard deviation of σ .

The likelihood $p(y|\theta)$ and the prior distribution $p(\theta)$ are considered based on the above assumptions. Since the measured data y follows the Gaussian distribution with a mean value of $f(x; \theta)$ and a standard deviation of σ according to equation (6), the likelihood is expressed as

$$p(y|\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - f(x_i; \theta))^2}{2\sigma^2}\right). \quad (7)$$

On the other hand, the prior distribution $p(\theta)$ must be considered for all parameters. In this study, the peak intensity I and the Gaussian ratio a are considered to be a uniform distribution with the range of $0 \leq I \leq A$ and $0 \leq a \leq 1$, respectively. Here, A is a sufficiently large peak intensity. Fifteen peaks in total were considered as the peaks which constitute the Raman spectrum. Those are the R6G molecule originated 13 peaks named a–m referred to the previous researches^{19), 27), 28)}, and the G and D bands of graphene. Their peak position μ (cm^{-1}) is assumed to be a Gaussian distribution with a mean of $\hat{\mu}$ (cm^{-1}) referred to previous research²⁷⁾ and a standard deviation of 5 (cm^{-1}). The FWHM Γ (cm^{-1}) is assumed to be a Gaussian distribution with a mean of 30 (cm^{-1}) and a standard deviation of 10 (cm^{-1}) based

on our own experience. Afterwards, the prior distribution $p(\theta)$ can be written as

$$p(\theta) = \prod_j \frac{1}{\sqrt{50\pi}} \exp\left(-\frac{(\mu_j - \hat{\mu}_j)^2}{50}\right) \cdot \frac{1}{\sqrt{200\pi}} \exp\left(-\frac{(\Gamma_j - 30)^2}{200}\right) \cdot \frac{1}{1} \cdot \frac{1}{A}. \quad (8)$$

Now, since the prior distribution and the likelihood function are defined, the posterior distribution can be estimated using the MCMC method. No-U-Turn-Sampler implemented in PyMC, that is a Python MCMC library, was used to execute the MCMC method.

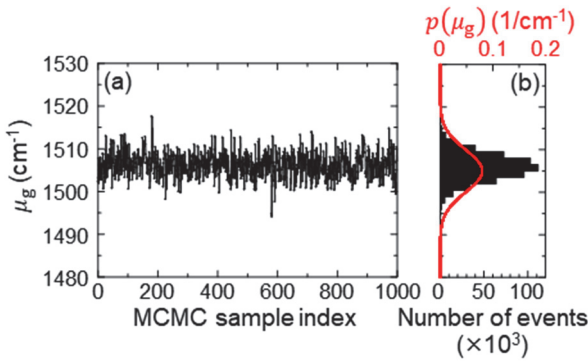


Figure 7 (a) 1,000 of μ_g obtained using the MCMC method. (b) Histogram of μ_g (black) and its prior distribution (red).

Figure 7 shows the g-peak position μ_g among 100,000 parameter sets θ obtained using the MCMC method for Raman spectrum of zGNM modified with R6G molecule of 10 μM concentration (figure 5(b)). Figure 7(a) shows 1000 of μ_g out of them, and the histogram of all 100,000 is shown by black lines in figure 7(b). It can be seen that μ_g takes various values probabilistically and converges within a certain range. In comparison with the prior distribution $p(\mu_g)$ shown by the red line in figure 7(b) (center at 1505 cm^{-1} with $\sigma = 5 \text{ (cm}^{-1}\text{)}$), the posterior distribution obtained using the measured data is found to have smaller variance, and the center value is shifted from $p(\mu_g)$.

Although the posterior distribution of each parameter can be estimated using the Bayesian method in such way, it is not straightforward to choose an optimal value of the distribution. In general, the maximum a posteriori (MAP) estimate, which is the value most frequently appears in the posterior distribution, and the expected a posteriori (EAP) estimate, which is the mean value of the posterior distribution, are discussed²⁵⁾. In this study, the EAP estimate is adopted.

4.4 Evaluation of the Raman spectrum analysis

Finally, the validity of the spectrum analysis performed in this study is evaluated. Figure 8 compares the Bayesian method (figure 8(a)) and the least-squares method (figure 8(b)). Note that figure 8(a) is reiteration of figure 5(b). At first glance, the least-squares method also fits the data well. However, some peaks are overestimated, the peak positions are shifted significantly, and obvious noise is evaluated as a peak. In contrast, such mis-analysis does not occur and the spectrum seems to be decomposed more appropriately in analysis using the Bayesian method.

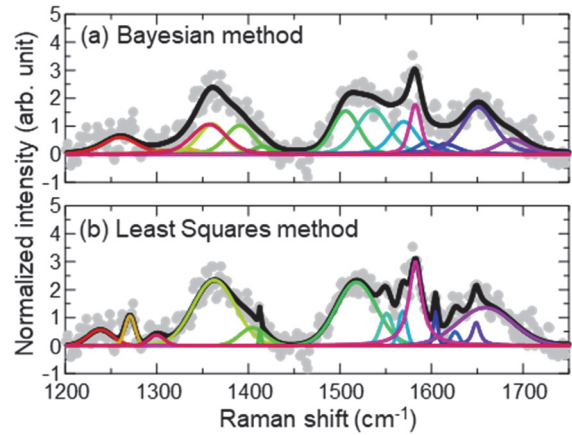


Figure 8 Results of peak decomposition of the Raman spectrum (gray dot) of R6G molecule of 10 μM concentration on a zGNM substrate using (a) the Bayesian method and (b) the least-squares method.

The validity of the analysis can also be confirmed by considering the ratio of the graphene area in the pristine and the zGNM regions. As mentioned before, the peak intensity of the G-band in Raman spectrum is roughly proportional to the area of graphene. The G-band intensities of pristine and zGNM regions, and the ratio of them, extracted from the data in figure 8, and the areal ratio of graphene per unit device area evaluated from AFM images are shown in table 1. It can be confirmed that the areal ratio estimated from the G-band intensity is within about ± 0.1 of the actual value before modifying the substrate with R6G molecule (w/o R6G). The ratio of the G-band intensity between the pristine graphene and the zGNM is expected to take a similar value as far as the graphene area does not change by R6G modification (w/ R6G). Actually, the G-band intensity ratio obtained using the Bayesian method after R6G modification is almost the same as the actual areal ratio. On the other hand, in the case of the least-squares method, the intensity ratio changes

significantly, and the zGNM shows larger G-band intensity than the pristine graphene, although the zGNM is a mesh structure. This also suggests that the analysis using the least-squares method is not valid, while the Bayesian method is appropriate, at least, in this study.

Table 1 The G-band intensity in the Raman spectra shown in figure 5 and 8 and their ratio, together with the areal ratio of graphene per unit device area evaluated from AFM images of the device.

G-band intensity		Pristine graphene (P_G)	zGNM (M_G)	ratio (M_G/P_G)
w/o R6G		3.96	1.58	0.42
w/ R6G	Bayesian method	3.00±0.63	1.78±0.41	0.59±0.19
	Least Squares method	2.91	3.02	1.04
Graphene area ratio of zGNM estimated from AFM images				0.52

5 Summary

In this study, we focused on the SERS by the chemical mechanism (CM) based on the charge transfer, and investigated the enhancement of the Raman scattering via electrons localized at the zz edge of graphene. By measuring the Raman spectra of R6G molecule adsorbed on pristine graphene and zGNM with zz edges, it was found that the intensity of the Raman scattering originated from the R6G molecule is stronger on zGNM than on pristine graphene at any R6G concentration. This suggests that the Raman scattering is enhanced on zz edges of zGNM. It was also found by evaluating the contribution of the zz edge to the Raman scattering that the R6G molecule is selectively adsorbed on the zz edge.

Since the R6G molecule shows many peaks within the measurement range in the Raman spectrum, the least-squares method, which is often used in spectrum analysis, does not always decompose the spectrum into peaks appropriately. Therefore, the Bayesian method, which is commonly used in the field of the machine learning, was used in this study.

This study reveals that the Raman scattering is enhanced through the zz edge states of graphene. Since the SiO₂ substrate, which is exposed inside the nanopits, does not enhance the Raman scattering, zGNM loses the area that

contributes to enhance the Raman scattering. If zGNM could be transferred onto pristine graphene, stronger enhancement of the Raman scattering can be achieved because Raman scattering is enhanced inside the nanopits, too. Furthermore, more practical enhanced Raman scattering substrate could be expected, if chemically grown graphene with large surface area could be fabricated into zGNM and transferred onto graphene on a flexible transparent film.

Acknowledgement

A part of this work was supported by “Nanotechnology Platform Japan” and “Advanced Research Infrastructure for Materials and Nanotechnology in Japan (ARIM)” of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Grant Number JPMXP1222UT1064, JPMXP1223UT1029, and fabrication was conducted in Takeda Cleanroom with help of Nanofabrication Platform Center of School of Engineering, the University of Tokyo.

References

- 1) A. F. Chrimes *et al.*, Chem. Soc. Rev. **42**, 5880 (2013).
- 2) P. M. Pancorbo *et al.*, EPL **136**, 34001 (2021).
- 3) Y. Kitahama *et al.*, Anal. Methods, **15**, 1028 (2023).
- 4) M. Fleischmann *et al.*, Chem. Phys. Lett. **26**, 163 (1974).
- 5) J. Langer *et al.*, ACS Nano **14**, 28 (2020).
- 6) G. D’Agostini, Rep. Prog. Phys. **66**, 1383 (2003).
- 7) M. Fujita *et al.*, J. Phys. Soc. Jpn. **65**, 1920 (1996).
- 8) Y. Niimi *et al.*, Phys. Rev. B **73**, 084521 (2006).
- 9) Tomohiro Matsui, Anritsu Technical Review **29**, 27 (2021).
- 10) A.E.B. Amend, *et al.*, e-J. Surf. Sci. Nanotech. **16**, 72 (2018).
- 11) T. Matsui *et al.*, J. Phys. Chem. C **123**, 22665 (2019).
- 12) Masahiro Kamada *et al.*, Anritsu Technical Review **30**, 28 (2022).
- 13) T. Ochi *et al.*, Carbon **203**, 727 (2023).
- 14) T. Yokosawa *et al.*, e-J. Surf. Sci. Nanotechnol. **20**, 139 (2022).
- 15) M. Moskovits, Rev. Mod. Phys. **57**, 783 (1985).
- 16) J. R. Lombardi *et al.*, J. Chem. Phys. **84**, 4174 (1986).
- 17) C. Zong *et al.*, Chem. Rev. **118**, 4946 (2018).
- 18) X. Ling *et al.*, Nano Lett. **10**, 553 (2010).
- 19) Y. Wang *et al.*, Appl. Phys. Lett. **99**, 233103 (2011).
- 20) X. Ling & J. Zhang, small **6**, 2020 (2010).
- 21) X. Ling *et al.*, Acc. Chem. Res. **48**, 1862 (2015).
- 22) inVia Raman microscope, Renishaw plc.
- 23) M means mol/l.

- 24) Masanori Hanada, So Matsuura, MCMC from Scratch: A Practical Introduction to Markov Chain Monte Carlo, Kodansha, ltd. (2020).
- 25) Ken-ichi Ishii, Naonori Ueda, “Zoku-Wakariyasui pattern ninshiki”, Ohmsha, ltd. (2014).
- 26) Hideo Toraya, Journal of the Crystallographic Society of Japan **34**, 86 (1992).
- 27) A. M. Michaels *et al.*, J. Am. Chem. Soc. **121**, 9932 (1999).
- 28) H. Watanabe *et al.*, J. Phys. Chem. B **109**, 5012 (2005).

Authors



Taisuke Ochi
Matsui Laboratory
Advanced Research Laboratory



Tomohiro Matsui
Matsui Laboratory
Advanced Research Laboratory

Overview of design method of millimeter-wave direct coupled resonator filters

Takahiko Mizuno

[Summary]

In recent years, the millimeter wave frequency band above 300 GHz has been developed for the practical application of next generation mobile communications. However, at a high frequency region such as a millimeter wave, several types of losses become pronounced, including dielectric losses of the substrate material. Thus, these losses limit the performance of analog frontend modules such as a millimeter-wave filter. This paper describes the design method for the type of direct coupled resonator as a typical waveguide filter. An example of design in the 300 GHz is compared with some previously reported microstrip filters focused on losses and filter performances. The limiting factors of filter performance and the necessity of precision machining for the waveguide filter processing are also discussed.

1 Introduction

The next generation information societies will consist of some key technologies: the Internet of Things (IoT), which enables big data collection and utilization, and the practical applications of virtual reality (VR) and augmented reality (AR). To realize these key technologies, high performance communication techniques, such as high speed, high capacity, low latency, and multiple simultaneous connections, will play an essential role^{1), 2)}. Current mobile communication technology, such as 5G, has begun utilizing not only microwave bands like the Sub 6 band but also millimeter-wave bands starting at 28 GHz, aiming to expand communication bandwidth. Furthermore, the next generation 6G is targeting even higher speeds and capacities and an increase in the number of connections by deploying in higher frequency millimeter-wave bands, such as the 300 GHz band. On the other hand, various applications of millimeter-wave sensing, which utilizes higher penetration compared to light and higher spatial resolution compared to microwaves, are also being explored^{3), 4)}. As such, the millimeter-wave band is actively being considered for usage. However, the design of passive devices in this frequency band requires a different approach than in the microwave band. This article outlines the design method of filters, those are the most important passive devices in the millimeter-wave range.

In the millimeter-wave range, circuit dimensions become comparable to the wavelength, and no inductors or capacitors are available with sufficient characteristics.

Hence, the filter design should be based on distributed element circuits consisting of the transmission line. Figure 1 shows typical transmission lines for microwave and millimeter-wave, those are utilized to constitute distributed element circuits. Filters consist of these transmission lines are compact such as less than several centimeters in the millimeter-wave region, whereas they are tens of centimeters in the microwave region. However, the decrease in size not only increases conductor losses, but also increases demands on manufacturing precision. The hollow waveguide shown in the upper part of Figure 1 has a low loss. However, due to its high cost of manufacture and large size in the microwave region, it has been limited to applications where high power handling is necessary. In the microwave region, a transmission line processed on a dielectric substrate, such as a microstrip, is widely used, as shown in the bottom part of Figure 1. These structures are easier to fabricate compared to hollow waveguides. However, the dielectric loss increases in a high frequency region such as the millimeter-wave.

Thus, appropriate filter geometries should be selected with consideration to both losses and ease of manufacture. Particularly in the millimeter-wave region, planar circuit filters on dielectric substrates are preferred when emphasizing integration density and ease of manufacture. In contrast, hollow waveguide filters are suitable when prioritizing low loss properties. In addition, the hollow waveguide filter can be directly connected to commonly used waveguides in the millimeter-wave band, thereby avoiding

waveguide type conversion losses. It is important in the front end of millimeter-wave measurement equipment that handles small level signals.

This article reviews the design procedure for bandpass filter (BPF) for millimeter-wave region and presents a design example of a 300 GHz band directly coupled resonator BPF based hollow waveguides. The designed BPF is compared with previously reported filters, and gains or losses of planar circuit filters on dielectric substrates versus the hollow waveguide filter, as well as the limiting factors of filter performance, are discussed.

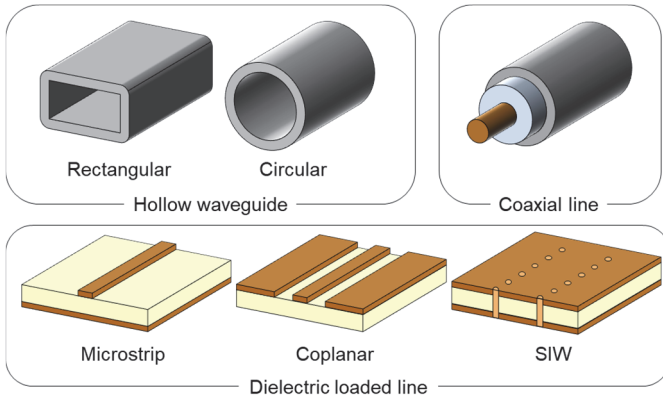


Figure 1 Various transmission lines used in the microwave and millimeter-wave bands

2 Design method of direct coupled resonator bandpass filter

BPF is one kind of passive circuit that allows only signals from a specific frequency band to pass through. For a method of its design, the insertion loss method is usually used, which was developed by G. Matthaei, L. Young, E. Jones, S. Cohn, and others in the 1950s. This method is based on the synthesis theory of networks and aims to achieve the desired frequency characteristics (insertion loss) by determining the circuit configuration and element values. In this procedure, first, a prototype low-pass filter (LPF) with normalized frequency and source impedance (admittance) is designed. Then, the circuit is transformed for ease of construction, followed by frequency scaling to convert it into a BPF. Inductors and capacitors deviate significantly from their ideal characteristics at the millimeter-wave range; hence, a discontinuity on a transmission line with reactance equivalent to these components is used to construct the filter. Electromagnetic field analysis is used to determine the

correspondence between the discontinuity dimensions on the line and the equivalent reactance. This relationship determines the structural dimensions that achieve the reactance designed by the insertion loss method. This chapter provides an overview of the above design procedure. For more details, please refer to references 5) to 8).

2.1 Frequency characteristics of bandpass filter

Typical specifications of frequency characteristics required for BPF include passband frequencies $f_{\text{pass,L}}$ and $f_{\text{pass,H}}$ with transmittance $|H_{\text{p,max}}|$, as well as stopband frequencies $f_{\text{stop,L}}$ and $f_{\text{stop,H}}$ with attenuation $|H_{\text{s,min}}|$. Figure 2 shows the typical relationship between BPF specifications and frequency characteristics. Filter design parameters such as cutoff frequencies f_1 , f_2 , and filter order N are defined to satisfy these specifications. A larger filter order provides a steep slope of the amplitude pass characteristic from near the cutoff frequency to the stopband.

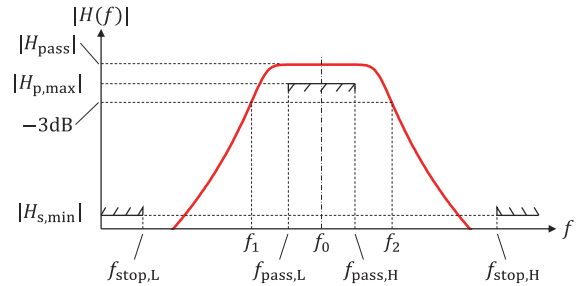


Figure 2 Typical BPF characteristics

2.2 Design of prototype LPF

A prototype LPF is a basic filter shape with the source impedance (or admittance) and the cutoff frequency normalized to unity. Figure 3 shows an equivalent ladder network of the prototype LPF, which consists of inductors and capacitors. Normalized element values (L'_i and C'_i) are represented as g_i . Figure 3 shows a sixth order design; the first reactance element is set to an inductor with a value of g_1 . $g_0 (= 1)$ and g_7 are the admittance of the source and the load, respectively. The same characteristics can also be composed with the other form of prototype LPF; the first reactance is set to a parallel capacitor (i.e., $g_1 = C'_1$), and every g_i represents impedances. These g parameters are set appropriately to obtain the required frequency response. Table 1 lists the typical frequency characteristics of some prototype filters, and Figure 4 shows the typical amplitude pass characteristics of each prototype LPF.

