IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 1, 2019, accepted December 16, 2019, date of publication December 24, 2019, date of current version January 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2961943

An Integrated Detection Circuit for Transmission Coefficients

MING-CHE LEE[1] AND CHI-YO HUANG[2]
[1]Department of Mechanical Engineering, National Taipei University of Technology, Taipei 106, Taiwan
[2]Department of Industrial Education, National Taiwan Normal University, Taipei 106, Taiwan

Corresponding author: Chi-Yo Huang (cyhuang66@ntnu.edu.tw)

This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 106-2221-E-003-019-MY3, and in part by National Taiwan Normal University.

ABSTRACT As the applications of radio-frequency (RF) circuits continue to prosper, scattering parameters (S-parameters) play an essential role in the verification of a variety of chips. The traditional way to measure the S-parameters of RF integrated circuits (RFICs) is by using vector network analyzers (VNA). However, measuring RFICs with VNAs is very expensive and likely to reduce the profits of IC products. An implementation of the embedded circuit for S-parameter measurement can greatly reduce the costs of using expensive VNAs. Another reason to embed the circuit for S-parameter measurement is to increase the portion of a chip that can be measured. Besides, novel technologies, such as three-dimensional ICs, will require advanced methods for on-chip verifications of RF circuits since many RF nodes may be buried deep inside a chip stack. In view of these needs, this paper proposes a simple network that can realize on-chip S_{21} measurements. The greatest advantages of this circuit are the easy implementation and technology independence. To verify the feasibility of the circuit, we fabricated the test chips by using the 0.18-μm IBM 7RF process. The measurement results show the expected behavior and demonstrate the feasibility of the design concept.

INDEX TERMS Radio frequency (RF), S-parameters, integrated circuit (IC), vector network analyzer (VNA), automatic test equipment (ATE), divider (DIV), device under test (DUT), embedded testing, bandwidth, calibration.

I. INTRODUCTION

Due to the difficulty of achieving perfect open and perfect short at high frequencies, traditional 2-port parameters, such as Z-parameters, Y-parameters, and ABCD-parameters, are insufficient to fully characterize the properties of a 2-port network. To accommodate the insufficiency of traditional parameters, the scattering parameters (S-parameters), defined in terms of traveling waves, were proposed. In the 1970s, the S-parameters started attracting engineers' attention due to the launches of new network analyzers, which enable S-parameter measurements with ease [1]. Since then, S-parameters have become the mainstream parameters for characterizing the behavior of circuits, chips, or systems for radio-frequency (RF), microwave, and millimeter-wave applications. As RF circuits can nowadays be found in a variety of telecommunication systems, S-parameters play an essential role in the modeling and verification of a variety of devices, chips, packages, and transmission lines. In principle, S-parameters are defined as the ratios of the forward and backward propagating waves inside a circuit or a system. These waves generally contain both the magnitude and the phase information. Hence, being the ratios between these propagating waves, S-parameters are measured by complex values. This fact makes S-parameters exceptionally prominent and radically different from traditional measurement methods of analog, mixed-signal, and digital integrated circuits (ICs), in which only real-valued voltages and currents are measured.

Since an instrument for S-parameter measurement reads complex-number data, its design is normally more difficult and complicated than that of measuring devices that can measure only real-valued readings from analog, mixed-signal, or digital ICs at comparatively lower frequencies. Traditionally, S-parameters are measured by vector network analyzers (VNAs), which are very complex in terms of both

The associate editor coordinating the review of this manuscript and approving it for publication was Nosherwan Shoaib.

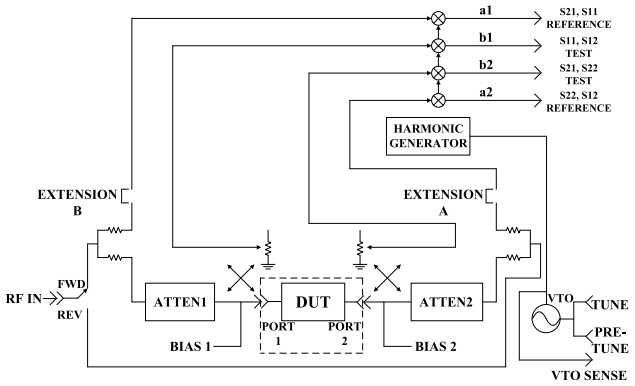

FIGURE 1. Block diagram of the hardware core of a VNA.

the microprocessor-based hardware and software. The hardware is responsible for imposing appropriate adjustments and guidance on the input RF signals and for correctly down-converting and sensing the output signals at the baseband. The software should be able to correctly filter out the nonidealities within the hardware, calculate and display the final measurement results, and serve as the interface with the microwave computer-aided design (CAD) software (e.g., HFSS, ADS, MATLAB). Fig. 1 demonstrates a typical block diagram of the hardware core portion of a VNA [2], [3]. According to the illustration, RF signals are attenuated and projected into Port 1 and Port 2 of a 2-port device under test (DUT). The responses of the DUT to the input RF stimuli are then appropriately manipulated by directional couplers, attenuators, and other switches. Finally, these guided signals are down-converted to the baseband. The resultant output signals, annotated as a1, b1, a2, and b2, can be further sensed by other amplitude and phase detectors.

Although the testing costs of analog or mixed-signal ICs are already higher than those of digital ICs, measuring S-parameters of RF ICs (RFICs) with VNAs is even more expensive than testing analog or mixed-signal ICs by automatic test equipment (ATE). This high testing cost of RFICs by measuring their S-parameters with VNAs sharply reduces the profit of microwave IC products, especially when the trend of systems-on-a-chip (SoC) is so prevalent today. Aiming to reduce the extremely high testing costs of available VNA-based solutions, this paper proposes the solution of embedded S-parameter measurement.

Further, embedded S-parameter measurement can increase the portion of a chip being measured. As the fabrication technology of semiconductors advances so rapidly nowadays, chip complexity is also increasing at an explosive rate. For a complicated product undergoing the design cycle, such high complexity makes it difficult for the internal nodes within the system to be tested. Hence, the product may become irreparable due to the very limited information available. Besides, designers may not know how to redesign the chip to debug or improve performance. As in the scan chain concept in memory testing, embedding the circuit for S-parameter measurement can increase the number of measurable nodes inside the RFICs. The diagnosis time will be reduced, which in return will reduce the testing costs and the time-to-market and increase the profitability. The new technologies, such as three-dimensional (3D) IC, will also require methods for on-chip verification of RF circuits because RF nodes may be buried deep inside a chip stack. In addition, measurement and verification of RF performance are also key applications of embedded circuits for S-parameter measurement.

Further, at least three other advantages of the embedded S-parameter measurement circuit can be identified. The first advantage is that circuits of interest can be directly and precisely measured without the intervention caused by the addition of a package. Without embedded circuits for S-parameter measurement, a chip can only be measured externally by a VNA. To do so, the chip needs to be placed inside a package, which helps bridge the chip and VNA on the outside. On the one hand, the addition of any package, regardless of the type, can only deteriorate a chip's performance because the very existence of a package inevitably forms an unnatural barrier that can attenuate, deflect, and delay the transmission of electromagnetic waves flowing through it. On the other hand, because the package issue is easy to overlook, and because the selection of packages often comes as an afterthought when the chip has already been fully developed, a high-performance die enclosed by an inferior package is a situation that is often seen. That is, this lack-of-prescience design style worsens the influence caused by the package. What is even worse, inferior packages are sometimes intentionally chosen to house high-performance silicon just to keep the package cost low, making the overall chip performance degrade even more. However, by measuring circuits on chip or on wafer, the problems caused by adding packages can be avoided. When the circuit being measured works in its normal operation mode, its performance will not be deteriorated by the redundant package. If the circuit works in its measurement mode, i.e., when it is being measured, the measurement results will be much more accurate without the noise and parasitic loading caused by the package. In other words, having embedded S-parameter measurement helps a chip work better in its normal operation mode and helps enhance the measurement accuracy when the chip is being measured.

The second advantage is the possibility of abandoning test fixtures. Because many RF packages cannot be probed directly, test fixtures are required to serve as a bridge between the package and the cables, which are further connected to the VNA. Although test fixtures have the advantages of lower costs and higher flexibility of port arrangement when compared to wafer probes and probe stations, they do have drawbacks as well. For example, test fixtures may impose unwanted parasitic loadings and other unfavorable effects, such as ground loops and RF reflections, at the interfaces. Besides, how to accurately de-embed becomes a formidable task because the calibration plane is often set at the edge of test fixtures instead of at the pads of chips. Fortunately, having embedded S-parameter measurement can help avoid these cumbersome problems.

The third advantage is that measurement results will more likely to be accurate than erroneous. A typical RF measurement system is composed of a VNA, RF cables, bias cables, and an interface to the DUT. The interface can be either on-wafer probes or test fixtures. Regardless of the type of interface, the only part of the whole measurement system mentioned above that really needs to be measured is the DUT. Except for the DUT, all the other components are extrinsic. These external components, responsible for measuring the DUT, are unfortunately exposed to the outside world, greatly increasing the chances of these components becoming susceptible to external unwanted influences. If any of these components suffers strong enough interference, the measurement results can turn out to be bad. For example, a VNA may malfunction due to extreme temperature or poor calibration quality; a cable may behave oddly due to material damage or accidental bend; an on-wafer probe may err due to tip damage or unbalanced position; or a test fixture may fail due to improper connections. If any of the above problems happens, the measurement system will be prone to generate wrong measurement results. Moreover, the above interference instances are just a portion of what can occur in the real world.

On the other hand, an embedded S-parameter measurement system, being rigidly fabricated on the same chip as the measured DUT, is less vulnerable to environmental influences from the outside world. Since the outside influence is too weak to cause measurement errors in the embedded S-parameter measurement system, the focus should be placed on the inside, which includes the quality of the fabrication technology, of the circuit design, and of the calibration methods. That is, if an embedded S-parameter measurement system can utilize a decent fabrication technology with little process variation, and can be well-designed and well-tested, it can generate data with much higher reliability than external measuring devices due to the embedded system's hidden and fixed structure for measurement. Hence, in terms of measurement error reduction, embedded S-parameter measurement is also favorable for the production of RFICs.

To verify the feasibility of the circuit, we fabricated the test chips targeted at the operating frequency of 11.25 GHz by using the 0.18-μm IBM 7RF process. All the devices are tuned to the desired performance around the targeted operating frequency. Each test chip occupies a total die area of approximately 1.2 mm $\times$ 1.2 mm, including the pads, and has three I/O ports, of which Port 1 is used as an input port and Port 2 and Port 3 are used as output ports. As for the pad configurations, Port 1 uses the famous G-S-G configuration, while Port 2 and Port 3 adopt the G-S configuration to achieve a better place-and-route arrangement. Measuring the S_{21} and S_{31} of the whole test chips with a VNA yields a ratio of S_{31} over S_{21} that is very close to the S_{21} of the DUT, an indication showing the validity of the proposed S_{21} detection circuit.

The remainder of this paper is organized as follows. Section II reviews various related works in the past. Section III explains the emergence of the design ideas, expatiates on the design specifications of the whole detection network, and elucidates the design concepts through a concrete derivation of mathematical equations. Section IV provides the overall results of this research, including the simulation results, the fabrication results, and the measurement results. Based on the findings in Section IV, Section V continues with further analyses, containing an in-depth examination of the measured data within the frequency range of interest, the possible integration of detectors in the future, and the limitations and niches of the proposed S_{21} detection network. Finally, Section VI summarizes the contribution of this research and discusses future research directions.

II. RELATED WORKS

Various scholars [4]–[9] have attempted to perform on-chip and off-chip S-parameter or S-parameter-like measurements. Although the design goals of various works differ, which can lead to distinct setups for tasks and can result in the adoption of dissimilar approaches, reviewing these related works beforehand is still beneficial. Hence, this section reviews several related works.

Jayaraman [4] used an on-chip peak detector to sense the output of a reconfigurable RF low-noise amplifier (LNA) operating at 2.4 GHz. The measured data are then fed into an off-chip calibration block, which is designed to generate corresponding feedback signals to tune the C_g and C_l of the LNA. By modifying the C_g and C_l, Jayaraman [4] claimed that the input matching and the gain of the LNA can be improved, which have been demonstrated through the measured magnitudes of the S_{11} and S_{21} of the LNA. The success of Jayaraman [4] can be attributed to the well-designed amplitude detector, the calibration block, and a reconfigurable RF LNA. In other words, Jayaraman [4] did not aim to measure the S_{11} and S_{21} of an arbitrary DUT on chip; neither did the work target to measure the S_{11} and S_{21} of the reconfigurable LNA. Instead, Jayaraman's work [4] focuses on using the on-chip peak detector and the off-chip calibration algorithm to generate suitable feedback signals, which, through the control by a varactor, can improve the matching and gain of the reconfigurable RF LNA.

Following the work of Jayaraman [4], a sequence of works that adopted the concepts of self-testing, self-correcting, and self-healing works [5]–[7] was proposed. Albeit all of these works proposed how to dynamically adjust the performance of LNAs, different approaches were adopted to attain the goals of self-healing. In general, the most obvious differences of the proposed approaches by the works mentioned above [5]–[7] can be classified into four main parts. The first part is the method in which RF signals are sensed. For example, the work of Goyal [5] uses a band-pass filter to sense the RF signals of an RF LNA. The second part concerns how to process the sensed data and how to establish effective healing algorithms to generate appropriate feedback signals. For example, the work of Wu [6] adopted a 6-bit successive approximation register controller and a 6-bit digital-to-analog converter to generate signals that are fed back to the RF LNA.

The third part deals with the mechanisms through which the feedback signals can be applied to rectify the performance of LNAs. For example, the work of Howard [7] realized the control of feedback by enabling the feedback signals to tune the bias current of the RF LNA through the bias knobs. The fourth part of major differences lies in the parameters of LNAs that these works aim to remedy. For example, the work of Howard [7] attempts to heal the output third-order intercept point (OIP3), the gain, and the image rejection ratio (IRR) of the mixer. The work also attempts to heal the OIP3, the gain, and the noise figure (NF) of the LNA.