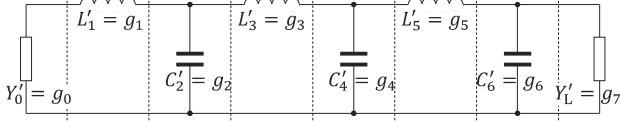


Figure 3 Prototype LPF equivalent circuit

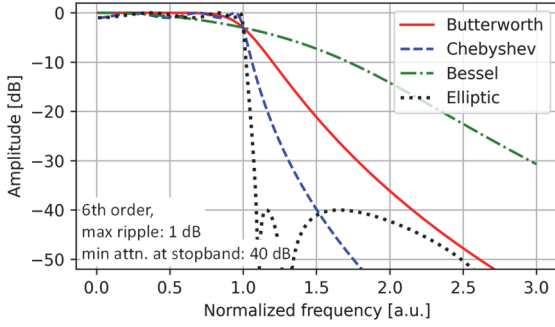


Figure 4 Amplitude pass characteristics of typical prototype LPFs

Table 1 Typical filter characteristics

Characteristic	Advantage	Disadvantage
Butterworth (maximally flat)	Highly flat response in passband	Not steep attenuation slope
Chebyshev (equiripple)	Steep attenuation slope	Passband is not flat, has ripple
Bessel (linear phase)	Linear phase change in passband, good temporal response	Very gradual attenuation slope, low frequency selectivity
Elliptical (equiripple in pass- and stop- band)	Attenuation slope can be steeply with placing attenuation pole properly.	Complicated structure due to interlaced coupling

In this paper, a Butterworth prototype LPF is used as an example. A Butterworth prototype LPF can be obtained by setting the element value g_i in Figure 3 to,

$$\begin{aligned} g_0 &= 1, \\ g_i &= 2 \sin \left[\frac{(2i-1)\pi}{2N} \right], \quad i = 1, 2, \dots, N, \\ g_{N+1} &= 1, \end{aligned} \quad (1)$$

where, N is the order of the filter, which corresponds to the number of ladder stages. The cutoff slope is determined to be $6N$ dB/oct., which corresponds to the filter order N . For example, in a sixth order Butterworth prototype LPF, the g parameter is shown:

$$\begin{aligned} g_0 = g_7 &= 1.0000, & g_1 = g_6 &= 0.5176, \\ g_2 &= g_5 &= 1.4142, & g_3 = g_4 &= 1.9319. \end{aligned}$$

The transfer function $H(s)$ of the Butterworth prototype LPF with such a g parameter set can be expressed as below,

$$H(s) = \begin{cases} \prod_{i=1}^{N/2} (s^2 + g_i s + 1)^{-1} & N: \text{even}, \\ \frac{1}{s+1} \prod_{i=1}^{(N-1)/2} (s^2 + g_i s + 1)^{-1} & N: \text{odd}. \end{cases} \quad (2)$$

2.3 Admittance Inverter

The prototype LPF shown in Figure 3 consists of both series and parallel connections of reactance elements. To facilitate implementation, use impedance inverters or admittance inverters to standardize either series or parallel connections. Here, we use an admittance inverter to convert the series inductor shown in Figure 5(a) to a parallel capacitor. An admittance inverter is a two-port network whose input admittance is J^2/Y_L when a load of admittance Y_L is connected. Here, J is called the admittance parameter (J parameter), and the J inverter's F matrix (transfer matrix) is shown,

$$F_J = \begin{bmatrix} 0 & 1 \\ -jJ & 0 \end{bmatrix}.$$

Here, the F matrix of the circuits shown in Figure 5(a) and (b) are below, respectively,

$$F_a = \begin{bmatrix} 1 & j\omega L \\ 0 & 1 \end{bmatrix}, \quad F_b = \begin{bmatrix} -\frac{J_2}{J_1} & -j\frac{\omega C_a}{J_1 J_2} \\ 0 & -\frac{J_1}{J_2} \end{bmatrix}. \quad (3)$$

If $F_a = F_b$, the following relation is obtained,

$$J_1 = -J_2, \quad L = -\frac{C_a}{J_1 J_2} = \frac{C_a}{J_1^2} = \frac{C_a}{J_2^2}.$$

Under the constraints imposed by this equation, circuits in Figure 5(a) and (b) can be made equivalent. Furthermore, as a general property of the circuit in Figure 5(b), the diagonal elements of F_b are not independent of each other, so F_b has two degrees of freedom: the (1,1) element and the (1,2) element. On the other hand, since the circuit has three parameters, J_1 , J_2 , and C_a , J_1 and J_2 can be determined such that $F_a = F_b$ for any C_a .

Therefore, the series inductors (L'_1 , L'_3 , L'_5) are shown in Figure 3 can be equivalently transformed into two J inverters and a parallel capacitor. Furthermore, all capacitances, including C'_2 , C'_4 , and C'_6 can be changed arbitrarily using the degree of freedom of J which was described above. As a result, the prototype LPF shown in Figure 3 can be converted into a circuit consisting only of parallel capacitors and J inverters, as shown in Figure 6. A J parameter $J'_{i,i+1}$ is given as follows for an arbitrary capacitance $C'_{a,i}$,

$$\begin{aligned} J'_{0,1} &= \sqrt{\frac{Y'_0 C'_{a,1}}{g_0 g_1}}, \\ J'_{i,i+1} &= \sqrt{\frac{C'_{a,i} C'_{a,i+1}}{g_i g_{i+1}}}, \quad i = 1, 2, \dots, N-1, \\ J'_{N,N+1} &= \sqrt{\frac{C'_{a,N} Y'_L}{g_N g_{N+1}}}. \end{aligned} \quad (4)$$

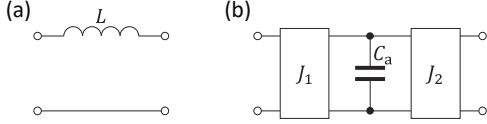


Figure 5 Circuit conversion by J inverter:
(a) series inductor, (b) capacitor between two J inverters

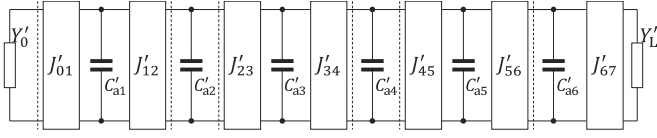


Figure 6 Equivalent circuit of prototype LPF with J inverter

2.4 Conversion from prototype LPF to BPF

Admittances are scaled from the prototype LPF shown in Figure 6 to the objective BPF. This scaling is performed by the conversion of the normalized angular frequency Ω in the prototype LPF and the angular frequency ω of the BPF as follows,

$$\Omega = \frac{\Omega_c}{\Delta} \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right) = \frac{1}{\Delta} \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right). \quad (5)$$

Where ω_0 is the central angular frequency, which is expressed using cutoff angular frequencies $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$ as follows,

$$\omega_0 = \sqrt{\omega_1 \omega_2}. \quad (6)$$

Then, the fractional bandwidth Δ is represented as below,

$$\Delta = \frac{\omega_2 - \omega_1}{\omega_0}. \quad (7)$$

According to this conversion, when $\omega > \omega_0$, it exhibits characteristics similar to the original LPF, and when $\omega < \omega_0$, it exhibits high pass characteristics in which the LPF is inverted at ω_0 . The cutoff angular frequency Ω_c of the prototype LPF corresponds to the cutoff angular frequency of objective BPF via the fractional bandwidth Δ .

The parallel capacitance C'_a in Figure 6 is converted as shown in Figure 7 using Eq. (5), then C_r and L_r are shown as follows,

$$\begin{aligned} j\Omega C'_a &= j\omega \frac{\Omega_c}{\omega_0 \Delta} C'_a + \frac{1}{j\omega} \frac{\omega_0 \Omega_c}{\Delta} C'_a, \\ C_r &= \frac{\Omega_c}{\omega_0 \Delta} C'_a, L_r = \frac{\Delta}{\omega_0 \Omega_c} \frac{1}{C'_a}. \end{aligned} \quad (8)$$

Then Eq. (4) is converted as follows,

$$\begin{aligned} J_{0,1} &= \sqrt{\frac{\omega_0 \Delta Y_0 C_{r,1}}{\Omega_c g_0 g_1}}, \\ J_{i,i+1} &= \frac{\omega_0 \Delta}{\Omega_c} \sqrt{\frac{C_{r,i} C_{r,i+1}}{g_i g_{i+1}}}, \quad i = 1, 2, \dots, N-1, \\ J_{N,N+1} &= \sqrt{\frac{\omega_0 \Delta C_{r,N} Y_L}{\Omega_c g_N g_{N+1}}}. \end{aligned} \quad (9)$$

Note that each capacitance $C_{r,i}$ can be set arbitrarily. This procedure provides a BPF consisting of the parallel LC resonant circuit and J inverter shown in Figure 8.

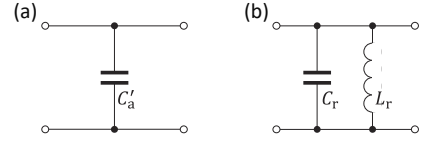


Figure 7 Circuit element conversion from prototype LPF to BPF:
(a) parallel C in prototype LPF, (b) parallel LC resonant circuit in BPF

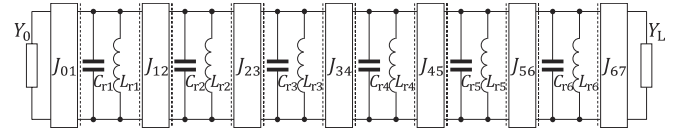


Figure 8 BPF consisting of parallel LC resonant circuit and J inverters

2.5 Conversion from lumped-element circuit to distributed-element circuit

The parallel LC resonant circuit in Figure 8 can be approximated using a cavity resonator constructed on a distributed element circuit. For this approximation, the susceptance of the cavity resonator must be equal to those of the parallel LC resonant circuit, and each first order derivative factor with respect to the frequency of the susceptance at the resonant frequency also must be equal to these simultaneously. The susceptance is defined as $B(\omega)$, and the slope parameter b respected to ω is defined as follows,

$$b = \frac{\omega_0}{2} \frac{dB(\omega)}{d\omega} \Big|_{\omega=\omega_0}. \quad (10)$$

Figure 9 shows a case of a transmission line used as a resonator. The susceptance of the parallel LC resonant circuit in Figure 9(a) is shown as follows,

$$B(\omega) = \omega C_r - \frac{1}{\omega L_r}, \quad \frac{dB(\omega)}{d\omega} = C_r \left(1 + \frac{\omega_0^2}{\omega^2} \right).$$

Since $B(\omega_0) = 0$ at the resonant angular frequency $\omega = \omega_0$, b was determined to,

$$b = \omega_0 C_r = \frac{1}{\omega_0 L_r}. \quad (11)$$

J parameters of each J inverters are determined by substituting Eq. (11) into Eq. (9), as follows,

$$\begin{aligned} J_{0,1} &= \sqrt{\frac{\Delta Y_0 b_1}{\Omega_c g_0 g_1}}, \\ J_{i,i+1} &= \frac{\Delta}{\Omega_c} \sqrt{\frac{b_i b_{i+1}}{g_i g_{i+1}}}, \quad i = 1, 2, \dots, N-1, \\ J_{N,N+1} &= \sqrt{\frac{\Delta b_N Y_L}{\Omega_c g_N g_{N+1}}}. \end{aligned} \quad (12)$$

On the other hand, for the transmission line with characteristic admittance Y_0 , phase constant β , length l , and open ended shown in Figure 9(b), the admittance $Y_{\text{line}}(\omega)$ seen from the input at angular frequency ω is shown as follows,

$$Y_{\text{line}}(\omega) = jY_0 \tan[\beta(\omega)l]. \quad (13)$$

Near the resonant angular frequency $\omega = \omega_0 + \delta\omega$ of the half-wavelength resonator is configured, assuming the phase constant β is linear around ω , the β is shown,

$$\beta(\omega_0)l = \pi, \quad \beta(\omega_0 + \delta\omega)l = \pi \left(1 + \frac{\delta\omega}{\omega_0}\right). \quad (14)$$

The first order derivation of the susceptance around ω_0 is shown,

$$\left. \frac{dB(\omega)}{d\omega} \right|_{\omega=\omega_0} = \frac{1}{j} \left. \frac{dY_{\text{line}}(\omega)}{d\omega} \right|_{\omega=\omega_0} \approx \frac{\pi Y_0}{\omega_0}. \quad (15)$$

From Eq. (10),

$$b \approx \frac{\pi}{2} Y_0, \quad (16)$$

is obtained. J parameters using resonator cavity are derived from Eq. (12) with $\Omega_c = 1$ and $Y_L = Y_0$, as follows,

$$\begin{aligned} \frac{J_{0,1}}{Y_0} &= \sqrt{\frac{\pi\Delta}{2} \frac{1}{g_0 g_1}}, \\ \frac{J_{i,i+1}}{Y_0} &= \frac{\pi\Delta}{2} \sqrt{\frac{1}{g_i g_{i+1}}}, \quad i = 1, 2, \dots, N-1, \\ \frac{J_{N,N+1}}{Y_0} &= \sqrt{\frac{\pi\Delta}{2} \frac{1}{g_N g_{N+1}}}. \end{aligned} \quad (17)$$

Furthermore, the J inverter can be converted to the π -type circuit shown in Figure 10. The J parameter normalized by characteristic admittance Y_0 is expressed as below⁵⁾,

$$\begin{aligned} \frac{J}{Y_0} &= \left| \tan \left[\frac{\phi}{2} + \tan^{-1} \left(\frac{B_p}{Y_0} \right) \right] \right|, \\ \phi &= -\tan^{-1} \left(2 \frac{B_s}{Y_0} + \frac{B_p}{Y_0} \right) - \tan^{-1} \left(\frac{B_p}{Y_0} \right). \end{aligned} \quad (18)$$

Here, B_p and B_s are the series susceptance and parallel susceptance of the π -type circuit, and ϕ represents the phase delay due to the π -type circuit. The line length l_i of the i -th resonator, when phase delay is considered, is represented by the equation below,

$$l_i = \frac{\lambda_{g0}}{2\pi} \left[\pi + \frac{1}{2} (\phi_i + \phi_{i+1}) \right], \quad i = 1, 2, \dots, N. \quad (19)$$

Here, λ_{g0} is the guide wavelength of the center frequency.

By conversion of the above procedure, Figure 8 is replaced with a $\lambda_g/2$ resonator and a π -type equivalent circuit, as shown in Figure 11.

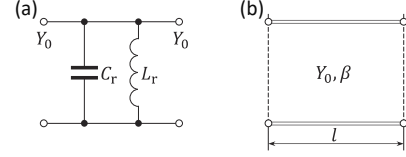


Figure 9 Conversion of a LC resonant circuit to a distributed-element circuit:

(a) parallel LC resonant circuit, (b) cavity resonator based on a waveguide with the electrical length π

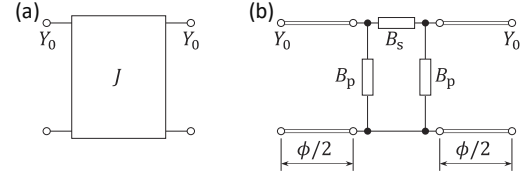


Figure 10 Example of equivalent circuit of J inverter:

(a) four-terminal model of J inverter, (b) symmetric π -type equivalent circuit containing a waveguide with the electrical length ϕ

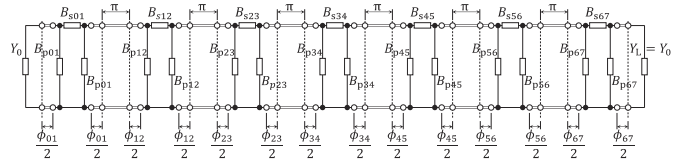


Figure 11 A BPF circuit constructed by $\lambda_g/2$ cavity resonators and π -type circuits

2.6 Discontinuous structure on waveguide and its electrical characteristics

The electrical characteristics of a discontinuous structure on a waveguide can be approximated by a π -type equivalent circuit. In this section, the inductive window in the rectangular waveguide is approximated by a π -type circuit. An inductive window in a rectangular waveguide has a conductor partition plate, which is inserted from either of the left and right sides, or both sides of a waveguide. With a sufficiently thin plate, its characteristics can be approximated by a parallel inductor. Figure 12 shows the procedure of calculating the parameters of a π -type equivalent circuit for an inductive window using an electromagnetic field (EMF) simulator (e.g., 3D EMF simulator based on a finite element method). As shown in Figure 12(a), a model is created for an inductive window with a width w and thickness t . The scattering matrix S is determined by an EMF simulator for each dimension parameter (w in here) of the discontinuous structure as in Figure 12(b). The susceptance of the π -type circuit can be determined from the S using the following equation (Figure 12(c)),

$$j \frac{B_p}{Y_0} = \frac{1 - S_{12} - S_{11}}{1 + S_{11} + S_{12}}, \quad (20)$$

$$j \frac{B_s}{Y_0} = \frac{2S_{12}}{(1 + S_{11})^2 - S_{12}^2}.$$

The width of the inductive window dimension (width: w) can be determined from the arbitrary equivalent susceptance.

With the above preparations, the dimensions of inductive windows w and cavity lengths l are determined as follows. J and ϕ can be found from Eq. (18) when w is given. Therefore, it is possible to find w such that J matches with $J_{i,i+1}$ designed from Eq. (17). Since J is a coupling coefficient between resonators, there always exists w that gives any J in the range of 0 to 1. Once the aperture width w is determined, the phase ϕ can also be found from Eq. (18), and the cavity length l of the $\lambda_g/2$ resonator can also be obtained from Eq. (19).

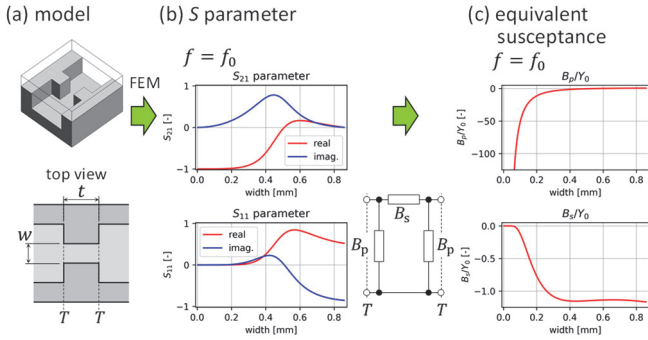


Figure 12 Equivalent parameter calculation process of a discontinuous structure using EMF simulator:

- (a) inductive window in rectangular waveguide (cut model),
- (b) S parameters obtained from EMF simulation,
- (c) equivalent susceptance in a π -type circuit

3 Design example of millimeter-wave filter and EMF simulation

3.1 Design of 300 GHz-band direct coupled resonator BPF

Table 2 shows the specifications and design parameters of the designed 300 GHz-band BPF. For the designed specifications, a center frequency and a passband width were set to 300 GHz and 10 GHz, respectively. The BPF was constructed in a WR-3.4 standard rectangular waveguide (operating frequency range is 220 to 330 GHz). A sixth order Butterworth filter was selected to satisfy the specifications, and cutoff frequencies were set to $f_1 = 293$ GHz and $f_2 = 307$ GHz. Following the design procedure in Chapter 2, direct coupled six $\lambda_g/2$ cavity resonators, which are

separated by seven inductive windows in the waveguide, were designed. The thickness of inductive windows was set to $t = 0.2$ mm. Table 3 shows the dimensions of each part determined through the design process. Figure 13 shows a cut-model of the designed BPF. The filter in which some resonators are directly coupled, as shown in Figure 13, hence it is called a direct coupled resonator filter.

Table 2 Design specifications and parameters

Required specifications			
Waveguide: WR-3.4 ($W \times H = 0.864 \times 0.432$ mm)			
$f_{\text{pass,L}}$: 295 GHz	$f_{\text{pass,H}}$: 305 GHz	$ H_{\text{p,max}} $	>-5 dB
$f_{\text{stop,L}}$: 280 GHz	$f_{\text{stop,H}}$: 320 GHz	$ H_{\text{s,min}} $	<-40 dB
Design parameters			
Filter type: sixth order Butterworth BPF			
f_1 : 293 GHz		f_2 : 307 GHz	
Window type: inductive window, $t = 0.2$ mm			

Table 3 Designed parameters and dimensions of Butterworth BPF using inductive window on WR-3.4 waveguide

	Width [mm]	Length [mm]	J / Y_0	ϕ [rad.]
Window 0-1/ Cavity 1	0.5094	0.3939	0.2466	1.8264
Window 1-2/ Cavity 2	0.3749	0.4992	0.0368	2.6052
Window 2-3/ Cavity 3	0.3268	0.5248	0.0190	2.7286
Window 3-4/ Cavity 4	0.3155	0.5248	0.0163	2.7524
Window 4-5/ Cavity 5	0.3268	0.4992	0.0190	2.7286
Window 5-6/ Cavity 6	0.3749	0.3939	0.0368	2.6052
Window 6-7	0.5094	—	0.2466	1.8264

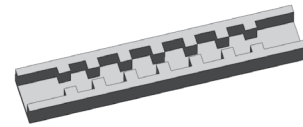


Figure 13 Picture of designed direct coupled resonator BPF (cut model)

3.2 Evaluation of designed filter characteristics

The designed filter was simulated by an EMF simulator (Femtet, Murata Software Co., Ltd.) to evaluate frequency characteristics. Figure 14 shows the reflectance and transmittance characteristics obtained by the simulation. Here, Figure 14(a) shows a model without loss, and (b) shows

results for some lossy models. In Figure 14(a), the 3 dB passband and the pass-band return loss were obtained to 293.8 to 306.2 GHz and <-30 dB, respectively. These results satisfy the design specification in Table 2. In Figure 14(b), to consider the effects of conductor losses, the following four models were compared: (i) with loss-less, (ii) with smooth copper (conductivity 6.0×10^7 S/m), (iii) with smooth aluminum alloy (conductivity 2.0×10^7 S/m) and (iv) with aluminum alloy with rough surface ($R_{\text{rms}} = 0.4$ μm). The maximum losses within the passband ($f_{\text{pass,L}}$ to $f_{\text{pass,H}}$) were 0.0 dB, 1.5 dB, 2.0 dB, and 3.3 dB, respectively. As the conductivity decreases or the surface becomes rougher, the loss increases. This result will be discussed again in Section 4.2. Note that even aluminum alloy with surface roughness $R_{\text{rms}} = 0.4$ μm satisfies the specifications in Table 2.