Different from the works concerning the allocations of self-healing adjustments on LNAs [4]–[7], the work of Nehring [8] focuses on implementing a miniature two-port VNA. The main idea of Nehring's work is to replace traditional VNAs, which are larger and more expensive, with the proposed miniature VNA. In other words, Nehring attempts to reduce the cost of VNAs using the proposed miniature VNA to measure the DUTs of interest. With such realization as the VNA front-end chip being mounted on boards, the miniature VNA being proposed by Nehring [8] measures DUTs externally. This is why the total physical dimensions of the miniature VNA of Nehring's work are as large as 100 mm $\times$ 90 mm $\times$ 38 mm. Given that the size of the whole measuring system is too large, the miniature VNA of Nehring's work cannot be used for embedded S-parameter measurement. The most critical drawback of measuring S-parameters externally is that performing correct de-embedding will be difficult especially when DUTs are buried deeply within ICs, as previously explained in Section I. The experiment of Nehring's work [8] demonstrates satisfactory results because the proposed miniature VNA avoids this most critical drawback by measuring biomedical materials rather than electronic devices. That is, the miniature VNA is used to measure the S-parameters of biomedical materials under test, which are composed of a mixture of ethanol and methanol. In contrast to the measurement of S-parameters of electronic DUTs, which may be located under layers of circuits, the measured object in Nehring's work has an excellent uniform distribution of density; thus, there is no need to perform de-embedding because the calibrated terminals of the miniature VNA directly touch the uniformly distributed materials under test. Compared with a traditional commercial VNA, the size of the miniature VNA proposed by Nehring is tiny. However, to embed a miniature VNA on chip, the circuit size needs further reductions. However, as the size reduces, the circuit complexity inevitably needs to be simplified as a compromise, while the simplification often causes the degradation of the measurement precision in return. That is, the effectiveness of functionalities and the size of a VNA are the two major trade-off design criteria that often fight each other.

Philippe [9] successfully demonstrated a design for on-chip S_{11} measurement at approximately 120 GHz using the design concepts derived from a six-port reflectometer, which is another famous methodology for measuring S-parameters. The theory of the six-port reflectometer is well-known for the simplicity in circuit design and the robust and complete calibration algorithms being proposed. Since nonideal terms inside circuits normally burgeon as operating frequencies increase, using the concepts of a six-port reflectometer to measure S_{11} values in the work of Philippe [9] is both reasonable and necessary because the solid calibration algorithms can effectively suppress nonideal terms and thus improve the measurement precision. However, the shortage of a six-port reflectometer is the complicated calibration procedures which often need the support of efficient external software to execute the sophisticated calculation. In particular, the work of Philippe [9] uses a transmission line that replaces the conventional phase shifter in a six-port reflectometer, together with 312 power detectors to measure the one-port complex reflection coefficient. Since a transmission line is a typical distributive device, it is suitable for applications at high frequencies, such as the frequency of 120 GHz in the work of Philippe [9]. For applications at lower frequencies, the large layout area of a transmission line almost makes the ideas of implementing transmission lines on chip impossible.

In conclusion, the works of Jayaraman [4], Goyal [5], Wu [6], and Howard [7] aim to apply self-healing adjustments on RF LNAs or mixers by utilizing some algorithms and S-parameter-like measurement results. However, these works can further be enhanced from three aspects. First, these works limit their DUTs of interest to only RF LNAs and mixers. The restriction on DUTs significantly reduces the applicable range because not only DUTs with different healing mechanisms will be excluded from candidacy, but a much greater portion of applications that do not require adjustments will also be disregarded. In other words, the works of Jayaraman [4], Goyal [5], Wu [6], and Howard [7] cannot serve as a universal solution that is applicable to a majority of DUTs. Second, because the works of Jayaraman [4], Goyal [5], Wu [6], and Howard [7] need to perform certain types of algorithms to generate feedback control signals, the possibilities to fully embed these works on chip will be limited. Even if the circuits responsible for executing the feedback algorithms are realized on chip, the resultant layout areas will be comparatively larger. Thus, these works will be less suitable for embedded measurements. Third, the works of Jayaraman [4], Goyal [5], Wu [6], and Howard [7] only measure and use partial information of the S-parameters, such as the amplitudes of the output of LNAs. However, S-parameters are more informative and can provide a much wider variety of applications. Obviously, using limited information also limits the applications of these prior works.

For the work of Nehring [8], the large dimension (i.e., 100 mm $\times$ 90 mm $\times$ 38 mm as reviewed earlier) limits the possibilities to place the proposed VNA close to a DUT of interest in the same chip. Instead, the miniature VNA measures the S-parameters of DUTs externally. As a result, the miniature VNA proposed by Nehring's work [8] does not enjoy the benefits of embedded S-parameter measurement. How to correctly de-embed the effects of other irrelevant

devices located on the paths of testing RF signals still remains unsolved. As today's semiconductor technologies continue to evolve, the tasks of performing proper de-embedding only become more and more challenging because new technologies are normally granted more numbers of layers, and 3D ICs are becoming increasingly popular nowadays. Therefore, as more and more devices standing on the paths in which testing RF signals are transmitted back and forth, or equivalently as the DUTs of interest are buried deeper and deeper inside ICs, the miniature VNA proposed by Nehring [8] faces the same problems as today's commercial VNAs do. The reason why Nehring's experiment shows no sign of having trouble performing de-embedding and hence demonstrates satisfactory measurement results is the miniature VNA is used to measure the S-parameters of biomedical materials under test, which, as a mixture of ethanol and methanol, enjoys the excellent advantage of being uniformly distributed.

The work of Philippe [9] aims to utilize the theory of a six-port reflectometer to measure the S_{11} of DUTs at approximately 120 GHz. However, the calibration algorithm of a six-port reflectometer is complicated. Although a set of robust and sophisticated calibration algorithms can help reduce unwanted interferences caused by nonideal terms burgeoning at such a high operating frequency as 120 GHz, it is less favorable for implementing embedded measurement of S-parameters. That is, if the circuits responsible for executing the calibration algorithm are to be implemented on chip rather than off chip, the resultant extra layout areas may make the overall measuring system too large to meet the requirements of embedded measurement. On the other hand, if the calibration algorithm is designed to be executed off chip, then the measuring system cannot be claimed as a fully embedded system capable of performing S-parameter measurement. Moreover, at lower operating frequencies, when the interferences caused by nonideal terms are not as severe, using the theory of a six-port reflectometer to perform S-parameter measurement may become overkill and may hinder the implementation of on-chip S-parameter measurement. In addition, the work of Philippe [9] adopts a distributive transmission line to replace traditional phase shifters inside six-port reflectometers. At lower operating frequencies, the design formulas of distributive passive components will demand much larger layout areas that can prohibit the work of Philippe [9] from being practicable for real-world applications. Based on the above discussions, the work of Philippe [9] is less possible for the implementation of on-chip S-parameter measurement at lower operating frequencies due to the complicated calibration algorithm and the too-large resultant chip area.

From the drawbacks analyzed above, several important challenges can be summarized as follows: First, aiming at specific DUTs will significantly limit the range of applications of the proposed measuring circuits or systems. Second, granting too complicated functionalities on the proposed measuring circuits or systems will considerably reduce the possibilities for embedded measurement. Third, measuring only partial information of S-parameters can be insufficient and may result in the failure of providing a general solution. Fourth, having too complex calibration algorithms can hinder the realization of on-chip S-parameter measurement. Fifth, using the theory of a six-port reflectometer may become overkill for applications at lower operating frequencies. Sixth, implementing the full functions of a commercial VNA into a measuring circuit or system is detrimental to the realization of that measuring circuit or system being embedded.

By understanding these related works, the values, the positioning, and the design spirit of our work can be further elucidated and appreciated. The pros and cons of our work can also be better clarified so that readers can be well aware of which option to choose when given a mission. In the following section, the idea and implementation of the S_{21} detection circuit will be disclosed.

III. THE IDEA AND IMPLEMENTATION OF THE S_{21} DETECTION CIRCUIT

This section first provides a background introduction on how traditional RF signal manipulation is achieved and why the proposed S_{21} detection network has a niche these days. Then, the detailed implementation, ranging from the circuit design to the system construction, is thoroughly elaborated, followed by the mathematical interpretation of the RF signal flows within the proposed S_{21} detection circuit.

A. THE IDEA

As introduced in Section I, S-parameters are measured by complex numbers, with which both the magnitude and the phase of the ratios of RF waves can be fully specified. Previous studies [10]–[13] show that various amplitude and phase detectors are able to directly measure high-frequency signals with out the need for down-conversion. Based on these mature measuring techniques in detector circuits, we assume that all the RF signals can be processed directly; that is, all the traditional down-conversion circuitries can be discarded. Although high-frequency amplitude and phase detectors are suitable for building embedded S-parameter measurement systems, these detectors are responsible solely for measuring the amplitudes and phases of RF signals. That is, the detectors are incapable of manipulating signals at the front end and are generally placed at the back end for measuring signals that have been well processed. Therefore, some networks are necessary to properly manipulate the flows of RF waves and to create suitable signals for the detectors to sense. In view of the great importance of these signal manipulation networks, we aim to explore and construct a network that can help implement the on-chip measurement of the S_{21} of a DUT. When the RF signal flows are properly guided by the proposed detection network, state-of-the-art amplitude and phase detectors can easily be appended at the back end. These cascaded high-frequency detectors can then correctly sense the preprocessed signals. Thus, the goal of embedded S_{21} measurements can be achieved without the help of extra down-conversion circuitry.

The indispensable RF signal manipulation networks are generally composed of two kinds of important passive circuits: the directional coupler (DRC) and the power divider (DIV). The DRC and DIV can normally be implemented in either a distributive form [14]–[16] or a lumped form [17]. Of these two forms, the distributive form requires a relatively large chip area compared with the lumped form. Hence the distributive-form-based circuits have rarely been implemented in ICs. However, to design well-qualified DRCs and DIVs in their lumped forms on ICs, the semiconductor fabrication technology needs to support quality passive electronic components, such as low-loss on-chip inductors and capacitors and precise on-chip resistors. Furthermore, at low operating frequencies, the areas of these passive inductors and capacitors can become extremely large. The extraordinary large areas of these devices are untoward factors that have stymied the realization of on-chip lumped passive electronic components in the past. Fortunately, thanks to the rapid advances in semiconductor fabrication technology and the quick boost in the operational frequency of RFICs nowadays, the two bottlenecks limiting the implementation of high-quality passive electronic components in their lumped form start to disappear. As a result, a growing number of essential passive components for RF measurement, such as DIVs, start to become realizable in their lumped form [17], which occupies much less area than their distributed counterpart.

B. CIRCUIT IMPLEMENTATION

Trying to exploit the emerging possibility of fabricating quality lumped passive electronic components on chip, the S_{21} detection network proposed in this paper was implemented in the 0.18-μm IBM 7RF semiconductor fabrication process. With a simple systemic circuit structure, the proposed network is composed mainly of three lumped DIVs [17], each of which is composed of four metal-insulator-metal (MIM) capacitors, two spiral inductors, and one K1 back-end-of-line (BEOL) resistor. A MIM capacitor is a capacitor that is made by a metal-insulator-metal configuration. A K1 BEOL resistor is a special resistor that is supported by the IBM 7RF technology. The advantages of K1 BEOL resistors include lower sheet resistance, higher resistance precision, and lower parasitic capacitance. The schematic of the DIV together with the inner components' design values are shown in Fig. 2. The resistor R1 in Fig. 2 aims to make the S_{11} of the DIV even smaller around the targeted operating frequencies so that the reflected RF signals at Port 1 of the DIV can be reduced further. That is, the purpose of inserting the resistor R1 is to adjust the overall performance of the DIV. If such resistor causes tremendous signal losses as measuring frequencies continue to rise in the future, reducing or even removing the resistance of R1 is appropriate and reasonable.

As a three-port network, the DIV mainly serves as a signal splitter. Fig. 3 demonstrates a simplified signal flow graph of the DIV when its three ports are assumed to be correctly terminated. If an RF signal is fed into Port 1, ideally the

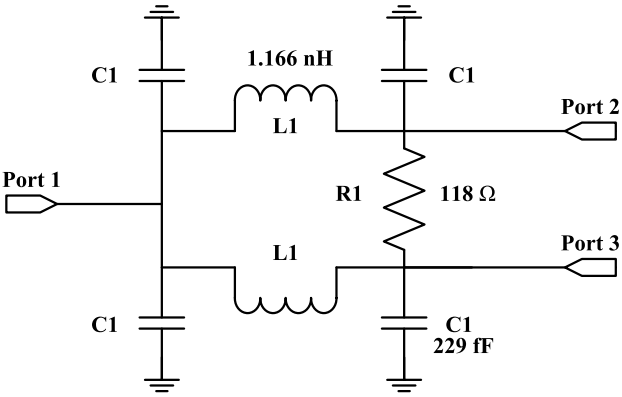

FIGURE 2. Schematic and design values of the lumped passive integrated DIV.

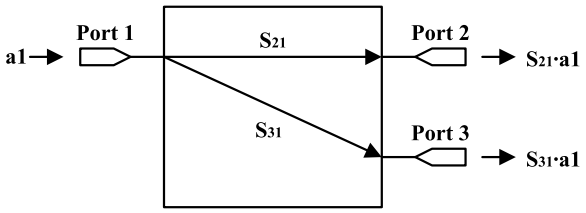

FIGURE 3. Simplified signal flow graph of the lumped passive integrated DIV.

DIV will divide the input signal into two signals, which are of equal magnitude and identical phase, and output these signals to its Port 2 and Port 3, i.e., $S_{21} = S_{31}$. With this unique capability of signal splitting, generating two or even more identical RF signals from one given RF signal source becomes feasible.

This signal duplication technique is essential in S-parameter measurement systems because it is feasible to let one signal remain unchanged while the other identical signal travels a path whose characteristics are of interest. With such an arrangement, it is possible to compare the modified signal, which is intentionally made to travel a certain path of interest, with the reference signal, which is kept unchanged. Through this kind of comparison, the differences between the two originally identical signals can be obtained, and the characteristics of the additional path can easily be measured. For this reason, the DIV is a very useful and indispensable component in the construction of embedded S-parameter measurement systems.