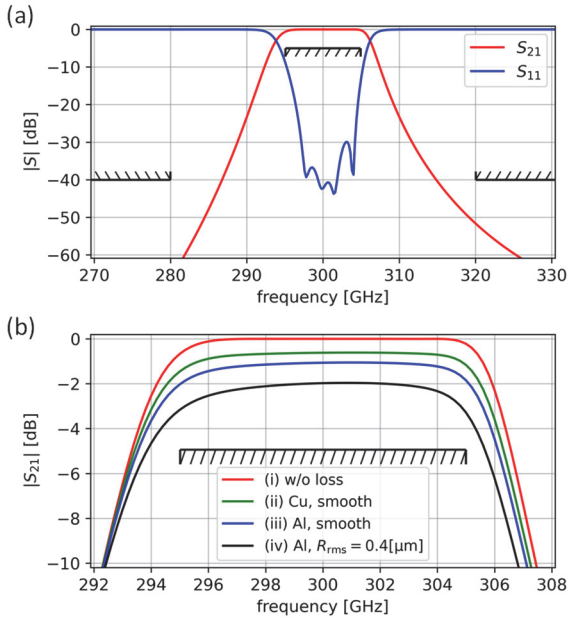


Figure 14 Simulated characteristics of the designed BPF:
(a) S_{11} and S_{21} under lossless condition,
(b) S_{21} with considering several varieties of losses; (i) loss-less, (ii) copper with smooth surface, (iii) Al alloy with smooth surface, (iv) Al alloy with rough surface ($R_{\text{rms}} = 0.4$ μm)

4 Notes for designing millimeter-wave filter

4.1 Filter characteristics variation by manufacturing process

The frequency characteristics of a filter composed of distributed-element circuits depend on each element's shape. Hence, manufacturing accuracy is required to be sufficiently appropriate for the objective frequency range. In this section,

the dependence of millimeter-wave filter characteristics on cavity dimensional error was evaluated using Monte Carlo simulation. Here, an 11th order Butterworth filter with a center frequency of 275 GHz and a passband width of 30 GHz was designed for both inductive and capacitive windows. Normal random numbers with a standard deviation of 2.5 μm were added as errors to the original design dimensions, and the transmittance characteristics $|S_{21}|$ were calculated for several pairs of dimensions. Figure 15 shows the simulation results; the dark line shows the design value model without any errors, and the light line shows the model with errors added. In the results of the inductive window Figure 15(a), (b), and capacitive window (c), deterioration is observed in the cutoff shoulder on the frequency characteristics due to dimensional errors. Figure 15(d) shows that sensitivity to a dimensional variation can decrease when the capacitive window is used, and the thickness is set to $t = \lambda_g/4$. The advantage of using a quarter wavelength thickness capacitive window for the cavity coupler was shown⁹⁾.

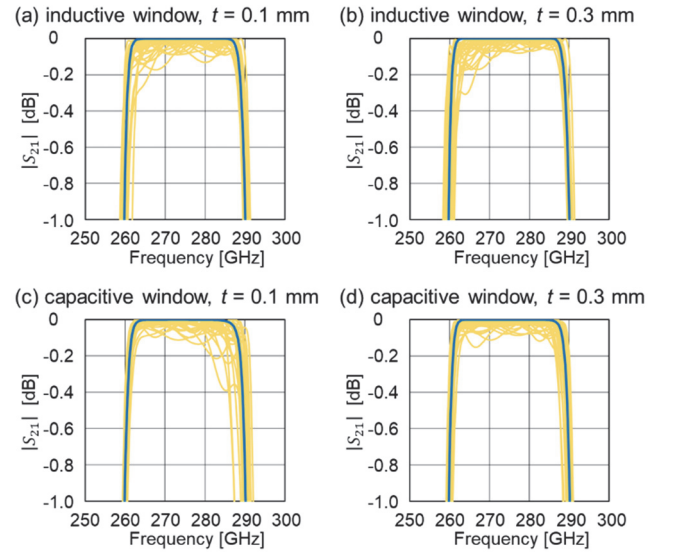


Figure 15 The filter characteristics depend on type of coupling window with dimensional variation, which simulates manufacturing accuracy

4.2 Filter performance limitations due to resonator losses

The loss in a resonator increases the insertion loss in the passband and limits the steep cutoff characteristics. The time-averaged ratio of the energy loss for a single cycle inside the resonator and the stored energy in the resonator is expressed as $2\pi/Q_u$. Here, Q_u is a dimensionless quantity

called the unloaded Q factor of the resonator. The smaller loss of the resonator becomes the larger Q_u and the steeper the resonance characteristics. The Q_u can be expressed by using Q_c due to conductor loss, Q_d by dielectric loss, and Q_r by radiation loss, as follows,

$$\frac{1}{Q_u} = \frac{1}{Q_c} + \frac{1}{Q_d} + \frac{1}{Q_r} \quad (21)$$

Note that magnetic loss is ignored here. For increasing Q_c and Q_d , it is necessary to use a low loss waveguide. A typical microstrip line (MSL) has more considerable radiation losses than hollow waveguides because of the electromagnetic field of an MSL existing in open space. Hence, shields are usually used to suppress radiation loss for increase Q_r . Assuming that the unloaded Q values of each resonator are the same, each loss can be compared by replacing to $s \rightarrow s + \delta$ in the transfer function $H(s)$ of prototype LPFs⁷⁾, where δ is a positive real number representing the loss, and $\delta = 1/(\Delta \cdot Q_u)$ for BPF. In particular, $|H(\delta)|$ represents the transmittance at the center frequency. Since the denominator of the transfer function is a polynomial for the variable s with positive coefficients, the transmittance decreases with smaller Q_u or narrower relative bandwidth. For example, the four models with different losses in Figure 14(b) verify the validity of this calculation. Table 4 shows the results of calculating the insertion loss at the center frequency from the transfer function. The relative bandwidth is $\Delta = 12.4/300$ from the 3 dB bandwidth 12.4 GHz of the lossless model in Figure 14(b). The unloaded Q for this verification was used based on the theoretical value of the TE₁₁₀ mode. As Table 4 shows, the theoretical values estimated from the transfer function agree well with the results of an EMF analysis.

Table 4 Comparison of insertion loss based on EMF analysis and theoretically from Eq. (21)

	EMF analysis	Theoretical value (Estimated)	
	Loss [dB]	Loss [dB]	unloaded Q
Copper (smooth)	0.64	0.52	1560
Aluminum alloy (smooth)	1.04	0.88	925
Aluminum alloy ($R_{rms}=0.4 \mu m$)	1.96	1.68	490

Table 5 shows the characteristics of previously reported filters using various waveguides and the filter designed in Section 3 for comparison. The bottom two rows in Table 5 show the unloaded Q and insertion loss converted to a center

frequency of 300 GHz, a passband width of 10 GHz, and a sixth filter order. Here, a Butterworth filter is assumed, and the unloaded Q is assumed to be inversely proportional to the square root of the center frequency. The loss calculation from the transfer function described above was used to calculate the insertion loss. These differences are due to the loss of each transmission line that constitutes the filter.

Table 5 Comparison of various BPF on several type of waveguides

Waveguide type	MSL	MSL	SIW (Si)	SIW (GaAs)	Hollow	Hollow (WR-3.4)
Center freq. [GHz]	89.6	152.4	270	328.5	286.6	300
Bandwidth [GHz]	11.7	10.5	5	46	16	10
Filter order	3	3	5	4	3	6
Attenuation @ f_0 [dB]	2.3	2.5	9.2	3.6	0.41	1.96*
References	(10)	(10)	(11)	(12)	(13)	—
Estimated Q @ 300 GHz	31.5	71.6	154	47.0	742	490
Estimated attenuation (sixth) [dB]	8.1	6.8	11.6	5.1	0.81	1.68

* Simulated in this paper. Al, $R_{rms}=0.4 \mu m$

To compare the losses in more detail, the cause-specific losses of typical waveguides, an MSL and a hollow waveguide, were calculated. Figure 16 shows losses of an MSL and a hollow in the 150 GHz and 300 GHz bands. The hollow waveguides were WR-6.5 and WR-3.4 as standard waveguides corresponding to each band, and the loss was determined for pure copper, aluminum alloy, and several surface roughness. The MSL was designed so that only the quasi-TEM mode propagated up to 170 GHz or 330 GHz, and its loss was simulated. The relative dielectric constant of the MSL substrate was set to $\epsilon_r = 2.0$ or 3.8, and the dielectric loss tangent was set 0.001 regardless of frequency. Each loss was calculated from the model in reference 6), considering the effect of surface roughness based on the Morgan-Hammastad model¹⁴⁾. Table 6 shows the attenuation constant of hollow waveguides, which is much lower than that of the MSL. Thus, an MSL has advantages such as a smaller footprint, but it should only be used for short distance connections where some loss is acceptable. In contrast, using the hollow waveguide, the attenuation constant of WR-3.4 remains at 20 dB/m with sufficiently low

surface roughness. Thus, it is effective when constructing a high-order filter for high-Q characteristics in the millimeter-wave region, or when suppressing insertion loss as much as possible is desired. To obtain even lower loss characteristics, surface finish processing such as lapping, which achieves a surface roughness of less than $R_{rms} = 0.1 \mu\text{m}$, is required. Therefore, it is desirable to realize a precision processing technology for a millimeter-wave filter that has both of the following specifications: low variation of manufacturing accuracy as shown in Section 4.1, low roughness of the surface as shown in Section 4.2.

Table 6 Comparison of attenuations of an MSL and a hollow waveguide

		Frequency	
		150 GHz	300 GHz
Attenuation [dB/m]	MSL	106	218
	Hollow waveguide (Cu, $R_{rms}=0.1 \mu\text{m}$)	6.2 (WR-6.5)	18.4 (WR-3.4)

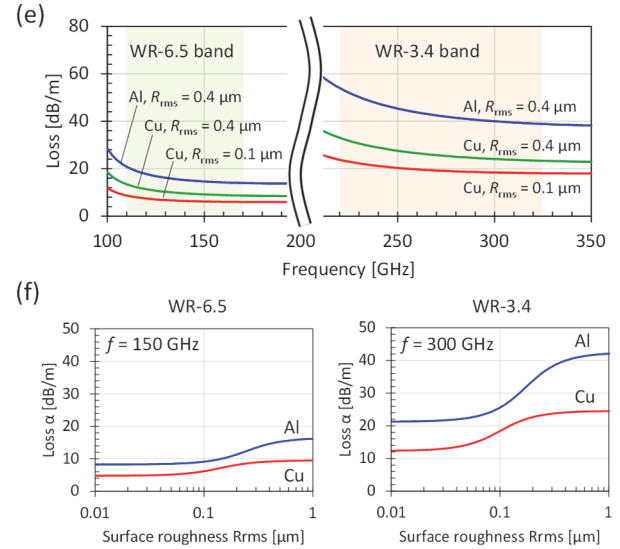
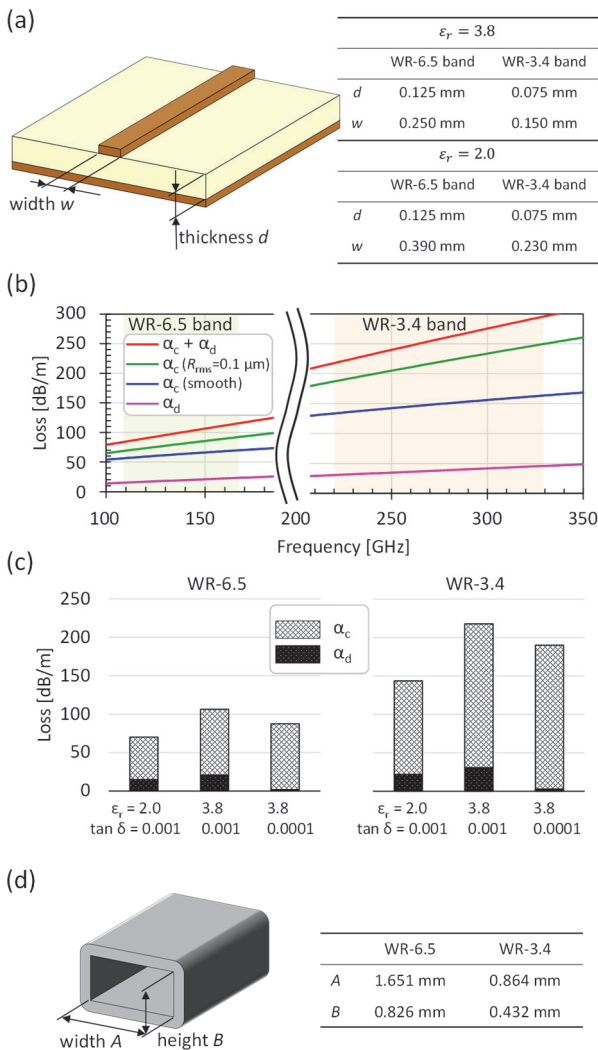


Figure 16 Propagation loss of an MSL and a rectangular hollow waveguide

- (a) MSL dimensions,
(b) calculated loss results, α_c shows conductor loss, and α_d is dielectric loss, respectively,
(c) loss comparison with various permittivity and dielectric loss tangent condition of MSL,
(d) rectangular hollow waveguide dimensions,
(e) calculated loss results due to various materials and surface roughness,
(f) variation of loss due to waveguide surface roughness

5 Conclusion

This paper provided an overview of a filter design method and some notes for design in the millimeter-wave region. In next generation mobile communication, the use of frequency bands that extend to the sub-terahertz region is considered. In such a region, reducing losses of a passive device such as a filter is important. Construction on a low loss waveguide provides an effective way to improve the performance of a filter, and it is necessary to use them appropriately along with a filter on a planar circuit, which has better packaging density and ease of manufacturing.

In filter manufacturing, dimensional accuracy and surface roughness must be managed to achieve the required performance. However, the dimensional accuracy required in a high frequency of millimeter-wave region, such as the 300 GHz band, is generally close to the limit of processing technology. The application of precise manufacturing technology is anticipated to play a crucial role in consistently achieving the high-performance requirements of millimeter-wave band filters. Furthermore, a high-precision device

manufacturing technology for the millimeter-wave band is expected to advance by reducing variation due to manufacturing error and using a correction technique using compensating for the variation in the post process of trimming. These advanced technologies will help accelerate the practical application of the millimeter-wave band.

References

- 1) A. Shahraki, M. Abbasi, M. J. Piran, and A. Taherkordi, "A Comprehensive Survey on 6G Networks: Applications, Core Services, Enabling Technologies, and Future Challenges," arXiv:2101.12475, 2021.
- 2) W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The Road Towards 6G: A Comprehensive Survey," IEEE Open J. Commun. Soc., vol. 2, pp. 334-366, Feb. 2021.
- 3) M. Alizadeh, G. Shaker, J. C. M. Almeida, P. P. Morita, and S. S.- Naeini, "Remote Monitoring of Human Vital Signs Using mm-Wave FMCW Radar," IEEE Access, vol. 7, pp. 54958-54968, Apr. 2019.
- 4) S. Kharkovsky, and R. Zoughi, "Microwave and millimeter wave nondestructive testing and evaluation - Overview and recent advances," IEEE Instrum. Meas. Mag., vol. 61, iss. 5, pp. 2046-2060, May 2013.
- 5) G. L. Matthaei, E. M. T. Jones, and L. Young, "Microwave Filters Impedance-Matching Networks and Coupling Structures," Artech House, Norwood, MA, 1980.
- 6) D. M. Pozer, "Microwave Engineering," 4th Ed., John Wiley & Sons, Inc., New York, NY, 2011.
- 7) R. J. Cameron, C. M. Kudsia, and R. R. Mansour, "Microwave Filters for Communication Systems: Fundamentals, Design, and Applications," John Wiley & Sons, Inc., Hoboken, NJ, 2007.
- 8) Yoshio Kobayashi, Yoshinori Kogami, Yasuo Suzuki, "Microwave dielectric filters," IEICE, Tokyo, CORONA Publishing Co., Ltd, 2007 (in Japanese).
- 9) Shigenori Mattori, "Monte Carlo simulation on dependency on dimensional error of all-pole waveguide filters," 2023 IEICE General Conference, no.C-2-33, Saitama, Mar. 2023 (in Japanese).
- 10) Y. Chen and S.-W. Zhang, "Design of millimeter-wave bandpass filters based on improved asymmetric parallel-coupled microstrip line structure for direct detection system of CubeSat radiometer," Microw. Opt. Technol. Lett. 64(1), pp.23-29, Jan. 2021.
- 11) A. Krivovitca, U. Shah, O. Glubokov, and J. Oberhammer, "Micromachined Silicon-Core Substrate-Integrated Waveguides at 220–330 GHz," IEEE Trans. Microw. Theory Tech. 68(12), pp.5123-5131, Dec. 2020.
- 12) T.-Y. Chiu and C.-H. Li, "Low-Loss Low-Cost Substrate-Integrated Waveguide and Filter in GaAs IPD Technology for Terahertz Applications," IEEE Access 9, pp. 86346-86357, Jun. 2021.
- 13) H. Yang, Y. Dhayalan, X. Shang, M. J. Lancaster, B. Liu, H. Wang, M. Henry, and P. G. Huggard, "WR-3 Waveguide Bandpass Filters Fabricated Using High Precision CNC Machining and SU-8 Photoresist Technology," IEEE Trans. Terahertz Sci. Technol. 8(1), pp.100-107, Jan. 2018.
- 14) B. C. Wadell, "Transmission Line Design Handbook," Artech House, Boston, MA, 1991.

Authors



Takahiko Mizuno
2nd Laboratory
Advanced Research Laboratory

Publicly available

Development of Pulse Optical Spectrum Measurement Technology using Time Averaging

Soichiro Kai, Maki Ueno, Shinji Morimoto, Ryosuke Nakagawara, Shinichi Seto

[Summary]

Operators face continuing demand for higher network speeds due to the spread of teleworking, online entertainment etc. The high optical output of high-bit-rate LD chips supporting faster networks causes heat-generation issues at manufacturing inspection tests. There is increasing demand for spectrum measurement of optical pulse to solve this issue, but conventional methods do not permit both stable results and short test times. We have developed a stable and fast optical pulse spectrum measurement technology using a time averaging method with the Optical Spectrum Analyzer MS9740B.

1 Introduction

Recently, higher-output and higher-reliability Distributed FeedBack Laser Diodes (DFB-LD) are being investigated for applications in data-center high-bit-rate optical transceivers, optical-sensing light sources, Light Detection And Ranging (LiDAR), etc.

These high-output DFB-LD are evaluated using optical spectrum analyzers (OSA) but the increasing temperature of the LD chip while emitting light causes changes in the spectrum characteristics, such as wavelength drift and lower optical power level (Figure 1).

As a result, pulsed light measurement is used as a method to suppress rising LD chip temperature when evaluating peak wavelength, peak power, Side Mode Suppression Ratio (SMSR), etc.

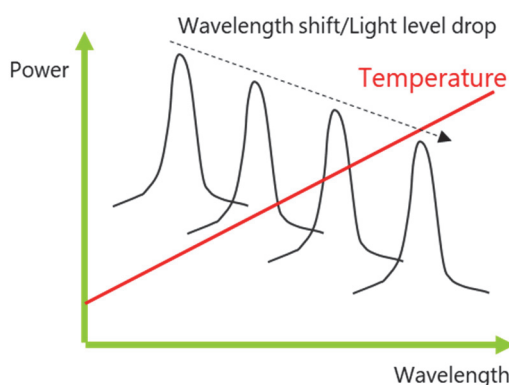


Figure 1 Changes in Spectrum Characteristics with LD Chip Temperature

Unlike Continuous-wave (CW) emission, increases in LD chip temperature can be controlled by emitting light periodically by turning the driving current on and off alternately (Figure 2).

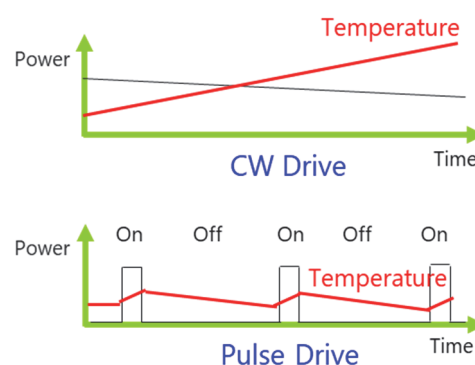


Figure 2 Change in LD Chip Temperature at CW and Pulse Driving

As future LD outputs become higher, this pulsed light emission will probably be essential when testing and evaluating optical spectrum characteristics. In particular, more accurate and faster SMSR testing at LD chip manufacturing inspection will be increasingly important in productivity improvement.