The system-level architecture of the proposed on-chip S_{21} detection network is demonstrated in Fig. 4, in which Port 1 serves as the input port and Ports 2 and 3 serve as the output ports. DUT2 stands for the specific device under test that is used in this research. The S_{21} of the DUT2, denoted as S_{21_DUT2}, is the target that the proposed S_{21} detection network tries to measure. There are three reasons for naming this device under test DUT2 rather than DUT1 or DUT. The first reason is that DUT is a general term for "device under test" and hence inappropriate to be used to indicate a specific circuit. The second reason is that DUT2 can convey the main

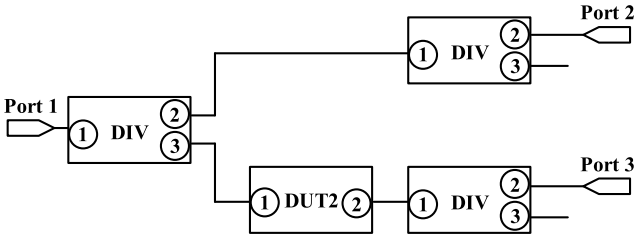

FIGURE 4. System-level architecture of the proposed on-chip S_{21} detection network.

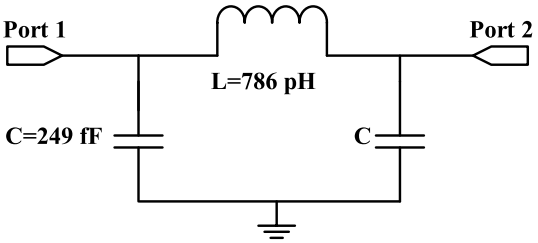

FIGURE 5. Schematic and design values of the DUT2 in Fig. 4.

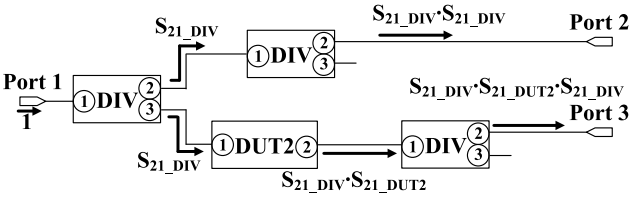

FIGURE 6. Demonstration of signal propagation within the embedded S_{21} detection network.

purpose of this paper, which is to measure the S_{21} of a DUT. The third reason is that we hope to reserve the term DUT1 for a publication about an embedded S-parameter measurement circuit that is dedicated to measuring the S_{11} of a DUT.

In addition to the specification of the system-level ports, port definitions inside each component are also given by the circled numbers. The original system can even be simplified further by keeping only the DIV on the left and the DUT2. The two DIVs on the right serve as connectors while the amplitude and phase detectors are integrated. For the time being, Port 3 of these two additional DIVs on the right is left open-connected in order to observe how well the system deals with unmatched conditions. In the future, when Port 3 of the two DIVs on the right is connected to the amplitude or phase detectors, the impedance looking out of Port 3 of the two DIVs on the right will be better matched. The performance of the detection circuit will hence bear a much closer resemblance to the simulation.

Constructed with such a simple structure containing only three DIVs and one DUT2, as shown in Fig. 4, the proposed system can report the S_{21} of DUT2 (S_{21_DUT2}). More specifically, the S_{21_DUT2} can be derived by dividing the S_{31} of the system (S_{31_SYS}), i.e., the signal transmitted from Port 1 to Port 3, by the S_{21} of the system (S_{21_SYS}), i.e., the signal transmitted from Port 1 to Port 2. In Subsection V.C, it will be demonstrated how extra back-end detectors can be added to measure S_{31_SYS} and S_{21_SYS}.

Fig. 5 demonstrates the schematic and design values of the DUT2 that was used in the circuit of Fig. 4. Since the DUT2 is merely used to help verify the functionality of the proposed S_{21} detection network, the only requirement for the DUT2 is that its S_{21} value can vary with frequency within the frequency range of interest. In view of this, the design of the DUT2 is rather simple—it comprises only two MIM capacitors and one spiral inductor.

C. MATHEMATICAL DEDUCTION OF THE S_{21} DETECTION SYSTEM

By defining the ratio of S_{31_SYS} over S_{21_SYS} as $S_{21\_DUT2\_DERIVED}$, (1) shows that this newly defined term equals the S_{21} of the DUT2 in the system arrangement in Fig. 4.

$$\frac{S_{31\_SYS}}{S_{21\_SYS}} \equiv S_{21\_DUT2\_DERIVED} = S_{21\_DUT2} \tag{1}$$

In fact, the accuracy of the equality between $S_{21\_DUT2\_DERIVED}$ and S_{21_DUT2} in (1) is based on the degree to which the following two assumptions are satisfied. First, the coupling between Port 2 and Port 3 of a DIV is assumed to be negligible. As the first-order nonideality contributor of the whole system, the coupling between Ports 2 and 3 of a DIV is regarded as the strongest interference source. Hence, in order not to spoil the signal integrity between the two deliberately separated branches, the coupling between Ports 2 and 3 of a DIV should be reduced to a negligible level. That is, S_{23_DIV} and S_{32_DIV} should be negligible. Secondly, the high-order reflections within the network also need to be assumed negligible. Because the validity of the proposed system is based on the first-order approximation of the flow graph of the RF signal, undesired high-order reflections may also damage the performance of the proposed system. Therefore, to guarantee the efficacy of (1), we need to assume that the proposed S_{21} detection circuit is so well-developed that the high-order reflections are dissipated to a negligible level.

Based on these two assumptions, the measurement concept expressed by (1) can easily be understood by using Fig. 6. In this figure, with a unity RF signal input into Port 1 on the left, the propagation condition of this unity input within the proposed detection network can clearly be demonstrated. For the sake of simplicity, every S-parameter is appended with an additional attribute indicating the sub-circuit it belongs to. Based on the above two assumptions and Fig. 6, it is easy to see that the signals reaching Ports 3 and 2 of the whole detection network differ only by the response of S_{21_DUT2}.

IV. RESULTS

In this section, the results of the simulation, chip fabrication, and measurement are demonstrated. Based on the crucial experimental results described in this section, some important discussions and analyses are further detailed in the next section.

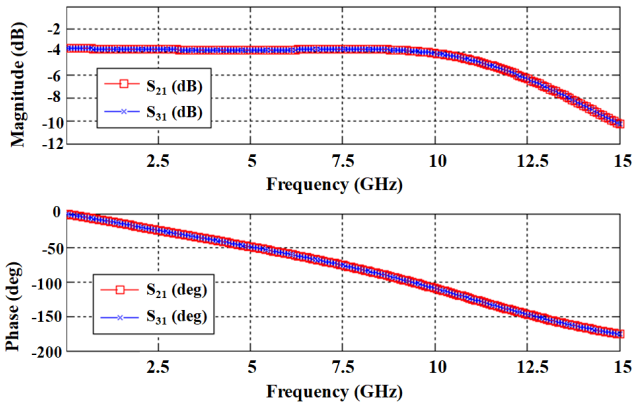

FIGURE 7. Simulation results of the S_{21} and S_{31} of the lumped passive integrated DIV.

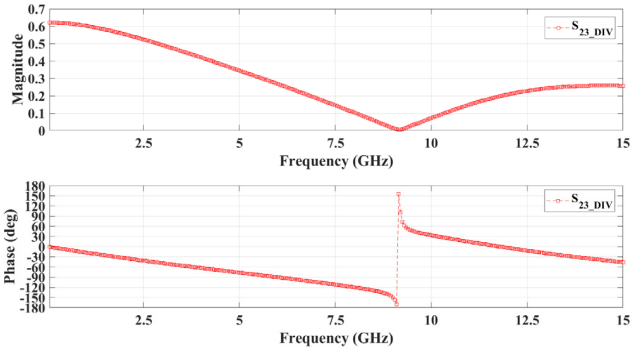

FIGURE 8. Simulation results of the S_{23} of the lumped passive integrated DIV.

A. SIMULATION RESULTS

Fig. 7 demonstrates the simulation results of the lumped passive integrated DIV presented in Fig. 2. Based on the simulation results, the S-parameters S_{21} and S_{31} are equal to each other in both magnitude and phase for the whole frequency range.

Fig. 8 demonstrates the isolation condition between Ports 2 and 3 of the DIV in Fig. 2. As explained in Subsection III.C, these two ports of the lumped DIV in Fig. 2 need to have a high level of isolation so that the term of the first-order non-idealities generated in the S_{21} detection circuit in Fig. 4 can be suppressed to the utmost, making the performance of the S_{21} detection circuit approach (1). As demonstrated in Fig. 8, the magnitude of the S_{23} of the DIV (i.e., S_{23_DIV}) is a small value around 10 GHz. In addition, because the DIV in Fig. 2 has a symmetrical circuit structure, the S_{32} should be similar to the S_{23} of the DIV.

By minimizing the magnitude of S_{23_DIV}, the coupling effects between Ports 2 and 3 of the DIVs in Fig. 4 can also be minimized. Of the three DIVs in Fig. 4, the most important isolation is the one between Ports 2 and 3 of the leftmost DIV. The reason is the leftmost DIV is responsible for dividing the input RF signals into two intentionally separated branches, one of which serves as the reference while the other tests the DUT of interest. Therefore, if the magnitude of S_{23_DIV} of the leftmost DIV in Fig. 4 is not sufficiently small, the two intentionally separated branches of RF signals will end up mixing with each other. The mixture of these two ideally independent branches of signals will make the measurement results chaotic because the reference signal is intertwined with the testing signal. What is even worst is that because these two branches of signals contain the ingredients that experience the least attenuation, the mixture of them can cause the most severe impacts on the measurement results. The isolation between Ports 2 and 3 of the DIVs, as demonstrated in Fig. 8, is good enough for situations in which measurement precision is not strictly demanded. For applications that require high measurement accuracy, the corresponding calibration procedures will be investigated in the future.

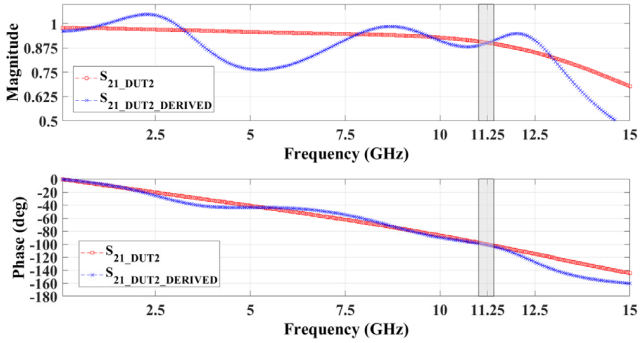

FIGURE 9. Simulation results of the S_{21} detection network.

Fig. 9 demonstrates the simulation results of the proposed S_{21} detection network. The S_{21_DUT2} curve is derived by directly simulating the DUT2 demonstrated in Fig. 5 using the CAD tool Cadence. The $S_{21\_DUT2\_DERIVED}$ curve, on the other hand, is derived by simulating the whole system demonstrated in Fig. 4 by using Cadence and then applying the resultant S_{21_SYS} and S_{31_SYS} to (1) using the CAD tool MATLAB. To clearly differentiate these two types of simulations, we call the first one "device simulation" and the second one "system simulation." As can be seen in Fig. 9, the S_{21_DUT2} curve and the $S_{21\_DUT2\_DERIVED}$ curve have an intersection at around 11.25 GHz. This result shows that (1) holds around the targeted design frequency when the two assumptions given in Subsection III.C are satisfied. Upon leaving the design frequency, unwanted signal coupling starts to burgeon, invalidating (1) and making the $S_{21\_DUT2\_DERIVED}$ system simulation deviate from the S_{21_DUT2} device simulation.

To evaluate the proximity level between a derived S-parameter and the ideal value, we propose the following equation:

$$error \equiv \frac{|S_A - S_B|}{|S_B|} \tag{2}$$

where S_A represents the derived value of an S-parameter and S_B stands for the ideal value of that S-parameter.

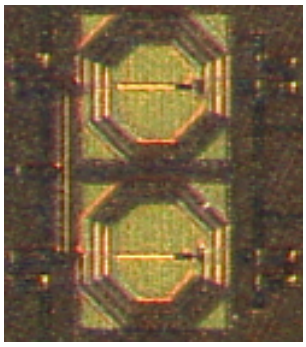

FIGURE 10. Die micrograph (270μm × 300μm) of the lumped passive integrated DIV.

In (2), if S_A is $S_{21\_DUT2\_DERIVED}$ and S_B is S_{21_DUT2}, then the error between these two data sets can easily be derived. If we define the bandwidth as the frequency range in which the error between $S_{21\_DUT2\_DERIVED}$ and S_{21_DUT2} is less than 3%, the proposed S_{21} detection network has a bandwidth of about 400 MHz, which ranges from 11.017 GHz to 11.417 GHz. The range of the 3%-error bandwidth is shaded in Fig. 9.

B. FABRICATION RESULTS

The detection network in Fig. 4 was fabricated by using the 0.18-μm IBM 7RF semiconductor fabrication technology. Due to the limited design wafer area, no stand-alone lumped passive integrated DIV was fabricated. As a result, there is no direct measurement result of the stand-alone DIV. Instead, DIVs are fabricated as part of the embedded S_{21} detection network so that more powerful and valuable functions of the systemic S_{21} detection network can be demonstrated and verified. Fig. 10 demonstrates the die micrograph of the fabricated DIV. In Fig. 10, the two spiral inductors, whose area dominates the overall layout of the DIV, are placed symmetrically to each other. Likewise, the four MIM capacitors are placed symmetrically at the four corners. The passive electronic components are deliberately placed by taking advantage of this symmetrical layout because better manufacturing quality can be assured and the DIV's performance can be kept close to the design specifications. The die of the DIV has a width of about 270 μm and a length of about 300 μm.