2 Optical Pulse Test Methods

There are three optical pulse test methods: external trigger, peak hold, and time averaging. Table 1 lists the advantages and disadvantages of each.

• External Trigger Test Method

Measures optical spectrum by synchronizing with emitted light timing

• Peak Hold Test Method

Measures optical spectrum by capturing maximum power within a constant period interval without requiring trigger signal

• Time Averaging Spectrum Test Method

Measures optical spectrum by averaging data captured over a constant period interval without requiring trigger signal

Table 1 Comparison of Optical Pulse Test Methods

Measurement Method	Advantage	Disadvantage
External Trigger	High-reproducibility measurement by synchronizing to optical pulse Supports peak power measurement	Long measurement time because requires repeated control and driving of diffraction grating in OSA
Peak Hold	Does not require trigger signal Supports peak power measurement	Increases noise level by holding of peak signal
Time Averaging Spectrum	Does not require trigger signal Fast stable measurement of optical spectrum by controlling measurement to optimize optical pulse	Cannot measure peak power but measures average power because uses additive averaging processing

3 Development of Spectrum Measurement Technology using Time Averaging Method

3.1 Need for OSA at LD Chip Manufacturing Inspection

LD manufacturing inspection requires productivity improvement by accurate and stable SMSR measurement. To meet these requirements, a pulse test method was selected because it measures SMSR quickly and stably.

3.2 Design Requirements

The important design points are listed below.

- 2.8 dB p-p SMSR reproducibility for pulses with 5 kHz-frequency and 1%-duty ratio
- Asynchronous measurement (trigger-less)
- Similar measurement time to CW light measurement
- Optimized for measurement system using MS9740B

3.3 OSA Functional Block Diagram

Figure 3 shows the functional block diagram for this system.

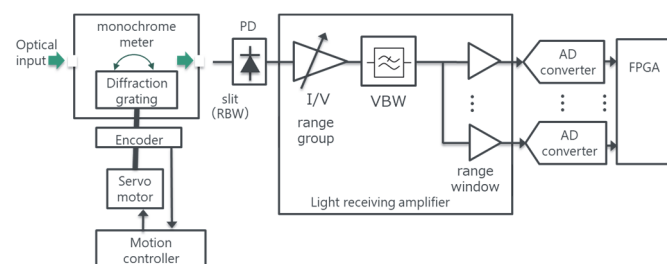


Figure 3 Optical Spectrum Analyzer Function Block Diagram

The light to be measured is diffracted by a diffraction grating and passes via a slit to a photodiode (PD) where it is O/E converted and the signal is amplified by the light receiving amplifier and is A/D converted. The measured wavelength is selected by rotating the diffraction grating.

To shorten measurement times using this developed system, the spectrum of the optical pulse is measured while wavelength sweeping (continuously variable wavelength).

Consequently, the timing of the wavelength sweeping and the asynchronous optical pulse measurement timing must be controlled.

Moreover, to shorten the measurement time also when using the light receiving amplifier, the simultaneous A/D conversion configuration with a wide level range (range window) is applied, which adjusts the signal gain (range group) depending on the power of the measured light.

3.4 Software Processing

Figure 4 shows the relationship between the pulse signal and averaging range. In the averaging range, values of the averaging points, which are specified by the software, are obtained and averaged for each period of reading the AD values (this is the averaging window).

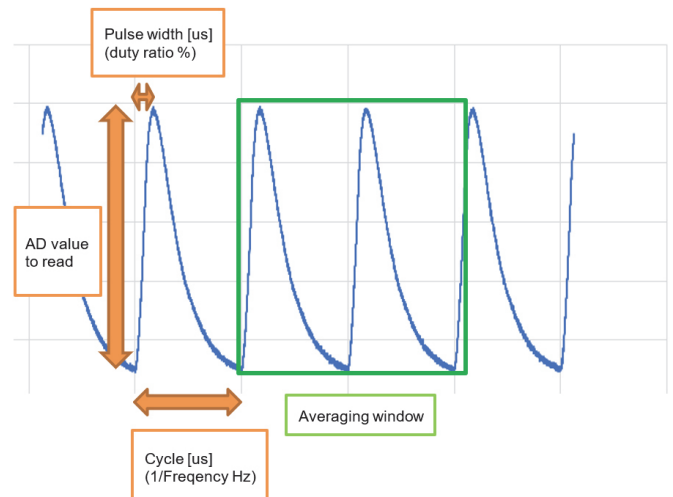


Figure 4 Relationship between Pulse Signal and Averaging Window

The number of averaging points is determined as the pulse period multiplied by an integer divided by the AD reading period. For example, the sampling interval is 4 μ s considering the sampling processing time per point when the measured light pulse cycle is 200 μ s (5 kHz pulse cycle), the pulse is 2 μ s (1% duty), and the VBW is 1 kHz.

The sampling time is set to be equivalent to the light pulse period multiplied by an integer, as shown in Figure 5, to shorten measurement time. In this time, we set the sampling time so that the number of averaging points (=100) was equivalent to two pulse period.

$$\text{Number of averaging points} = \frac{\text{Pulse cycle integer multiple time}}{\text{AD read cycle}}$$

Figure 5 Averaging Point Count Calculation

To minimize the times to switch the light receiving amp range to quickly measure a wide level range, the light receiving amp is configured for measuring a level range of 50 dB simultaneously and a range group that switches the range window in six steps with the difference of 10 dB, depending on the input level. Figure 6 shows the relationship between the range window and range group. This enables high-speed measurement from +10 to -90 dBm.

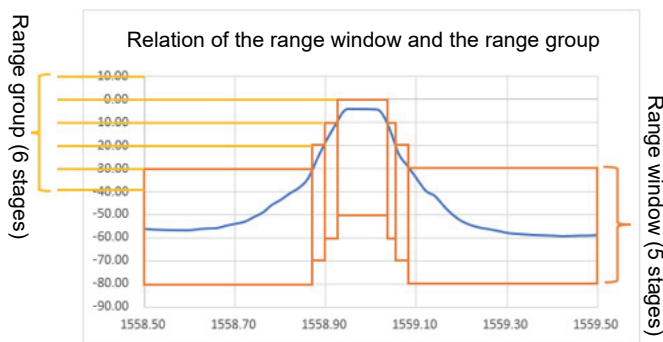


Figure 6 Relationship between Range Window and Range Group

To capture measurement data in the optimum range window, the average and maximum levels can be captured from the multiple sampling data at each measurement point, and the optimum level is determined based on the maximum level.

Obtaining process of the average (mean) and maximum level values are accelerated by a Field Programmable Gate Array (FPGA).

Additionally, we added a function for switching the Pulse Mode On/Off with the GUI. Figure 7 shows the setting screen when this option is enabled and Figure 8 shows the waveform display screen.

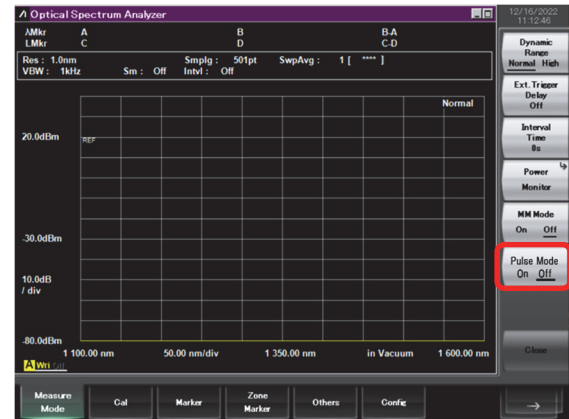


Figure 7 Added Pulse Mode Function

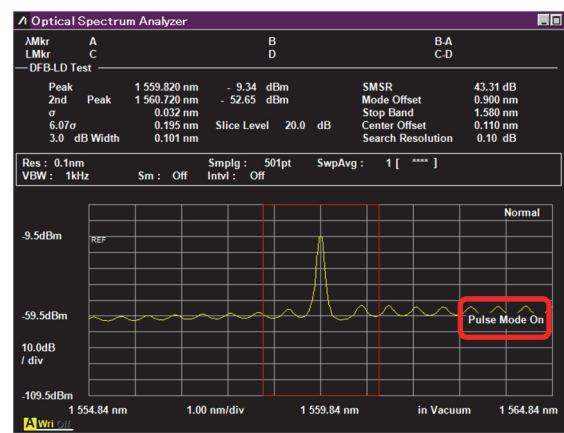


Figure 8 Waveform Display Screen

Based on the above functions, users can use the Pulse Mode just by updating the firmware version and by purchasing the option, without adding any hardware.

3.5 Time Averaging Processing using FPGA

Time averaging processing is handled by the FPGA. As shown in Figure 9, time-averaged sampling values are derived by averaging multiple AD data points obtained at a shorter cycle compared to the sampling points forming the total waveform (the measured optical signal level digital data value from the A/D converter is called AD data).

AD data points are used to average multiple points around a target sampling point on the time axis. For example, the temporal interval between AD data points is 4 μs and 100 points are averaged, as described in section 3.4. Under the measured pulse light conditions given in section 3.4, the two cycles of pulsed light shown in Figure 4 will be contained in the averaging window shown in Figure 9.

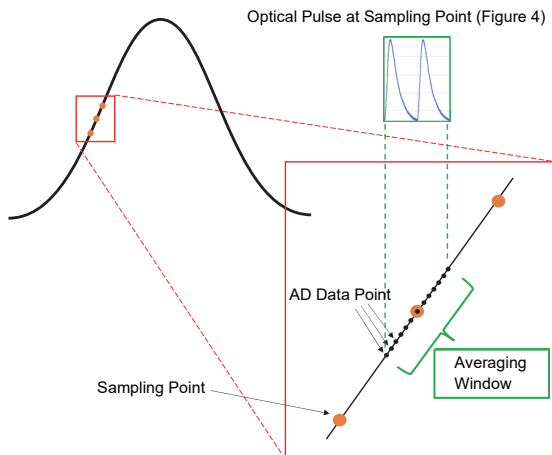


Figure 9 Time Averaging Schematic

The following section explains the FPGA operation for implementing the above-described operation.

As shown in Figure 10, the FPGA is composed of internal RAM for saving AD data points as well as RAM Controller and Calculation Process modules. The firmware handles the various settings for averaging processing.

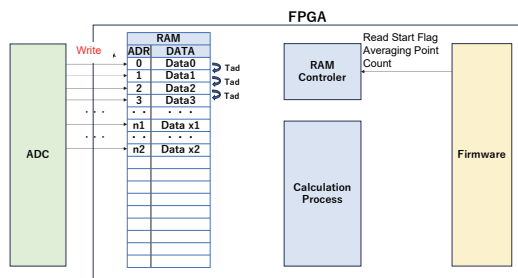


Figure 10 FPGA Internal Operation 1

Regardless of the firmware operation, the FPGA always reads values from the A/D converter at a fixed cycle T_{ad} and saves the result to RAM. The RAM uses a ring buffer structure and replaces the oldest data with latest data when it is full. Instead of using firmware timing, the FPGA independently writes AD data from the A/D converter to RAM, which enables the use of old saved AD data for averaging when a request for a data of a sampling point is received from firmware.

At the waveform sampling point acquisition timing, the RAM Controller receives information about the start flag and the number of averaging points from firmware. The RAM Controller then checks the address storing the latest AD data point at that point ($n2$ in Figure 10), and watches and waits until subsequent AD data points equal to half of the number of averaging points are added from the ADC.

Figure 11 shows the operation after the data write wait. Reading RAM starts when the write wait until address $n3$

ends. The data at the address $n1$ ($n2 - \text{averaging point count} \div 2$) to $n3$ ($n2 + \text{averaging point count} \div 2$) are read sequentially and transferred to the Calculation Process module. The Calculation Process module sums up the transferred data and the value is divided by the averaging point number, resulting in the average value. Additionally, the maximum value is also extracted from the data read from RAM. Both the calculated average and maximum values are passed to firmware and the maximum value is used for ranging processing by the firmware.

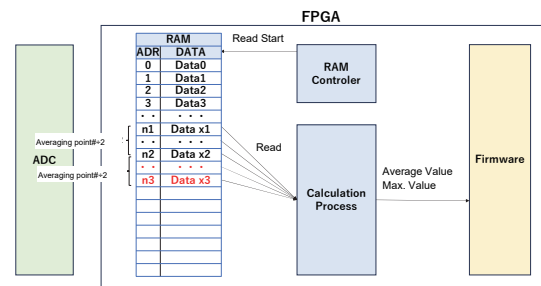


Figure 11 FPGA Internal Operation 2

Based on the above operation, averaging of multiple points in the temporal range around the sampling point is performed.

We consider the reason why the AD data points before and after the sampling points are averaged, from the perspective of processing time and the error between the theoretical value of the sampling points.

Table 2 compares the three methods: the AD data points for averaging are all before the sampling point, are all after that, and both before and after that.

Table 2 Advantages and Disadvantages Depending on AD Data Point Position Relative to Sampling Point

	AD Data Point for Averaging	Processing Time	Deviation from Theoretical Value
A	Before sampling point	○	△
B	After sampling point	×	△
C	Both before and after sampling point	△	○

From the viewpoint of FPGA operation, when the request of averaging comes from the firmware on reaching the sampling point, addition of data after the sampling point requires waiting time to write from the A/D converter to RAM. Consequently, using AD data only before the sampling point is fastest because no wait time is required. Conversely, using AD data only after the sampling point requires more time

because it is necessary to wait just for the averaging point number. When using AD data both before and after the sampling point, the processing time is midway between the other two methods because the averaging point number to be waited is half. Figure 12 compares the processing for each method when the sample number is n.



Figure 12 Processing Time Comparison

From the viewpoint of deviation from the theoretical value, although it depends on the averaging point number, the error rate is lowest compared to the theoretical value of the target sampling point's average power when the AD data is centered around (before and after) the original sampling point. When the AD data is either only before or after the sampling point, the average power is lower or higher than the theoretical value, compared to the first case. Therefore, the error is thought to be larger. Figure 13 depicts the comparison of the error from the theoretical value for each method.

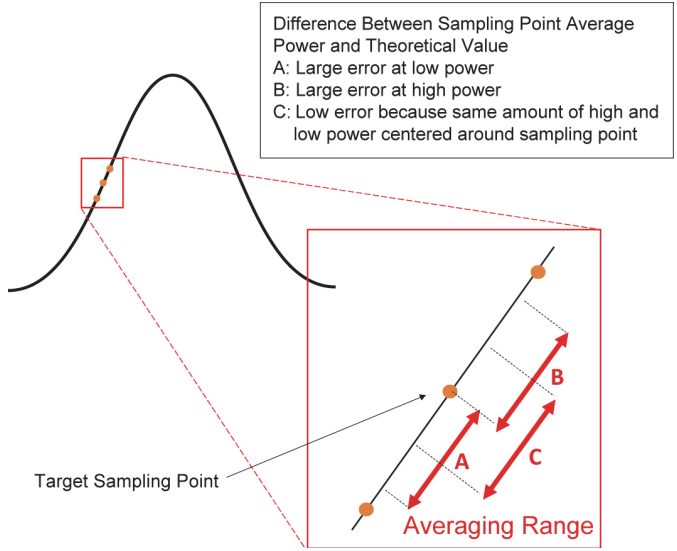


Figure 13 Sampling Point Average Power Theoretical Error Comparison

We adopted the type-C sampling method (AD data both before and after sampling point) in Table 2 because it satisfies both reasonable processing speed and low error rate.

4 Results

Table 3 and Figure 14 show the results with the time averaging method, as well as the conventional external trigger and normal sweeping measurement methods. SMSR stability was higher for the time averaging method than the normal sweeping method, and the measurement time was faster than the external trigger method. We confirmed the improved stability and measurement time with the time averaging method.

Table 3 Comparison of Time Averaging and Conventional Test Results

Optical Pulse Test Method	SMSR Value [dB]	SMSR Stability [dBp-p]	Measurement Time [s]
Time Averaging	44.00	0.31	0.482
External Trigger	43.94	0.10	1.172
Normal Sweep	31.92	46.82	0.450

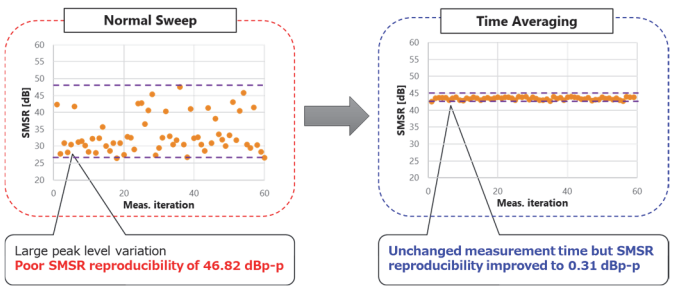


Figure 14 Comparison of Time Averaging and Normal Sweep SMSR Stability

5 Conclusions

Due to increasing demand from manufacturers for optical pulse measurements supporting high-output optical transceivers for faster and larger-capacity communications, we developed an optical pulse measurement technology using the time averaging method to support fast, stable optical spectrum measurements.

In the future, we aim to help optical network development by further improving optical pulse measurement methods.

Authors



Soichiro Kai
Product Development Dept.
Service Infrastructure Solutions
Test & Measurement Company



Maki Ueno
Solution Marketing Dept.
Service Infrastructure Solutions
Test & Measurement Company



Shinji Morimoto
Product Development Dept.
Service Infrastructure Solutions
Test & Measurement Company



Ryosuke Nakagawara
Product Development Dept.
Service Infrastructure Solutions
Test & Measurement Company



Shinichi Seto
Software Engineering Dept.
Engineering Div.

Comparative Study of Antenna Array Calibration Methods Based on Symmetrical Mutual Coupling of 4-Element Linear Array

Takashi Mori

[Summary]

Antenna array calibration is indispensable for accurate direction-of-arrival (DOA) estimation, because mutual coupling and amplitude/phase error on antenna array cause a significant error in the MUSIC spectrum. Although an asymmetric calibration method which independently calculates all elements of the calibration matrix enables precise array calibration, a large number of reference data set is required to find many unknown variables. On the other hand, the 28-GHz frequency band is used for high-speed wireless communication on the 5th generation mobile communication system, and a demand for DOA estimation will increase on the SHF band. In this paper, two calibration models using symmetrical mutual coupling of 4-element uniform linear array were compared to reduce the number of reference data set. As a result of 28-GHz band experiment, the DOA estimation error and MUSIC spectrum were improved by an appropriate calibration model and were close to that of the asymmetric calibration method with many unknown variables. Copyright © 2023 IEICE

1 Introduction

By using an array antenna, it is possible to estimate the direction of arrival of radio waves without mechanical scanning, and high-resolution direction-of-arrival estimation methods such as the MUSIC algorithm¹⁾ are known well. In the typical MUSIC algorithm, it is assumed that each antenna can ideally receive the incoming waves, but in actual array antennas, there are various error factors listed below, so it is difficult to obtain the theoretical results. This hinders accurate direction-of-arrival estimation. The first error factor is mutual coupling between antenna elements. When one antenna receives an incoming wave, re-radiation occurs due to the current excited on the element, and another antenna receives this, resulting in mutual coupling between elements. The second error factor is amplitude and phase errors of each element. Even if antenna elements, feed lines, and receivers with the same specifications are arranged in an array, there will be non-uniformity in the actual device, resulting in errors in amplitude and phase characteristics between the elements. The third error factor is considered the position error of the antenna element. In this paper, we assume that the position errors can be ignored and discuss an array calibration method that calibrates the mutual coupling between elements and the amplitude and phase errors of each element. Furthermore, in the 5th generation mobile communication system (5G), the 28-GHz band is used in addition to the sub 6-GHz band to perform high-speed, large-capacity wireless

communication using a wide frequency band, and direction-of-arrival estimation in the SHF band is expected to be necessary. As the carrier frequency increases, the loss of high-frequency devices and transmission lines generally increases, and the carrier phase difference due to line length differences also increases, so the amplitude and phase errors of each element tend to increase, and array calibration becomes more important.

The method of determining a calibration matrix using a wave source with known direction of arrival is considered useful in practice, since it is possible to calibrate the receiver's amplitude and phase errors as well. When the number of antenna elements is N , mutual coupling between elements is represented by a complex matrix of $N \times N$, and amplitude and phase errors are represented by N complex numbers. If we do not assume array symmetry and assume that all elements of the calibration matrix are independent, the mutual coupling between elements and the amplitude and phase errors can be expressed together by an $N \times N$ calibration matrix. This asymmetrical calibration method using the $N \times N$ matrix is versatile and can be applied to various antenna configurations, and enables highly accurate calibration, but it requires many calibration data sets because of the large number of unknown variables. On the other hand, with an antenna configuration such as a uniformly spaced linear array, it is possible to reduce unknown variables in the mutual coupling matrix due to its symmetry. However, since the

amplitude and phase errors of each element cannot be included in the mutual coupling matrix, the unknown variables of the mutual coupling coefficient and the amplitude and phase errors take the form of a product, making it difficult to directly solve each unknown variable. Previously proposed methods^{2), 3)} solve the product of the mutual coupling coefficient and the amplitude/phase error as a new unknown variable. There were two problems in that the number of unknown variables increased and the calibration results varied depending on the method of calculating the mutual coupling coefficients and the amplitude and phase errors from the new unknown variables. To overcome these problems, iterative calculation methods^{4), 5)} have been proposed for determining mutual coupling coefficients and amplitude/phase errors. The latter method allows calculations without increasing the number of unknown variables, so it has the feature of being able to perform array calibration with a small number of calibration data sets.

In this paper, we focus on the symmetrical mutual coupling coefficient and its iterative calculation method for array calibration with a small calibration data set and conduct an array calibration experiment using a 28-GHz band four-element uniformly spaced linear array antenna to compare several calibration methods. Specifically, array calibrations of a four-element uniformly spaced linear array are performed using the calibration model described in References 4) and 5) and a calibration model in which the order of the mutual coupling coefficient and amplitude/phase error is reversed. As a calibration result, we evaluated the error in direction of arrival estimation using the MUSIC algorithm and the peak value of the MUSIC spectrum and compared the two calibration models. We also compared the results of the calibration using only the amplitude and phase errors of each element and the asymmetrical calibration using the $N \times N$ matrix.¹

2 Array Calibration Models

2.1 Symmetrical Calibration Model #1

The number of elements of the receiving antenna is N , the positions of the receiving antenna elements are $x_1, \dots, x_N$, the number of calibration data sets is M , the complex

amplitude of the i th calibration signal is s_i , the angle of arrival is θ_i , and similar to the References 4) and 5), the received signal vector $\mathbf{r}_i$ is expressed using the calibration model of equation (1).

$$\mathbf{r}_i = \mathbf{C}\mathbf{\Gamma}\mathbf{a}(\theta_i)s_i + \mathbf{n}_i \quad (1)$$

$$\begin{aligned} \mathbf{a}(\theta_i) &= [a_{i,1}, \dots, a_{i,N}]^T \\ &= \left[e^{j\frac{2\pi}{\lambda}x_1 \sin \theta_i}, \dots, e^{j\frac{2\pi}{\lambda}x_N \sin \theta_i} \right]^T \end{aligned} \quad (2)$$

$$\mathbf{\Gamma} = \text{diag} \{ \gamma_1, \dots, \gamma_N \} = \text{diag} \{ \alpha_1 e^{j\phi_1}, \dots, \alpha_N e^{j\phi_N} \} \quad (3)$$

$$\begin{aligned} \mathbf{n}_i &= [n_1^{(i)}, \dots, n_N^{(i)}]^T \\ (i &= 1, \dots, M) \end{aligned} \quad (4)$$

Where, $\mathbf{a}(\theta)$ is the ideal mode vector of the linear array, and $\mathbf{\Gamma}$ is the diagonal matrix representing the amplitude error $\alpha_1, \dots, \alpha_N$ and phase error $\phi_1, \dots, \phi_N$ of each element, $\mathbf{C}$ is an $N \times N$ complex matrix representing mutual coupling between elements, $\mathbf{n}_i$ is the noise vector at the i th reception. The mutual coupling coefficient $\mathbf{C}^{(3)}$ in equation (5) is assumed based on the symmetry of the four-element uniformly spaced linear array antenna.

$$\mathbf{C} = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_1 & c_4 & c_5 & c_2 \\ c_2 & c_5 & c_4 & c_1 \\ c_3 & c_2 & c_1 & c_0 \end{bmatrix} \quad (5)$$

The unknown variables are $\{c_0, \dots, c_5, \gamma_1, \dots, \gamma_4\}$, and if the coefficients are normalized to $c_0 = 1, \gamma_1 = 1$, there are 8 unknown variables to solve. To calculate these unknown variables, it is necessary to obtain the received signal $\mathbf{r}_i$ when a wave source with a known arrival angle θ_i is received. Equation (1) includes the product of the unknown variables $\mathbf{C}$ and $\mathbf{\Gamma}$ and is difficult to solve directly, so a method that calculates the product of the unknown variables has been reported³⁾. However, the number of unknown variable products to be solved is 15, which is the same number as a 4×4 matrix (with any one element as the reference) that does not assume symmetry, and the effect of reducing unknown variables due to symmetry cannot be obtained. Furthermore, when the symmetry is not perfect, such as in actual measurement data, there is a problem in that the calibration results vary depending on the method used to convert the 15 unknown variable products into the 8 unknown variables. Therefore, in

¹ The material in this paper was presented in part at the IEICE Technical Report 9), and all the figures of this paper are reused from 9) under the permission of the IEICE.

this paper, we calculate the 8 unknown variables using the iterative calculations shown in references 4) and 5) to obtain the mutual coupling coefficient $\mathbf{C}$ and the amplitude/phase error $\mathbf{\Gamma}$.

2.2 Symmetrical Calibration Model #2

Considering that the mutual coupling $\mathbf{C}$ between receiving elements occurs at the receiving antenna, and the amplitude and phase errors $\mathbf{\Gamma}$ of each receiving element are caused by the feed lines and receivers after the receiving antenna, the order of $\mathbf{C}$ and $\mathbf{\Gamma}$ changes, and the calibration model is expressed as

$$\mathbf{r}_i = \mathbf{\Gamma} \mathbf{C} \mathbf{a}(\theta_i) s_i + \mathbf{n}_i \quad (6)$$

Reference 6) presents a calibration model for the order of $\mathbf{\Gamma} \mathbf{C}$, but does not assume the symmetry of mutual coupling, whereas this paper assumes the symmetrical mutual coupling of equation (5) for a four-element uniformly spaced linear array antenna, as described in the previous section. The unknown variables are $\{\alpha_0, \dots, \alpha_5, \gamma_1, \dots, \gamma_4\}$, and if the coefficients are normalized to $\alpha_0 = 1, \gamma_1 = 1$, there are 8 unknown variables to solve. To calculate these unknown variables, it is necessary to obtain the received signal $\mathbf{r}_i$ when a wave source with a known arrival angle θ_i is received. As described in the previous section, equation (6) includes the product of the unknown variable $\mathbf{C}$ and $\mathbf{\Gamma}$, and it is difficult to solve it directly. Therefore, we calculated the 8 unknown variables using the iterative calculations shown in references 4) and 5), and obtained the mutual coupling coefficient $\mathbf{C}$ and the amplitude/phase error $\mathbf{\Gamma}$.

3 Experiment

Figure 1 shows the experimental setup.

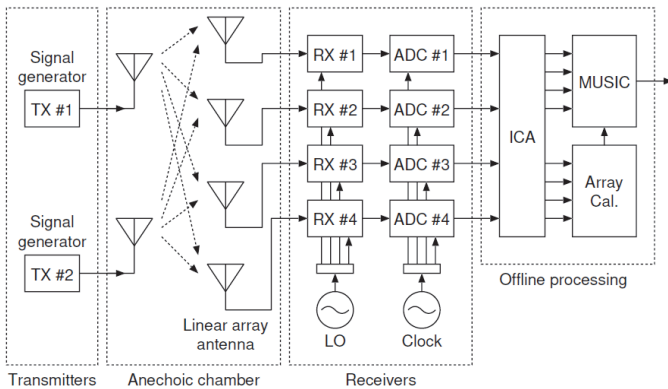


Figure 1 Experimental setup for array calibration

The transmitters (TX#1, #2) are general-purpose signal

generators to output OFDM signals with a carrier frequency of 28.25 GHz, a bandwidth of 100 MHz, a subcarrier spacing of 120 kHz, and QPSK modulation. The transmitting antenna is a horn antenna, and the receiving antenna is a four-element linear array with half-wavelength spacing using the sleeve antennas⁷⁾ formed on printed circuit boards. The distance between the transmitting and receiving antennas was 2.15 m, and the experiment was conducted in an anechoic chamber. The receiver consists of a 4-channel down converter (RX#1–#4) and a 4-channel A/D converter (ADC#1–#4) using a common local signal and a common A/D conversion clock signal for phase synchronization between channels. Although the signal wiring for each channel was designed to be as equal in length as possible, no individual phase correction was performed using variable phase shifters, etc., and phase differences due to cable length variations and individual differences in high-frequency devices remain. The A/D converters with simultaneous sampling capability of 4 channels digitize each IF signal from the down converter with a sampling frequency of 983.04 MHz, a resolution of 14 bits, and a time span of 4 ms, and then convert to I/Q signal with a sampling frequency of 245.76 MHz by the digital down converter (DDC). In the experiment, we received two signals with different directions of arrival, separated the signals using independent component analysis (ICA), calculated the covariance matrix, and performed array calibration. 25 data sets were obtained by setting the direction of arrival angle from -72° to $+72^\circ$ in 6° steps, and among them, 7 data sets of $\{-72^\circ, -48^\circ, -24^\circ, 0^\circ, +24^\circ, +48^\circ, +72^\circ\}$ were used for array calibration. The MUSIC spectrum after calibration was calculated by

$$P_{\text{MUSIC}}(\theta) = \frac{\tilde{\mathbf{a}}^H(\theta) \tilde{\mathbf{a}}(\theta)}{\tilde{\mathbf{a}}^H(\theta) \mathbf{E}_N \mathbf{E}_N^H \tilde{\mathbf{a}}(\theta)} \quad (7)$$

using the calibrated mode vector $\tilde{\mathbf{a}}(\theta) = \mathbf{C} \mathbf{a}(\theta)$ or $\tilde{\mathbf{a}}(\theta) = \mathbf{\Gamma} \mathbf{C} \mathbf{a}(\theta)$. Where, $\mathbf{E}_N$ is a matrix consisting of noise eigenvectors. To estimate the direction of arrival, we calculated the angle θ_i at which the MUSIC spectrum is maximum for each of the 25 data sets from -72° to $+72^\circ$, and calculated the peak value of the MUSIC spectrum $P_{\text{MUSIC}}(\theta_i)$ to evaluate the calibration accuracy.

4 Experimental results

Figure 2 shows the direction of arrival estimation error, and Figure 3 shows the peak value of the MUSIC spectrum.

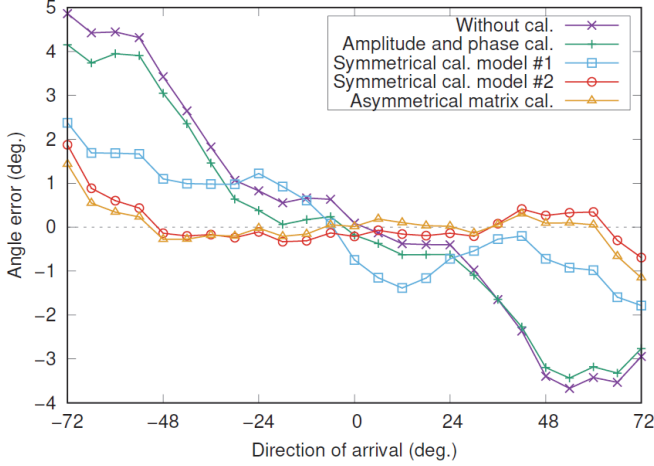


Figure 2 Direction of arrival estimation error

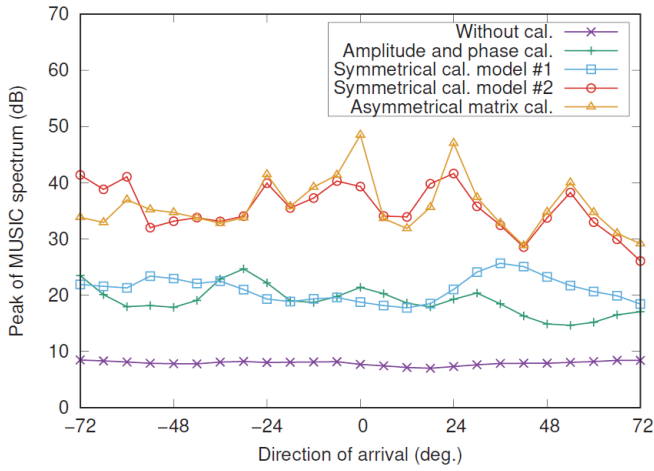


Figure 3 Peak value of MUSIC spectrum

“x” indicates no array calibration, “+” indicates calibration of only the amplitude and phase errors of each element, “□” indicates the symmetrical calibration model #1 ($\mathbf{CT}$), “○” indicates the symmetrical calibration model #2 ($\mathbf{\Gamma C}$), “△” indicates the asymmetrical calibration using 4×4 matrix. Although the peak value of the MUSIC spectrum is slightly improved by calibrating only the amplitude and phase errors of each element, the direction of arrival estimation error is not much different from the case without array calibration, and so the direction of arrival estimation error seems to be mainly due to mutual coupling between elements. The symmetrical calibration model #1 ($\mathbf{CT}$) shows an error in the direction of arrival over the entire angle, and the peak value of the MUSIC spectrum also hardly improves from the calibration of only the amplitude and phase errors. Therefore, the calibration model in which the mutual coupling exists after the amplitude and phase errors is considered to be inconsistent with the experiment. In the symmetrical calibration

model #2 ($\mathbf{\Gamma C}$), the direction of arrival estimation error decreases and the peak of the MUSIC spectrum increases, yielding results close to the asymmetrical calibration using a 4×4 matrix.

Figure 4 summarizes the relationship between the number of unknown variables and calibration accuracy. Figure 4(a) shows the Root-Mean-Square (RMS) value of the direction of arrival estimation error, and Figure 4(b) shows the average value of the MUSIC spectrum calculated using equation (8). Since the denominator of $P_{\text{MUSIC}}(\theta)$ is expressed as the sum of squares of the inner product of the mode vector and the noise eigenvector⁸⁾, the average value of the reciprocal number was used as the average value of the MUSIC spectrum.

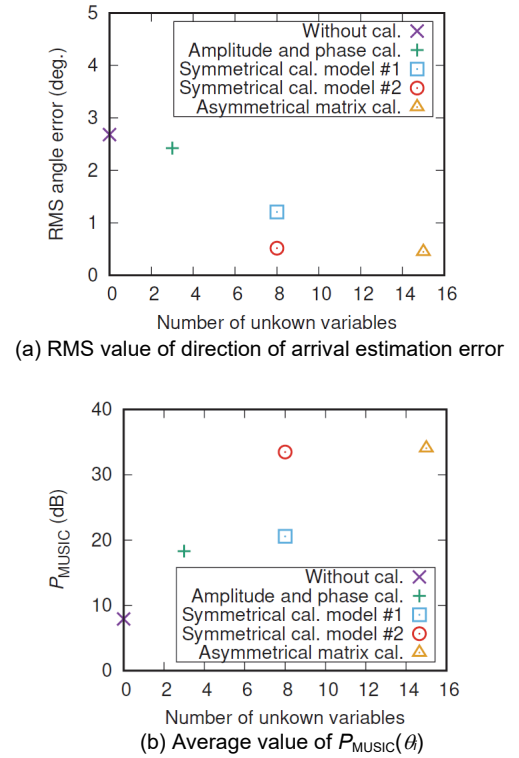


Figure 4 Comparison of array calibration results

$$P_{\text{MUSIC}} = \frac{1}{\frac{1}{L} \sum_{i=1}^L \frac{1}{P_{\text{MUSIC}}(\theta_i)}} \quad (8)$$

Where, L is the number of measurement points for the direction of arrival estimation, and in this experiment, $L = 25$. Figure 4 shows that the symmetrical calibration model #2 can achieve calibration accuracy close to the asymmetrical calibration by a relatively small number of unknown variables. It is considered that the symmetry of the mutual coupling coefficient shown in equation (5) and the calibration model shown in equation (6) are in good agreement with the experiment.

Table 1 shows the calibration values obtained using the symmetrical calibration model #2 (ΓC).

Table 1 Calibration value by calibration model #2

		Amplitude	Phase (deg.)
Amplitude and phase error	γ_1	1.00	0.0
	γ_2	0.81	-19.9
	γ_3	0.83	40.5
	γ_4	0.97	-15.9
Mutual coupling coefficient	c_0	1.00	0.0
	c_1	0.18	48.3
	c_2	0.09	-146.5
	c_3	0.06	28.6
	c_4	1.08	0.0
	c_5	0.19	46.0

The experimental results in Reference 3) showed a phase error of 21° (peak-to-peak) at a frequency of 2.4 GHz, whereas our experiment had a large phase error of 60° (peak-to-peak) due to the high frequency of 28 GHz, so it is thought that our experiment showed the large difference depending on the calibration models. Regarding the mutual coupling coefficient between elements, the diagonal components c_0 and c_4 have an amplitude of approximately 1, and the amplitude decreases in the order of the mutual couplings c_1 and c_5 between adjacent elements, the mutual coupling c_2 two elements apart, and the mutual coupling c_3 three elements apart. This trend is similar to the amplitude of the mutual coupling coefficient in reference 3) and is considered to be a valid result as a four-element linear array.

5 Conclusion

For the purpose of array calibration with a small calibration data set, we used an iterative calculation method assuming mutual coupling coefficients based on the symmetry of a four-element uniformly spaced linear array, and made comparison between the calibration model #1 which exists mutual coupling after amplitude and phase errors and the calibration model #2 which exists amplitude and phase errors after mutual coupling. As a result of experiments in the 28-GHz band, the calibration model #2 showed better characteristics in terms of direction of arrival estimation error and MUSIC spectrum. Therefore, the calibration model based on that mutual coupling occurs at the receiving antenna and

amplitude and phase errors occur at the feed lines and the receivers after the receiving antenna is considered more realistic. In addition, the calibration results using the calibration model #2 are close to the results of the asymmetrical calibration in which all elements of the 4×4 matrix are calculated independently, and so the assumption of mutual coupling coefficients based on symmetry is considered valid. From the above, it was found that highly accurate array calibration is possible with a relatively small number of calibration data sets using the symmetrical mutual coupling coefficients and the calibration model #2.

6 Acknowledgments

This paper includes the results of the research and development under “R&D for further advancement of the 5th generation mobile communication system” (JPJ000254) conducted by the Ministry of Internal Affairs and Communications.

References

- 1) R.O. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Trans. Antennas Propag., vol.AP-34, no.3, pp.276–280, March 1986.
- 2) T. Arai, R. Hara, H. Yamada, and Y. Yamaguchi, “On calibration of superresolution array using known sources,” IEICE Trans. Commun. (Japanese Edition), vol.J86-B, no.3, pp.527–535, March 2003.
- 3) K. Chiba, H. Yamada, and Y. Yamaguchi, “Experimental verification of antenna array calibration using known sources,” IEICE Technical Report, AP2002-41, pp.7–12, July 2002.
- 4) R. Hara, H. Yamada, and Y. Yamaguchi, “Iterative array calibration method with reduction of number of calibration data sets,” IEICE Trans. Commun. (Japanese Edition), vol.J86-B, no.9, pp.1906–1913, Sept. 2003.
- 5) R. Hara, H. Yamada, and Y. Yamaguchi, “On reduction of number of reference data sets for antenna array calibration,” IEICE Technical Report, AP2002-40, pp.1–6, July 2002.
- 6) T. Inaba, T. Sakamoto, R. Miura, M. Oodo, and K. Araki, “A study on mutual coupling compensation method of array antennas,” IEICE Trans. Commun. (Japanese Edition), vol.J85-B, no.10, pp.1757–1769, Oct. 2002.
- 7) T. Kawamura and K. Shioiri, “Design of a small size sleeve antenna,” Tech. on Instrumentation and Measurement, IEE Japan, IM-19-030, pp.13–17, Nov. 2019.

- 8) Y. Ogawa and K. Itoh, "High-resolution estimation using the MUSIC algorithm," IEEJ Transactions on Fundamentals and Materials, vol.116, no.8, pp.671–677, Aug. 1996.
- 9) T. Mori, "Comparative study of antenna array calibration methods based on symmetrical mutual coupling of 4-element linear array," IEICE Technical Report, AP2022-178, pp.6–11, Jan. 2023.