Constructed by the three lumped passive integrated DIVs and one DUT proposed in Subsection III.B, the embedded S_{21} detection network has a chip size of 1.2 mm x 1.2 mm. Together with the pads that surround and protect the core, the micrograph of the fabricated die of the S_{21} detection network is demonstrated in Fig. 11. Port 1, serving as the input port in Fig. 4, corresponds to the three middle pads on the left-hand side in Fig. 11. These three pads form the famous G-S-G configuration, where G symbolizes the ground and S represents the signal. Ports 2 and 3, serving as the output ports in Fig. 4, correspond to the two upper pads and the two lower pads in Fig. 11, respectively. Both Port 2 and Port 3 use the G-S configuration, in view of the geometric

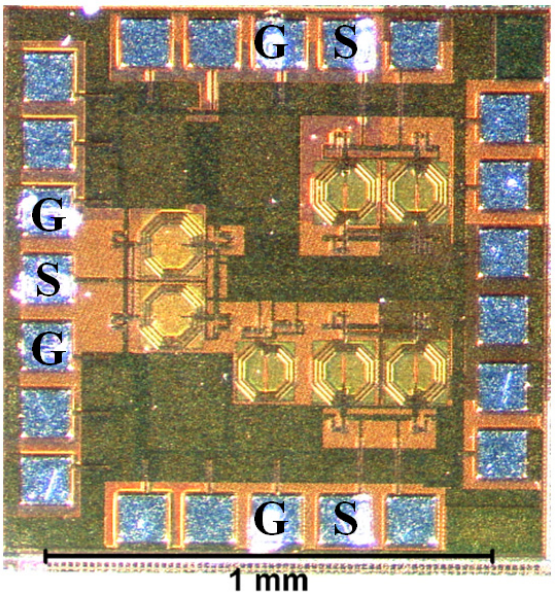

FIGURE 11. Die micrograph (1.2 mm × 1.2 mm) of the S_{21} detection network.

relationships between the DIVs and the pads, the arm location of the RF probe station, and the available types of RF probes. Although the G-S configuration is not as balanced as the G-S-G configuration, such a configuration is good enough to measure the test chips at around 10 GHz according to the probe specifications.

C. MEASUREMENTS RESULTS

As explained in Subsection IV.B, since no stand-alone lumped passive integrated DIV was fabricated, there is no direct measurement result of the stand-alone DIV. Instead, the whole chip of the S_{21} detection network in Fig. 11 is measured on an RF probe station with the help of the RF cables, the RF probes, and the VNA. The RF probe station serves as a base where the chip is secured and placed upon. The RF signal cables are used as the connection between the VNA and the RF probes. The RF probes are used to be landed on the pads of the G-S-G and G-S configurations so that the RF signals can be transmitted from and received by the VNA. The VNA is responsible for deriving the S_{21} and S_{31} of the chip in Fig. 11 and plays a main role in the overall measurement configuration.

This paper adopts a VNA as a core instrument to measure the test chips and, thus, to verify whether the performance of these chips is consistent with the design concepts. Indeed, the verification of the performance of the test chips is exactly equivalent to the verification of the efficacy of (1). Using a VNA to test the fabricated chips is logical because this paper focuses mainly on how to correctly manipulate input testing RF signals at the front end so that the measured S_{21} of DUT can be further detected by other sensors (e.g., amplitude and phase detectors) in the future. In other words, the information of the measured S_{21} of the DUT of interest is kept inside

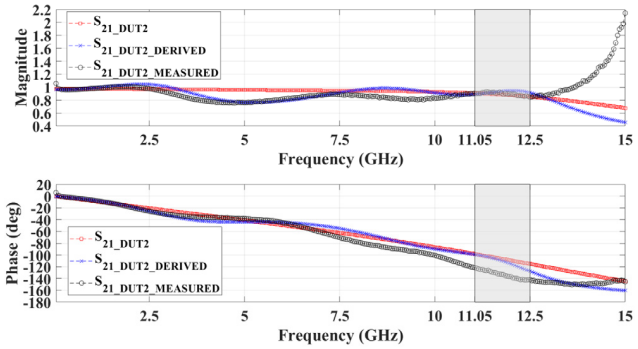

FIGURE 12. Measurement results of the S_{21} detection network.

the RF signals reaching Ports 2 and 3 in Fig. 4, whose die micrograph is in Fig. 11. Following the famous divide-and-conquer strategy, the validity of the S_{21} detection circuit is verified at first. Once the design concepts are successfully substantiated, other peripheral circuits, such as RF switches, detectors, and signal generators, can further be integrated into the proposed S_{21} detection circuit. In the meantime, when there are no amplitude and phase detectors in the test chips, adopting a VNA to verify the design concepts being expressed by (1) is convenient and appropriate because for a VNA working on the three-port mode, the measured S_{21} and S_{31} represent the S_{21_SYS} and S_{31_SYS} in (1), respectively.

Fig. 12 demonstrates the measurement results of the S_{21} detection network. The S_{21_DUT2} curve and the $S_{21\_DUT2\_DERIVED}$ curve are the same as the ones demonstrated in Fig. 9. The $S_{21\_DUT2\_MEASURED}$ curve represents the derived S_{21_DUT2} from the measurement results. That is, $S_{21\_DUT2\_MEASURED}$ is the result of measuring the S_{21} and S_{31} of the chip in Fig. 11 and applying the values measured to (1).

From Fig. 12, it can be seen that the $S_{21\_DUT2\_MEASURED}$ curve derived from system measurement and the $S_{21\_DUT2\_DERIVED}$ curve derived from system simulation have similar trends. That is, both $S_{21\_DUT2\_MEASURED}$ and $S_{21\_DUT2\_DERIVED}$ tend to approach the S_{21_DUT2} curve derived from device simulation around the frequency region shaded in Fig. 12. By taking a closer look, however, we can observe some differences in how these two curves intersect with the S_{21_DUT2} curve. The original $S_{21\_DUT2\_DERIVED}$ curve derived from the system simulation results intersects with the curve S_{21_DUT2} curve from device simulation at about 11.25 GHz for both the magnitude and phase parts. The measured $S_{21\_DUT2\_MEASURED}$ curve, on the other hand, intersects with the S_{21_DUT2} curve from device simulation at about 11.05 GHz only for the magnitude part. At 11.05 GHz, the phase of the measured $S_{21\_DUT2\_MEASURED}$ curve lags the phase of the simulated S_{21_DUT2} curve by about 20 degrees. That is, compared to the 11.25 GHz in Fig. 9, the intersection frequency is shifted to 11.05 GHz, with an additional error of 20 degrees in phase for the measurement results.

Compared to the phase error, which is around 20%, the error of the intersection frequency is only about 1.8%. Hence, the shift of the intersection frequency, from 11.25 GHz in the earlier simulation results to 11.05 GHz in the measurement of test chips, can be regarded as a minor change, which may be caused by tiny process variations or some real-world parasitic elements. For the phase error part, the variation seems to be much larger and is more problematic. Such a huge variation may be due to process variations in the path from Port 1 to Port 3 in Fig. 4.

Despite these phase errors, the design concept has been successfully demonstrated and verified based on the clearly similar trends between the simulation and the measurement results. In addition, it should be kept in mind that in Fig. 4 Port 3 of the two DIVs on the right is deliberately left open-circuited in order to observe how the system deals with the most unmatched condition. Therefore, we are highly confident that the discrepancy between the measurement and simulation results can be much less when the two DIVs are properly terminated with 50Ω resistance in real applications. Since this paper focuses mainly on the verification of the design concepts, more advanced analyses and reasoning about the potential causes of the errors will be studied in the future. To remove the effect of process variations on the detection network, corresponding calibration procedures will be given in a future publication.