Author

Takashi Mori
Technology Planning Dept.
Advanced Research Laboratory

RF Switch MMIC Based on InP DHBT Technologies for Sub-THz

Mohammad Salah, Alexander Chenakin

[Summary]

In the realm of RF/microwave industry the utilization of MMIC switches at Sub-THz emerges as a critical advancement for unlocking the full potential of the spectrum. The miniaturization and integration of multiple components along with switches on a compact single chip is a paramount consideration to enhance the versatility of Sub-THz systems. Sub-THz MMIC switches play a pivotal role in enabling Sub-THz communication systems, advanced imaging and sensing, and instrumentation. Optimizing insertion loss, isolation, and bandwidth of MMIC Switches is poised to lead the way in the next frontier of RF applications.

1 Introduction

Mm-Wave to Sub-THz system architectures sometimes use the same frequency for TX and RX which dictates using a frontend switch to select either TX or RX. The cost of power in the Sub-THz spectrum is relatively high compared to lower frequencies. This increases the trend of integrating switches on a single chip with other components because deploying them on separate chips requires either lossy interconnects or waveguide probes that increase the insertion losses. Which results in higher noise figures for RX and lower power and linearity for TX.

2 Technology

Quests have been made on which technology to use and the tradeoffs for designing wideband switches. GaAs PIN diodes have a proven record for low insertion loss and improved isolation. However, integration with other components on the same chip blocks exploiting this technology and limits it to a single package. Some pHEMT technologies have integrated PIN diode options but not at the sub-THz spectrum. The bandwidth of the switch is the first figure of merit that we look at, which is a direct function of the transistor used.

InP HBT-based switches have the advantage of high linearity³⁾ and higher frequency⁴⁾. However, the main challenges are isolation and power consumption. This is because the input bias on the base of the transistor is not highly isolated from the RF ports at the Collector/Emitter. This creates undesired effects on high frequencies and can be observed as the bias source moves away from the base of the transistor. The HEMT-based switches have the advantage of isolation but fall short.

on power and very low frequencies for ultrawideband applications. In this article, we will exploit the InP DHBT-based switches to enable the integration of switching to amplifiers and multipliers.

A multi-metal layer structure is used for this design. This not only gives a degree of freedom to move the RF ground but also to isolate some DC components and biases from RF circuit.

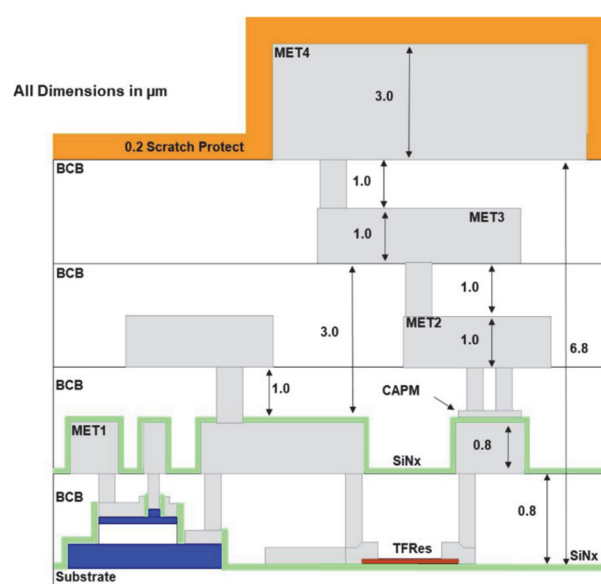


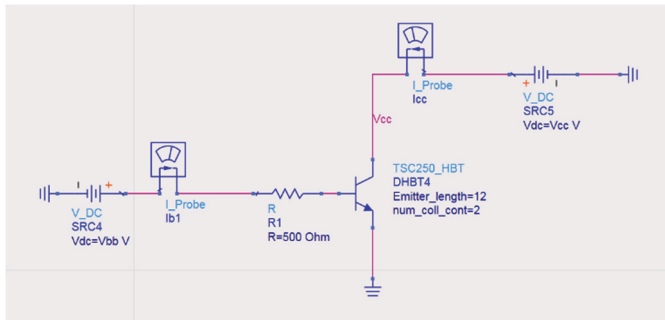
Figure 1 Representation to a cross-section of TSC250 IC technology, not to scale.⁶⁾

3 Design

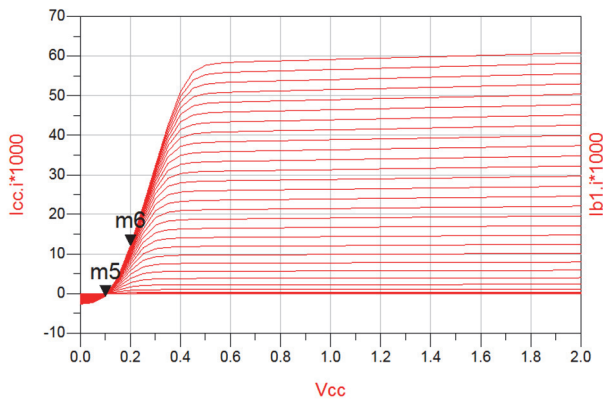
Starting with a 12 μm length DHBT transistor minimizes the parasitic capacitance at the expense of power handling. In this approach, the goal is to push the technology to its limits with the simplest topology to facilitate integration and achieve ultrawideband frequency response from DC to Sub-THz. Stacked transistors with matching circuits would play

a role in improving isolation which will be presented in a later section.

Figure 2 shows some bias points to be used. The realization of resistances and DC bias circuit in the layout of the chip plays a role in improving isolation and frequency response ripple. It is to be noted that it is not uncommon for design libraries to not have enough information about maximum base current limitations. This requires paying close attention to the physical structure of the transistor to estimate proper bias currents. The trade-off is between improving isolation versus damaging the transistor by overdriving it.



m5	m6
Vcc=0.100	Vcc=0.200
Icc.i*1000=-0.627	Icc.i*1000=12.194
Vbb=2.000	Vbb=2.000



m8
indep(m8)=1.300
plot_vs(Ib1.i*1000, Vbb)=1.131
Vcc=0.100

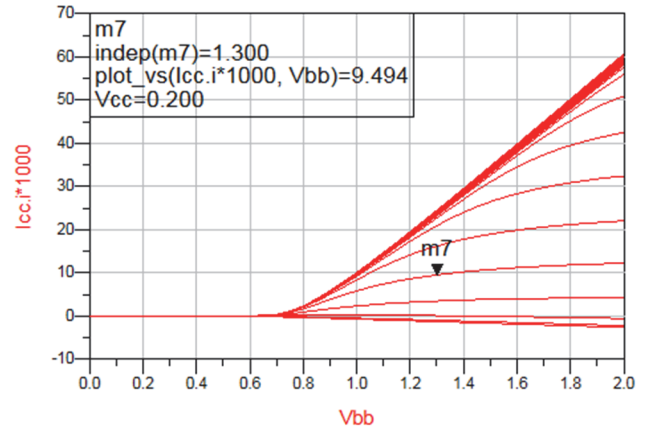
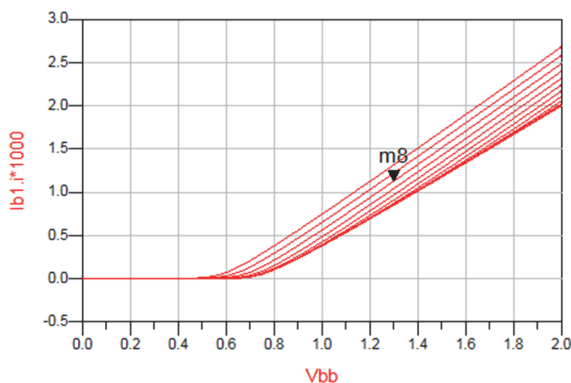


Figure 2 12 μ m DHBT I-V characteristics and bias points.

To better evaluate the potentials and limitations of the technology, several topologies are investigated. Figure 3 shows a simple series through a switch in which the insertion loss is not the limiting factor for the ON state, but the OFF position limits the isolation because of the off capacitance. There are multiple ways to address this issue such as series cascading or isolation with shunt transistors. The second one investigates shunt termination. Figures 3a and 3b can handle from DC to Sub-THz, only the SPDT in 3c cannot go to lower frequencies but can handle higher power. The series 12 μ m length transistor in a through configuration can achieve a 1 dB insertion loss from DC to 300 GHz in an on position as shown in Figure 4. However, the off capacitance was evaluated to be 8.75fF for the selected bias conditions. This has a direct limit to the maximum frequency. Interestingly, a stacked topology can be used to improve isolation and increase the maximum frequency at the expense of insertion loss. Since different systems have different switching requirements, the design topology of the switch can be adjusted to improve some figures of merit at the expense of others.

In general, achieving low insertion loss often comes at the expense of isolation and vice versa. Faster switching speeds generally lead to higher power consumption and broadening the frequency bandwidth of the switch may result in a larger physical size. It is crucial to understand the underlying physical implementation and tradeoffs to make a good design choice.

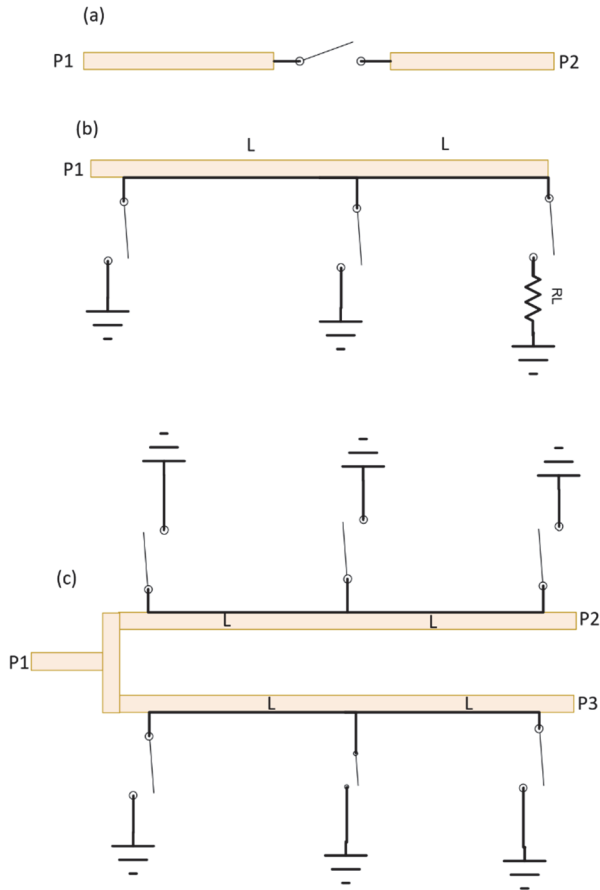


Figure 3 Block diagram of several Sub-THz switch approaches, a) simple series SPST b) Short, Load, and Open arm c) SPDT with multiple quarter wavelength sections.

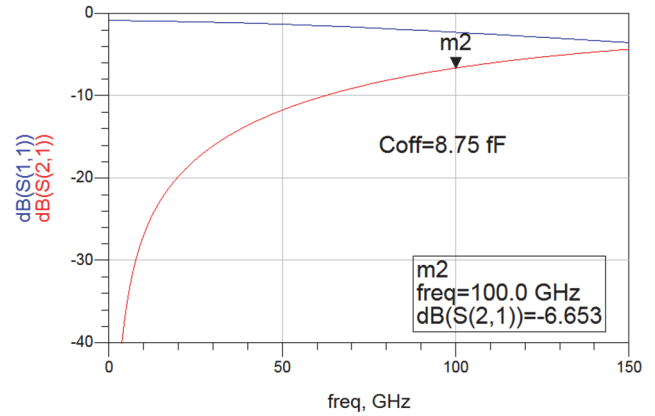
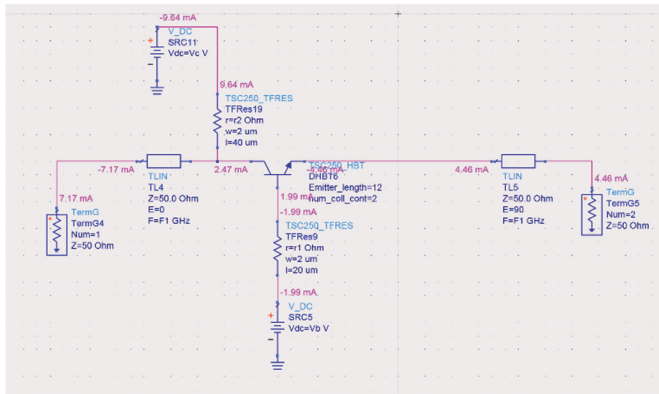


Figure 4 A Thru configuration where Ron of 12 ohms and Coff of 8.75 fF is achieved based on selected bias points. Coff limits the Bandwidth for isolation.

A Short, Open, and Load Switch arm shown in Figure 5 demonstrates frequency response from DC to 150 GHz. At the Sub-THz band, the electrical delay between open and short can be noticed on the Smith chart in Figure 5. The Capacitance on the off position can be manipulated through the bias conditions of each transistor as well as base resistance.

The challenge is that bias conditions cannot be easily modified after the manufacturing of the chip, this is because it is required to have the bias resistance on the chip close to the transistor.

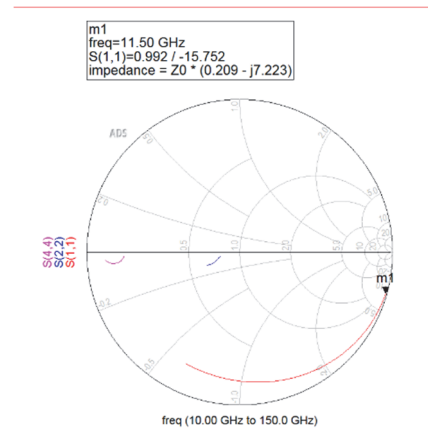
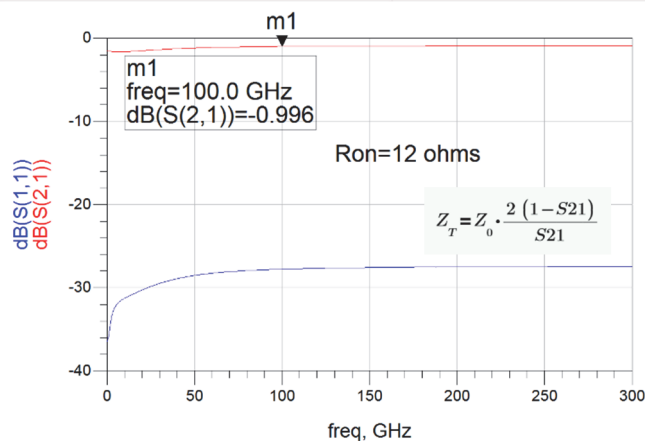
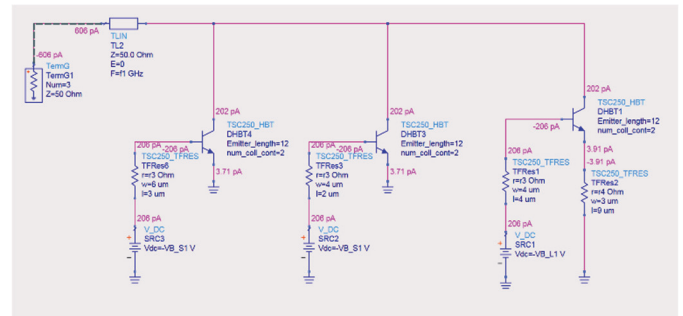


Figure 5 A Short, Open, and Load arm from DC to 150 GHz.

4 Physical Layout

The physical layout requires impedance matching and considerations of the 3D structure, since this is a multi-metal layer structure, this degree of freedom must be exploited to address the challenges of leakages or resonances that could happen anywhere in the spectrum⁵⁾.

The layout design takes into consideration the challenges associated with operating at these ultra-high frequencies, including signal losses, parasitic capacitances, and transmission line effects. Careful optimization of the InP HBT switch layout addresses these challenges, ensuring efficient signal integrity and minimizing the impact of parasitic elements.

The port launch demonstrated in Figure 6 is optimized on the multi-metal layer structure. The RF ground at the launch is different from the RF ground on the rest of the chip.

For waveguide packaging, this launch must be replaced with an on-chip probe.

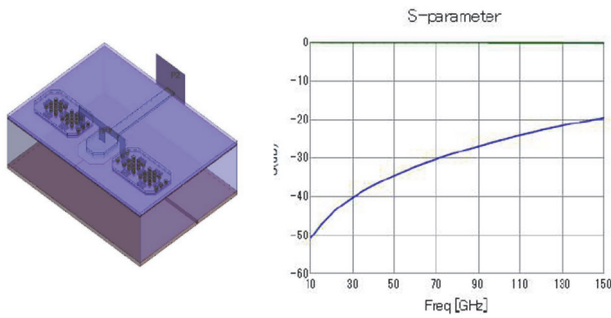


Figure 6 Port Launch optimization on a multi-metal layer structure.

The open-ground structure beneath the transmission line in Figure 7a is required for the design to be able to place circuits such as transistors and other biasing circuits.

This is problematic and requires another careful layout optimization for ultrawideband Sub-THz applications. Only this window can cause significant spikes as shown in Figure 7b, after optimizing the 3D structure in the layout an improvement can be achieved to avoid those resonances as shown in Figure 7c.

Figure 8 represents a series-thru configuration. note the necessity of the offset disconnect in the transmission line.

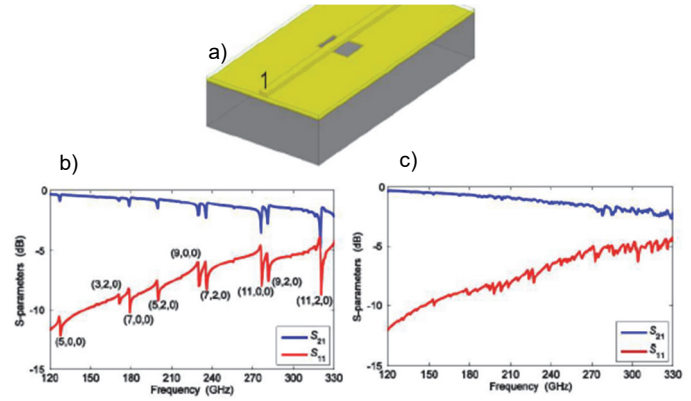


Figure 7 a) an open structure for transistors or bias has a significant effect when considering 330 GHz devices.
b) insertion and return loss before layout optimization.
c) insertion and return loss after layout optimization.

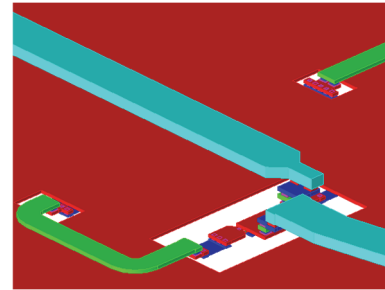


Figure 8 The physical layout of a series thru.

Finally, a complete EM simulation and optimization are done to address design rule checks (DRC) as well as consider manufacturing tolerances. A multi-switch module was successfully integrated into a 1×1 mm chip shown in Figure 9.

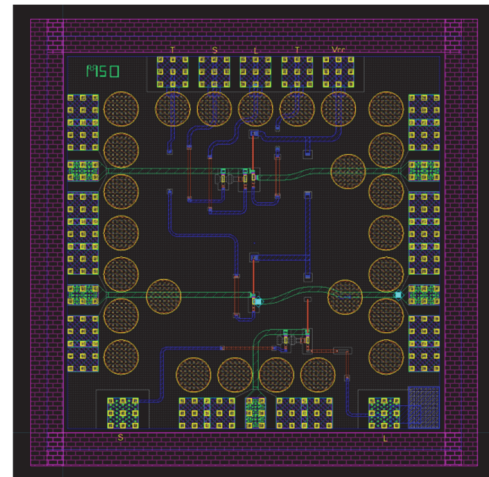


Figure 9 1×1 mm multi-switch chip module.

5 Conclusions

This article reports an investigation into RF Sub-THz MMIC switch utilizing InP DHBT which yielded promising results and demonstrated significant potential for miniaturized advanced RF systems in the from DC to Sub-THz

spectrum. The unique properties of InP DHBT such as electron mobility and thermal conductivity contribute to enhancing the performance and making them a compelling choice for advanced telecommunications.

Furthermore, the investigation has addressed key challenges such as ultrawideband from DC to Sub-THz, insertion loss, and isolation. Although DHBT is not the best choice for isolation, this poses a burden on the design effort to achieve the desired isolation and could be at the expense of maximum frequency. The increased power handling of DHBT over HEMT solidifies the foundation of using them on front-end ultrawideband transmitters. MMIC switches for use in emerging communication systems, terahertz imaging, and other high-frequency applications. As technology continues to advance, the outcomes of this investigation contribute to the ongoing evolution of communication networks and solidify the position of InP DHBT as a promising material for future MMIC switch innovations.

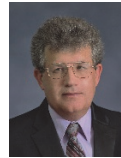
References

- 1) W. Khan et al., "A D-band (110 to 170 GHz) SPDT switch in 32 nm CMOS SOI," in IEEE MTT-S Int. Microw. Symp. Dig., May 2015, pp. 1–3.
- 2) A. Ç. Ulusoy et al., "A low-loss and high isolation D-band SPDT switch utilizing deep-saturated SiGe HBTs," IEEE Microw. Compon. Lett., vol. 24, no. 6, pp. 400–402, Jun. 2014.
- 3) T. Shivan et al., "Highly linear 90-170 GHz SPDT Switch with High Isolation for Fully Integrated InP Transceivers," 2019 IEEE MTT-S International Microwave Symposium (IMS), Boston, MA, USA, 2019, pp. 1011–1014.
- 4) T. Shivan et al., "220–325 GHz high-isolation SPDT switch in InP DHBT technology," in Electronics Letters, vol. 54, no. 21, pp. 1222–1224, 18 10 2018.
- 5) Johansen TK, Doerner R, Weimann N, Hossain M, Krozer V, Heinrich W (2018). EM simulation assisted parameter extraction for transferred-substrate InP HBT modeling. International Journal of Microwave and Wireless Technologies 10, 700–708.
- 6) Griffith, Zach, Miguel E. Urteaga, Petra Rowell and Richard L. Pierson. "190-260 GHz High-Power, Broadband PA's in 250 nm InP HBT." 2015 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS) (2015): 1-4.

Author



Mohammad Salah
Service Infrastructure Solutions
US Division
Test & Measurement Company



Alexander Chenakin
Service Infrastructure Solutions
US Division
Test & Measurement Company

Publicly available

Switching Speed Measurements of Barium Strontium Titanate (BST) Varactors

Tom Roberts, Jon Martens

[Summary]

Varactors are used in a wide variety of adaptive and reconfigurable RF and Microwave circuits. For applications that require circuits to have fast re-configurability, characterization of the device's switching speed is typically required. This paper describes switching speed measurements of Barium Strontium Titanate (BST) varactors using a Vector Network Analyzer over a broad range of frequencies.

1 Introduction

In the past couple of decades, the radio spectrum has expanded and become crowded as it continues to support a wide variety of services and applications. Spectrum resources continue to grow as new spectrum allocations and communication standards are defined and made available for use. Reconfigurable or adaptive radios and radio front ends offer a scalable approach in simplifying hardware by tuning to a frequency and standard that is least used at a given moment in time^{1), 2)}. Tunable passive RF/Microwave devices are the building blocks of reconfigurable or adaptive electronics and modern radio front ends. These devices are also critically important in the use of steerable antennas, antenna arrays and tunable filters.

The speed or state transition time at which a passive RF/Microwave device can change states is important since this will, in most cases, impact system level performance specifications. Ideally, the measured speed in which a device changes states is not limited by the measurement system. Often for tunable devices which exhibit fast switching characteristics such as BST devices, it is the measurement system that will limit the Device Under Test (DUT) characteristics observed. In this paper, an Anritsu MS464XB Vector Network Analyzer is used to measure the rise/fall time characteristics of varactors. This measurement process can be applied to any switching device but is specifically interesting when phase carries the transition information.

2 Measurement Block Diagram

2.1 Pulse Measurement VNA options

The switching speed measurement block diagram shown in Figure 1 uses an Anritsu MS464XB Vector Network Analyzer

(VNA) with options 035 and 042³⁾. Option 035 is the IF digitizer option that receives, processes and records digitally the analog IF information for specific measurement time durations. Option 042 called PulseView™ is a User Interface (UI) that adds the capability to generate, trigger and measure pulsed signals. The VNA uses an external pulse generator⁴⁾ and pulse amplifier^{5), 6)} to drive the varactor from 0 V to +20 V DC. An oscilloscope monitors the pulse input leading and trailing edges to the varactor. At any given CW frequency, the VNA is controlled to digitally capture the change in magnitude and phase of the transmission parameter S21 as the varactor pulse input transitions from 0 V to +20 V and +20 V to 0 V. The S21 magnitude and phase 10%-90% transition times are compared with those observed on the oscilloscope. Since the VNA measures the relative change in transmission magnitude and phase over each pulse transition, a VNA receiver calibration is not necessary. Also, either the change in transmission magnitude or phase (not necessarily both) is needed to quantify actual DUT switching speeds.

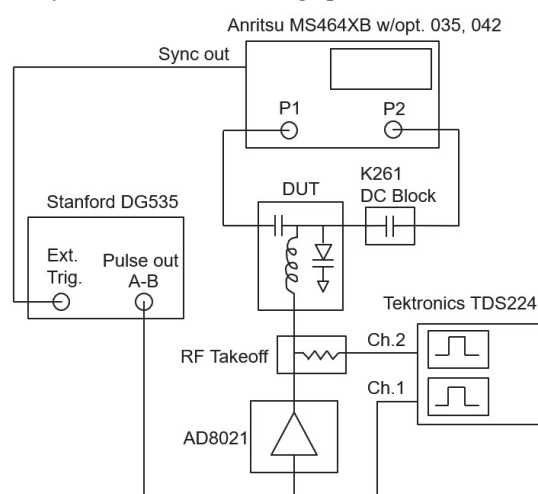


Figure 1 Switching speed measurement block diagram using Anritsu MS464XB w/options 035 and 042

2.2 Pulse Generation

Measuring and quantifying varactor switching speed over its full tuning voltage range of 0 to +25 V is desirable. The VNA externally triggers an external pulse generator which only has an output of 4 V peak and approximately 10 nS rise and fall times. A high-speed amplifier increases the pulse generator's output voltage amplitude to bring the signal level to +20 V while trying to preserve rise and fall times of the pulse generator. For the high-speed amplifier selection, compromises were made of both peak amplitude and rise/fall times. Analog Devices AD8021 was selected with a voltage gain of 5 giving a peak voltage output of +20 V and rise/fall times of about 450 nS and 150 nS, respectively. This amplifier was assembled and evaluated on an evaluation printed circuit board (PCB).

3 Varactor DUT Model Development

3.1 Electrical Models

Previous work⁷⁾ characterized Tecdia Varactors⁸⁾ to derive electrical models as a function of tuning voltage. This data was used to create an electrical model in AWR Microwave Office⁹⁾ to evaluate the change in transmission magnitude and phase vs. frequency over a range of tuning voltages.

Figure 2 shows a simplified electrical model of the DUT circuit. The tuning voltage is applied to "Pulse In" end of the RF choke (RFC). The RFC is modeled as a 3.5 mm long bond wire to provide sufficient inductance above 2-3 GHz but not so much inductance to limit the input pulse edge rise and fall times. Figures 3 and 4 show the change in transmission magnitude and phase respectively of the HBC2R2KY20X10X5A02 (2.2pF) varactor design. The absolute values in Magnitude and Phase are not so much of a concern but it is the relative change in magnitude and phase that gives confidence the VNA will be able make the measurement with enough vertical scale resolution. Note in Figure 4, the phase change at 10 GHz is small so measurement resolution will be limited. Determining the rise/fall times may be difficult at this frequency. This is not a problem, however, since the transmission magnitude shows sufficient relative change at this frequency. Again, measurement of either transmission magnitude or phase is necessary to determine DUT switching speed. The transmission magnitude and phase electrical model results of the HBCR22AY20X10X5A02 (0.22pF) varactor are similar and not shown.

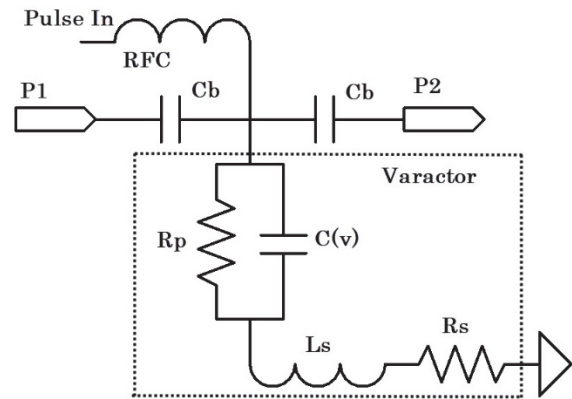


Figure 2 DUT electrical model

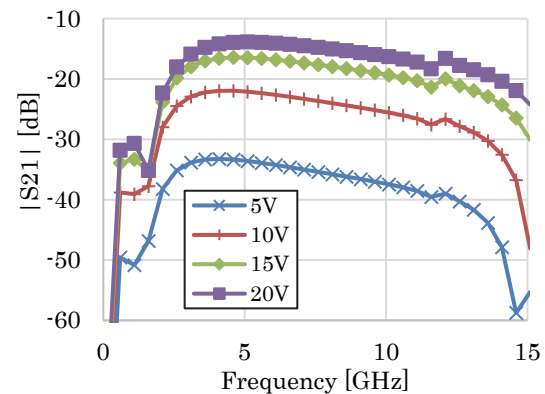


Figure 3 Transmission Magnitude vs. frequency vs. varactor voltage

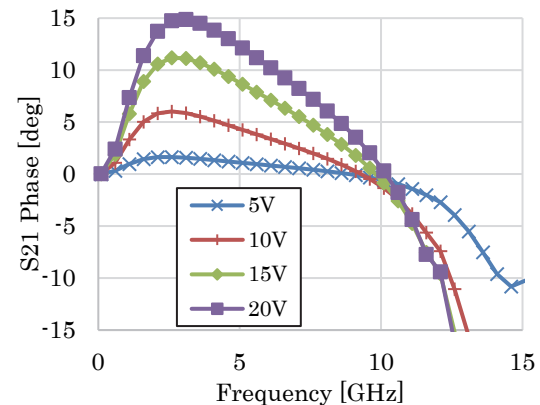


Figure 4 Transmission Phase vs. frequency vs. varactor voltage

3.2 Varactor DUT Modules

DUT physical models were built using a 10-mil thick Alumina circuit using 2.2 pF and 0.22 pF shunt mounted varactors. Segmented wire bonds were necessary to achieve the 3.5 mm bond wire length and required inductance value. Anritsu K connectors were used with a ribbon wrap pin-substrate connection on the connector's backside. An SMA connector was used to connect to the Bondwire RFC. Figure 5 shows the DUT internal circuit.

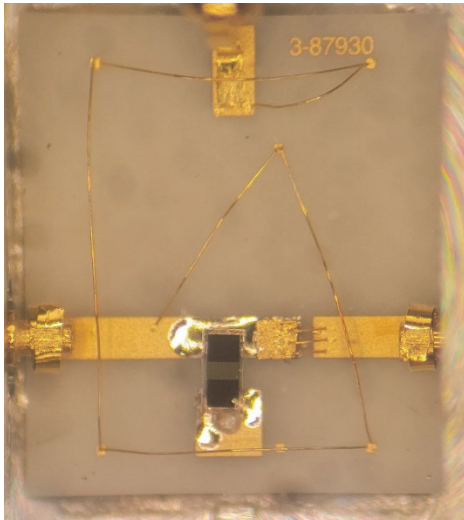


Figure 5 Shunt mounted varactor on 10-mil Alumina Substrate

4 Switching Speed Measurements

4.1 Instrument Settings

The Anritsu MS464XB VNA Option 042 PulseView™ (Option 042) UI settings are shown in Table 1. The measurement mode is selected as Pulse profile since structure of the data within the pulse is of interest. In this measurement, S21 magnitude and phase vs. time is captured at the pulse rising and falling edges. In this mode, both frequency and power are held constant. The pulse amplifier's 10%-90% rise and fall times are approximately 100 nS so the Measurement Width parameter is set to 10% of this value. It is expected the pulse amplifier will limit how fast the varactor can be switched. The VNA receiver is set to b2 (Port 2) which implies S21 Magnitude and Phase will be measured.

The Pulse Repetition Interval (PRI) describes the period of the pulse input to the DUT. The first half of the interval the pulse to the DUT will be at a high state and the last half of the interval, the pulse will be at a low state as a 50% duty cycle was employed. The Start/Stop interval of 99 uS to 101 uS captures the falling-edge state of the measurement. Conversely, a Start/Stop interval of 0 uS to 2 uS captures the rising-edge state of the measurement. The DG535 pulse generator's pulse repetition interval is set to 200 uS, with a duty cycle of 50% and a pulse amplitude of 4 V. Changing the duty cycle of the pulse generator requires a change in Start/Stop interval when measuring the falling-edge state of the DUT.

Table 1 Anritsu MS464XB PulseView™ configuration settings

Parameter	Value
Measurement Mode	Pulse Profile
Receiver	b2
# of Points	401
Measurement Width	10 nS
Start	99 uS
Stop	101 uS
Control Type	Synch
Pulse Repetition Interval	200 uS
Synch	Internal
Resolution	2.5253 nS

4.2 Pulse Amplifier Output

In Figure 1, the pulse amplifiers output is monitored using Ch. 2 of the TDS224 oscilloscope through a resistive take-off. The pulse amplifiers input is monitored using Ch. 1. Monitoring the pulse input to the DUT is useful to ensure the pulse amplifier is not loaded by the DUT resulting in a distorted pulse waveform. The pulse amplifier output response can also be compared to the VNA S21 magnitude and phase edge transition response. The AD8021 pulse amplifier exhibits asymmetry between its rise and fall times of about 450 nS and 150 nS respectively with the 150 nS fall time limited by the slew rate of the amplifier. Plots of the Oscilloscopes Ch. 1 and Ch. 2 rise and fall times are shown in Figures 6 and 7.

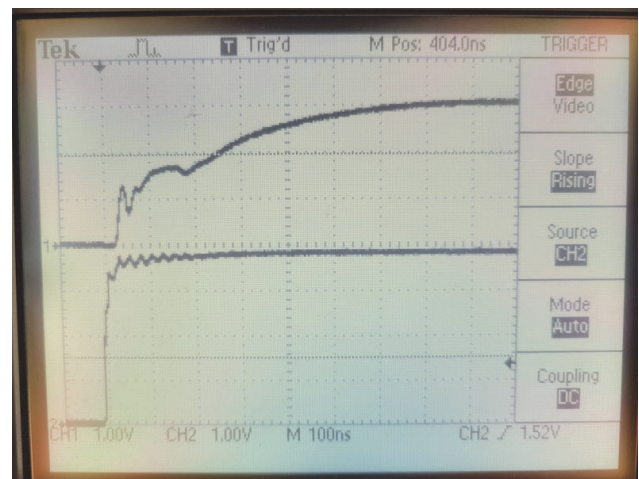


Figure 6 Pulse amplifier risetime, input (bottom trace) and output (top trace)

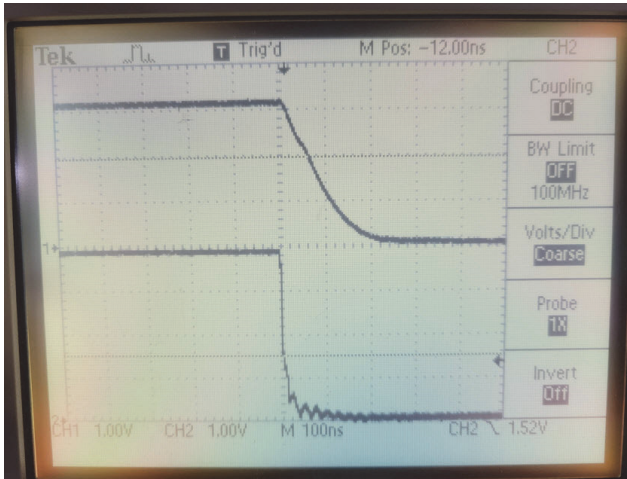


Figure 7 Pulse amplifier falltime, input (bottom trace) and output (top trace)

4.3 MS464XB VNA PulseView™ Measurements

If the switching speed of the varactor is limited by the 10%-90% rise/fall times of the pulse input to the DUT, the transmission magnitude and phase transition times will mirror the pulse rise/fall times. Using the MS464XB VNA PulseView pulse profile at continuous wave (CW) frequencies at 5, 10 and 15 GHz, the 10%-90% rise time can be measured from both S21 transmission magnitude and phase as shown in Figures 8, 9 and 10. Note that the relative trace noise is generally lower on the phase measurements which shows some of the value in using the complex S21 pulsed measurement. Since the DUT is a capacitor, this improved phase sensitivity would be somewhat expected. These risetimes are identical to DUT pulse input risetimes. For slight changes in either magnitude or phase, the measurement point scatter tends to increase or widen but risetimes can still be measured with transitions as small as 0.2 dB in magnitude and 0.5 deg in phase. The corresponding fall time plots are shown in Figures 11, 12 and 13 which show a 10%-90% fall time of about 150 nS.

Only three CW frequencies are presented in this paper but any CW frequency within the range of the VNA could have been used. Rise and fall times for other CW frequencies beyond 15 GHz were measured and all exhibited identical 10%-90% rise and fall time durations, albeit at different absolute magnitude and phase values.

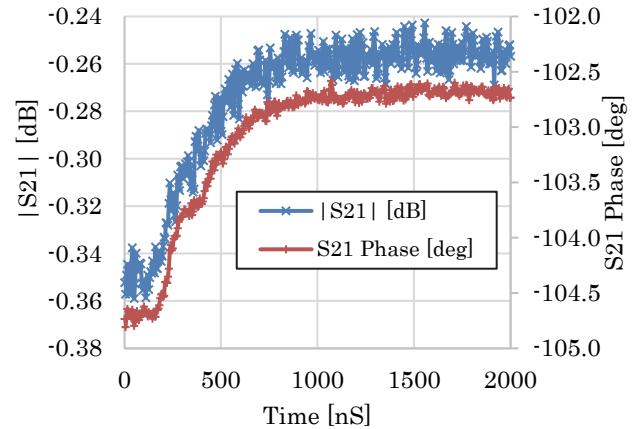


Figure 8 Risetime delta magnitude and phase at 5 GHz for 0.22pF varactor

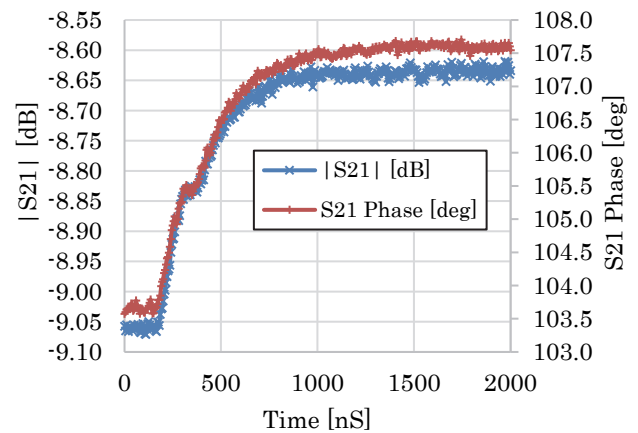


Figure 9 Risetime delta magnitude and phase at 10 GHz for 0.22pF varactor

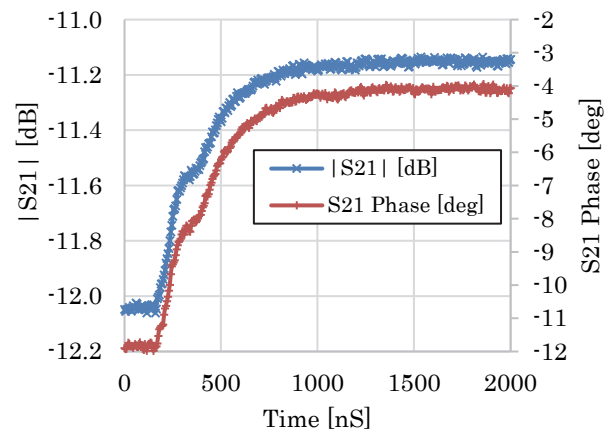


Figure 10 Risetime delta magnitude and phase at 15 GHz for 0.22pF varactor

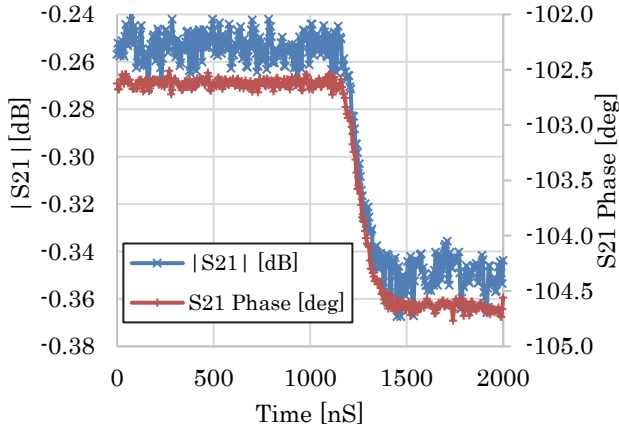


Figure 11 Falltime transmission magnitude and phase at 5 GHz for 0.22pF varactor

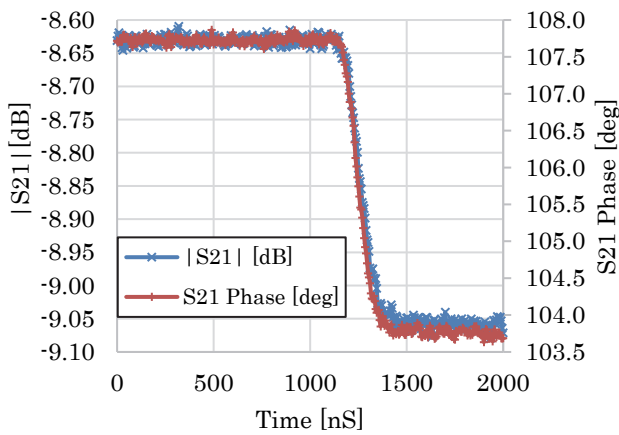


Figure 12 Falltime transmission magnitude and phase at 10 GHz for 0.22pF varactor

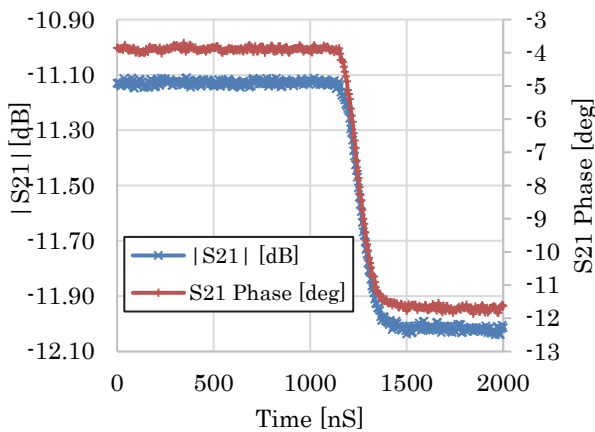


Figure 13 Falltime transmission magnitude and phase at 15 GHz for 0.22pF varactor

4.4 Varactor Current and Power Dissipation vs. Pulse Duty Cycle

All measurements in the previous sections were performed using a pulse duty cycle of 50%. With a duty cycle of 50%, the pulse amplifier positive current (I_{cc}) is 37 mA. The quiescent current of the pulse amplifier with no input drive is 10 mA.

Figure 14 shows Pulse amplifiers supply current vs pulse duty cycle with I_{cc} linearly proportional to the pulse duty cycle. The amplifier's negative current (I_{ee}) remained relatively constant at 6 mA. I_{cc} is used to calculate the varactor average power dissipation for a given duty cycle given by equation 1 and plotted in Figure 15.

$$P_v = \frac{\text{DutyCycle} [\%]}{100} * V_p * I_{cc} \quad (1)$$

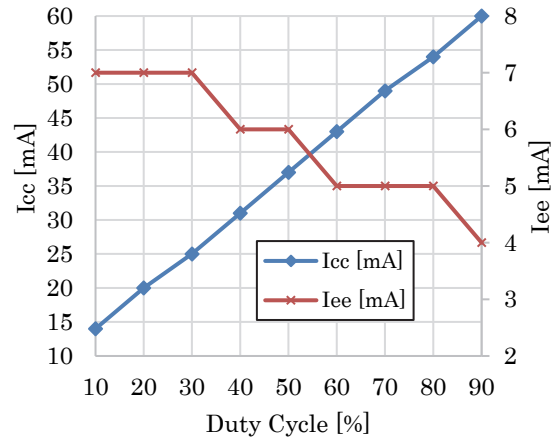


Figure 14 Pulse amplifier I_{cc} and I_{ee} vs. pulse duty cycle

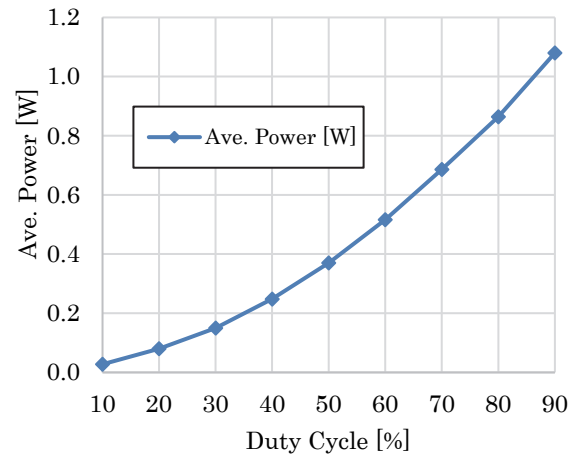


Figure 15 Varactor average power dissipation vs. pulse duty cycle

5 Conclusions

Characterizing the speed of switching devices is easily accomplished using an Anritsu MS464XB Vector Network Analyzer (VNA) with options 035 and 042. In this paper, switching speed of BST varactors were characterized using a digitized time interval sample of transmission magnitude and phase at specific frequencies. The change in either the transmission magnitude or phase with time can be used to determine device switching speed. The rise and fall times of either of these parameters is compared to the respective rise and

fall times of the pulse input to the DUT.

For this work, the pulse amplifier limited how fast the varactor could be switched over its voltage bias range of 0-20 V. Due to the limited availability of high-speed op amps allowing a full +20 V output voltage, one compromise could be selecting a higher speed op-amp with a lower output voltage of +5 V for example. This compromise would allow a greater selection of high-speed op-amps to choose from. The amplifier's lower output voltage requirement translates to a high Gain Bandwidth Product (GBP) since the amplifier's gain would be 5 V/4 V or 1.25 maximizing its operational bandwidth.

References

- 1) S. Gevorgian, "Agile microwave devices," in IEEE Microwave Magazine, vol. 10, no. 5, pp. 93-98, August 2009, doi: 10.1109/MMM.2009.932830.
- 2) J. Hong, "econfigurable Planar Filters" in IEEE Microwave Magazine, vol. 10, no. 6, pp. 73-83, September 2009, doi: 10.1109/MMM.2009.933590.
- 3) Anritsu VectorStar MS4640B Series <https://www.anritsu.com/en-us/test-measurement/products/ms4640b-series>
- 4) Stanford Research Systems DG535 Pulse Generator <https://www.thinksrs.com/products/dg535.html>
- 5) Analog Devices AD8021 High Speed Amplifier <https://www.analog.com/en/products/ad8021.html>
- 6) Analog Devices 8-Lead SOIC Amplifier Evaluation Board, PN: 038383B <https://www.analog.com/media/en/technical-documentation/user-guides/ug-755.pdf>
- 7) T. Roberts, D. Young "Broadband Device Modeling and Q Calculations of Barium Strontium Titanate (BST) Varactors," Anritsu Technical Review No. 31, March 2023
- 8) Teedia Inc. Tunable Dielectric Capacitor (Dielectric Varactor) PN: HBC2R2KY20X10X5A02 (2.2pF), PN: HBCR22AY20X10X5A02 (0.22 pF), <https://us.tecdia.com/products/hf/detail/slc/tunable-dielectric-capacitor/>
- 9) Cadence AWR Microwave Office V16, Cadence Design Systems, Inc, 2655 Seely Ave, San Jose, CA 95134, <http://www.cadence.com/>

Author



Tom Roberts
Service Infrastructure Solutions
US Division
Test & Measurement Company



Jon Martens
Service Infrastructure Solutions
US Division
Test & Measurement Company

Publicly available

Synchronized Vector Network Measurements Over Long Distances

Don Bradley, Paul Davis, Aaron Miller, Jon Martens, Navneet Kataria

[Summary]

The RF/microwave wireless industry has an increasing number of applications where synchronized vector network measurements over distance are required. Traditionally, elaborate systems involving long front-end cables have been employed to make these measurements. These historic systems are inherently unstable and therefore inaccurate. This article reports on a new vector network analyzer (VNA) architecture that dramatically improves measurement stability and accuracy for port separation distances of up to 200 meters.

1 Introduction

A key aspect of microwave VNA testing is minimizing the measurement effect of the physical interface to a device under test (DUT). Using high stability, low insertion loss test port cables and applying techniques like reference plane extension and fixture de-embedding, VNA users have been very successful for years in minimizing the effect of the path from the VNA port to the DUT on typical microwave and millimeter wave VNA s-parameter measurements. These techniques are less successful with longer distances or larger DUTs at high frequency as cable insertion loss and instability grows significantly over distance making it more difficult to compensate for the path using typical bench techniques.

Long cable runs between the VNA port and the DUT occur in applications like Over-the-Air (OTA) testing in chambers and outdoor ranges as well as measuring RF and microwave characteristics of large vehicles like ships and aircraft. One key issue to overcome in these types of applications is the insertion loss of the cable affecting the stability and accuracy of the VNA measurements.

2 Modular VNA

To overcome the loss issues, many have used external mixing structures¹⁾⁻³⁾ linked back to a parent VNA. While this can be successful, the setups can be quite complicated and stability can still be an issue. A novel solution to this issue is the concept of a modular (or distributed) VNA⁴⁾. Instead of having a single chassis instrument where its ports are co-located (and possibly an elaborate external mixing setup), a modular VNA has independent portable modules which allow them to get close to the DUT with minimal cabling (see Figure 1). With source and measurement capability in the

port modules, the main issue of insertion loss and distortion due to long coax or fiber interconnect disappears, and stability and accuracy of the measurements improves.

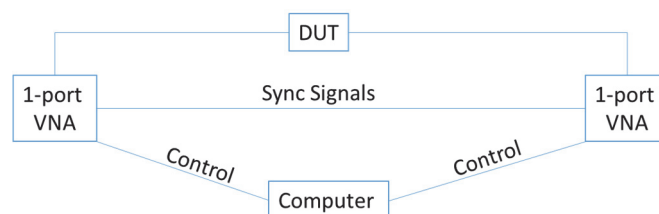


Figure 1 Modular VNA block diagram

Setup of a long-distance measurement is simplified as portable ports make it possible to move the VNA to the DUT eliminating the need for long interconnect cables. Calibration becomes much easier as the calibration plane moves close to or at the port rather than at the end of meters of cable through active components. Without active mixing or fiber optic translation in the measure path, attaining the DUT performance with a modular VNA is straightforward without requiring complex de-embedding.

While portable independent VNA ports eliminate the interconnect issues of long-distance measurement, it does create a major hurdle of providing the precise synchronization required between ports needed to support microwave vector (magnitude and phase) S-parameter measurements.

3 Practical Implementation

Assuming one is not driving both ports simultaneously in some forced phase, the main objectives for synchronizing the modules are:

1. A common LO for both transmit and receive modules
2. A means of sharing frequency references between modules so ADC clocks can be synchronous

The latter objective is met by one module, the Base module, that is defined at start-up time. The Base module supplies a frequency reference for the other module's synthesizers, and ADC synchronization. The former objective is met by literally sharing an LO between modules as shown in the diagram below. This is possible with an optical link (E/O conversion at one end and O/E conversion at the other) between modules to keep insertion loss to a minimum as shown in Figure 2. Not all of the details are shown in this figure for simplicity:

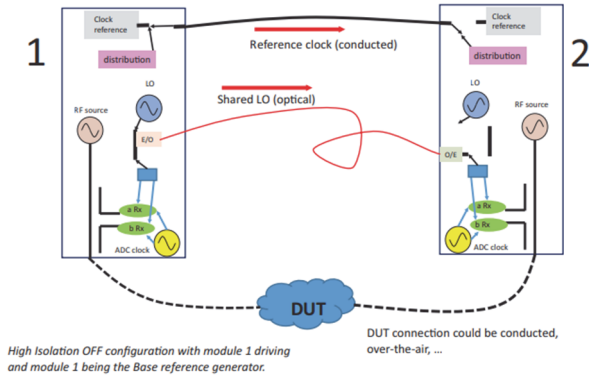


Figure 2 The basic assembly of two modular VNAs is shown here. A cable harness connects the ADC synchronizing signal (and control signals) from one module to the other (the module sourcing these signals is defined at startup) and an optical fiber allows the sharing of an LO. USB cables connecting each module to the PC are not shown here.

This LO connection is an analog signal link so the phase and amplitude information embedded in the LO is transferred (although the amplitude information is limited due to a strong limiting operation). While the insertion loss is low in the fiber link, the optical fiber is not immune to dielectric constant changes with temperature and/or flexure (although less than that observed with a typical coaxial cable) so there could be phase shift in the measured transmission response. To understand this, consider how (for example) S_{21} is constructed from the raw measurement data in an initial state with no LO path electrical length change:

$$S_{21}^{original} = \frac{b_2}{a_1} \approx \frac{K e^{j\theta_{LO}} B e^{-j\theta_{RF}}}{K e^{j\theta_{LO}} A e^{-j\theta_{RF}}} = \frac{B}{A}$$

If, now the phase on the LO path between the modules changes by ϕ , the relationship changes to:

$$S_{21}^{new} = \frac{\tilde{b}_2}{\tilde{a}_1} \approx \frac{K e^{j(\theta_{LO} + \phi)} B e^{-j\theta_{RF}}}{K e^{j\theta_{LO}} A e^{-j\theta_{RF}}} = \frac{B e^{j\phi}}{A}$$

Had the phase change been common to both the test and reference paths, the phase would have cancelled out. Since the receivers are separate and the longer run exposed, this kind of un-cancelled phase change is a possibility. There are unique methods on the MS46131A to minimize effects of changes to the link.

4 Phase Compensation

Because of the potential large separation of modules (over 100 m in some cases) and the dependence of optical fiber dielectric constant on temperature, pressure, etc., noticeable environmentally-induced phase change is possible. While this might not present too much of a problem in some measurements (e.g., magnitude-only antenna pattern measurements), it can be an issue in others (near-field measurements, for example). A phase compensation algorithm is available where an additional measurement of just the LO path is made periodically and compared to an initial measurement and the phase difference (on a per frequency basis) is computed and used to correct the data. The baseline 'initial' measurement can be re-initialized at any time. Thus monitor measurements are used to periodically assess the state (in an electrical length sense) of the optical fiber link. When this mode is turned on and a new user RF calibration is performed, the baseline phase is automatically re-initialized. Again, because it is the relative LO phase between modules that is of concern here, it is mainly transmission phase that is affected. In highly transmissive devices, there is a coupling to load match so reflection measurements in that class can also be affected.

As an example, consider a setup where a localized cold source is positioned near the optical fiber cable. S_{21} of a transmissive device is being measured and a reading is taken before the temperature change occurs and then again after a two-second blast of $\sim -40^\circ\text{C}$ cold spray. Admittedly this is not a precise thermal experiment but it is a bounding case for thermal shock events that can occur in practical measurement setups. The results with and without phase compensation applied (the optical cable was allowed to return to its initial starting temperature between the experiments) are shown in Figure 3.

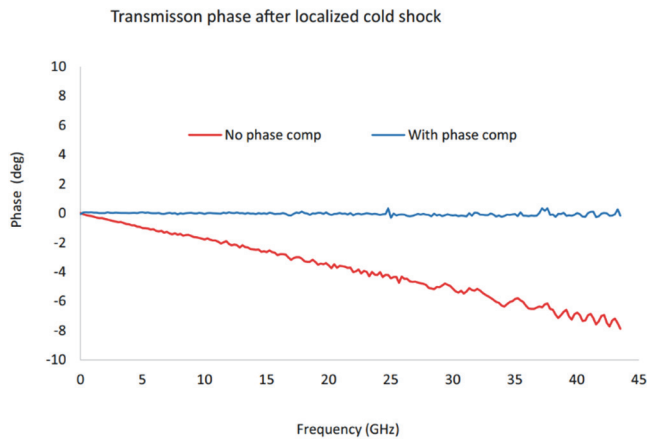


Figure 3 The stabilizing effect of the phase compensation process is illustrated here during thermal shock experiments.

Note that the measurement was done with two different calibrations and at two different points in time but the data is suggestive of the improvement. The algorithm is not limited to handling small phase excursions and can be useful in more thermally and mechanically dynamic environments.

It should be pointed out that, had conventional coaxial cables been used between the modules, not only would the separation distance have been more limited (due to cable loss) but the temperature sensitivity could have been similar or greater. In contrast, the LO sharing cable of the distributed VNA is only subject to phase drift (which can be compensated for as discussed); amplitude stays virtually unaffected.

5 Distributed Power and USB Technology

There exists a control module in this distributed VNA system which is located near the control PC running the VNA application software. This control module provides both power and USB communication to the individual port modules located x meters away. USB communication is over fiber and uses standard commercial technology. Distance for this link is not a limiting factor. Dual isolated floating 36 V supply voltages are transported to the individual VNA modules x meters away. Once received, each VNA module internally regulates the 36 V DC to 12.5 V DC needed for VNA operation. 36 V DC allows up to 200 m of power cable before IR drop prevents regulation at the VNA head. The DC voltage from the control module is supplied through a shielded coax cable and is floating with respect to control module ground. There is a 1 Meg Ohm resistance and 20 nF capacitance to control module ground which prevents power return current

from traveling along the PhaseLync™ Electrical (PLE) return shield. The DC voltage is derived from a switching power supply operating at 500 kHz. A 50 mV 500 kHz filtered sin wave ripple voltage is present on the 36 V DC. There are no 500 kHz magnetic fields emanating from the shielded power cable as the equal and opposite fields cancel due to the floating source.

6 Conclusion

For VNA applications over distance like far-field antenna measurements and large vehicle shielding and propagation characterization, several methods exist to deal with the insertion loss and other negative aspects of long interconnect cables. A new modular VNA architecture solves the issue by bringing the VNA port to the DUT eliminating most if not all of the cable interconnect for long distance applications. Measurement stability and accuracy are improved by removing the significant insertion loss due to long cables and by providing a compensation algorithm for the residual phase drift present in the fiber optic LO distribution cable. Setup and calibration are simplified by bringing the calibration plane to the VNA port and removing any active signal conditioning in the measurement path.

References

- 1) D. R. Novotny, A. J. Yuffa, R. C. Wittmann, M. H. Francis, J. A. Gordon, "Some Advantages of Using Bi-directional S-Parameters in Near-Field Measurements," 40th Proc. Antenna Meas. Techn. Assoc. Conf., Nov. 2018.
- 2) I. Ida, H. Yoshimura and K. Ito, "Reduction of drift error of a network analyser in small antenna measurement," IEE Proceedings -Microwaves, Antennas and Propagation, vol. 148, no. 3, pp. 188-192, June 2001.
- 3) J. Kang, J. Kim, N. Kang and D. Kim, "Antenna measurement using S-parameters," 2012 Conference on Precision electromagnetic Measurements, Washington, DC, 2012, pp. 658-659.
- 4) US patent 11,846,663, "Remote multiport modular vector analyzer system and method of use," and preceding patents.

Author



Don Bradley
Service Infrastructure Solutions
US Division
Test & Measurement Company



Paul Davis
Service Infrastructure Solutions
US Division
Test & Measurement Company



Aaron Miller
Service Infrastructure Solutions
US Division
Test & Measurement Company



Jon Martens
Service Infrastructure Solutions
US Division
Test & Measurement Company



Navneet Kataria
Service Infrastructure Solutions
US Division
Test & Measurement Company

Publicly available

Contribution to other Publications

Contribution to other Publications (April 2023 to March 2024)

Auther	Affiliation	Title	Publication
Takanori Saitoh	D	High accuracy OFDR by coherence compensation method	28th International Conference on Optical Fiber Sensors Th6.68 (2023).
Hayakawa Yuki Yusaku Narita	I I	Nanoscale fabrication of Graphene by Hydrogen-plasma Etching	e-Journal of Surface Science and Nanotechnology 20, 139-144 (2022)
Rin Takehara Yuki Uto Tomohiro Matsui Masahiro Hara	*1 *1 R *1	In situ thermal noise measurements under nanoindentation of suspended graphene	Nanotechnology 34 , 475704 (2023).
Ken-ichi Sasaki Kenichi Hitachi, Masahiro Kamada Takamoto Yokosawa Taisuke Ochi Tomohiro Matsui	*2 *2 R R R R	Introducing Corrections to the Reflectance of Graphene by Light Emission	C 10 , 18 (2024).
Yusa Takemoto Shigenori Mattori	R R	The Experimental Study of Four Types of an Elliptical Choke Operating in the WR-3 Full Band	IEICE Technical Report MW2023-80 (in Japanese)

R : Advanced Research Laboratory
I : Infivis Company
D : Sensing&Devices Company

*1 : Kumamoto University
*2 : NTT Basic Research Laboratories



ANRITSU CORPORATION

<https://www.anritsu.com>
5-1-1 Onna, Atsugi-shi, Kanagawa 243-8555, Japan