V. DISCUSSION

The measurement results in Subsection IV.C are presented at a relatively large scale, which allows an overall observation of the performance of the fabricated S_{21} detection network. As demonstrated previously, the measurement and simulation data have similar performance trends in the frequency region of interest. Following this observation, Subsection V.A zooms in and discusses in detail the network's measurement results near the operating frequencies of interest. To further emphasize the aim of this work, Subsection V.B compares the goals of this work and that of related works in the past [4]–[9]. In addition, to make the proposed S_{21} detection network a more integrated measuring circuit, Subsection V.C demonstrates how amplitude and phase detectors can be cascaded in the future. Finally, to provide clear instructions, Subsection V.D explores the possible upper and lower bounds of the applicable operating frequencies of the detection circuit. Despite the potential limitations, Subsection V.D also proposes strategies that can help stretch the usable frequency range and examines the situations in which the proposed S_{21} detection network can play a vital role.

A. FURTHER ANALYSES OF THE MEASUREMENT RESULTS

To better observe the S_{21} values around the frequencies shaded in Fig. 12, Fig. 13 shows a partially enlarged portion of Fig. 12. As can be seen, for the magnitude part, the $S_{21\_DUT2\_MEASURED}$ curve intersects with the S_{21_DUT2} curve at a lower frequency than the $S_{21\_DUT2\_DERIVED}$ curve. Assuming that the fabricated DUT2 suffers no process variation (i.e., the device simulation of S_{21_DUT2} remains the same), Fig. 14 shows the error percentage of the

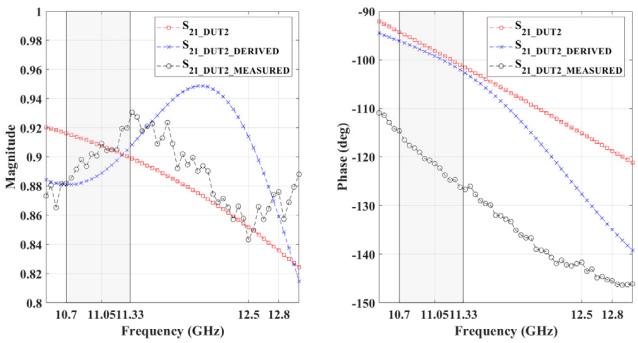

FIGURE 13. The partially enlarged picture of Fig. 12.

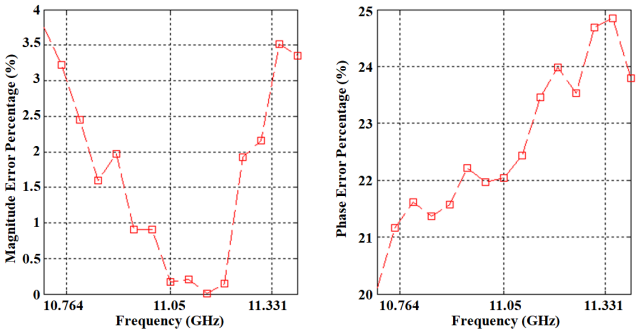

FIGURE 14. Error percentage of the measurement results of the S_{21} detection network.

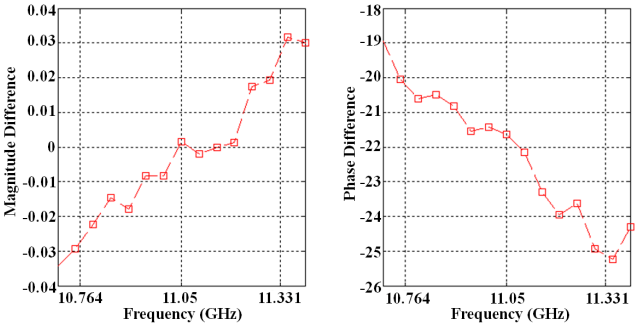

FIGURE 15. Difference in the measurement results of the S_{21} detection network.

magnitude and the phase of S_{21} between the S_{21_DUT2} and $S_{21\_DUT2\_MEASURED}$ curves for the shaded region in Fig. 13. As stated in Subsection IV.C, the error percentage in phase is relatively large and may be caused by process variation. When we neglect the phase error, the 3%-error bandwidth from the measurement can be found to be 560 MHz, ranging from 10.764 GHz to 11.331 GHz. On the other hand, the phase error percentage ranges approximately from 21% to 25% within such a frequency window.

Fig. 15 shows the difference in magnitude and phase between the S_{21_DUT2} and $S_{21\_DUT2\_MEASURED}$ curves for the shaded region in Fig. 13. As can be seen, from 10.764 GHz to 11.331 GHz, the difference in magnitude ranges approximately from -0.03 to 0.03, and the difference in phase ranges approximately from -20 to -25 degrees.

TABLE 1. Average and variance of the measurement results of the S_{21} detection network.

Frequency (GHz)	Magnitude of S_{21}		Phase of S_{21} (degree)	
	Mean	Variance	Mean	Variance
10.95	0.907098	0.000016	-119.539802	0.594578
11	0.905301	0.000013	-120.025170	0.628977
11.05	0.913969	0.000015	-120.783017	0.655643
11.1	0.909182	0.000016	-121.873961	0.674856
11.15	0.909432	0.000014	-123.613379	0.736022

For the measurement of the S_{21} detection network, there are a total of 8 chips being measured, and all the measurement results are consistent with each other. This high consistency guarantees the quality of the measurement execution. To show the excellent congruency among the measured data of the 8 chips, the average and variance of the magnitude and phase of the derived S_{21} (i.e., $S_{21\_DUT2\_MEASURED}$) at around the targeted frequencies are given in Table 1, in which the variance values can be seen to be very small.

Due to the limited silicon area available, there was not enough room to tape out some stand-alone lumped DIVs and the DUTs. Instead, the whole detection network, including the three DIVs and one DUT, was manufactured on one chip, a complex that makes the exact verification of the measurement results extremely difficult. Fortunately, thanks to the simple circuit structure, the careful arrangement of the layout, and the high reliability of the IBM process, the design concept can be shown to be successful since high similarity exists between the simulation and the measurement data.

However, minor discrepancies can also be found between the simulation results and the measured ones. The simulation results demonstrate that the system-simulation $S_{21\_DUT2\_DERIVED}$ intersects with the device-simulation S_{21_DUT2} at about 11.25 GHz, but the measurement results show that the system-measurement $S_{21\_DUT2\_MEASURED}$, obtained by measuring the S_{21} and S_{31} of the chip and applying them to (1), intersects with the device-simulation S_{21_DUT2} at about 11.05 GHz, with an additional error of approximately 20 degrees in phase. That is, there is a downward shift in the operating frequency combined with a phase error in the real implementation. Of the two main changes in real chip implementation, the extent of the frequency shift, which is only about 1.8%, is relatively minor and can be regarded as a result of reasonable process variations. On the other hand, the phase error is more severe and hard to ignore. Although the direct cause of such a deviation is difficult to prove in a rigorous way, it is still possible to retrieve useful information from the measured data and identify some potential culprits for such variations. Based on more detailed analyses of the simulation and measurement data, it is found that the process variations on the path from Port 1 to Port 3 in Fig. 4 is the main possible factor contributing to the phase error in a real implementation. Because of space limitations, the detailed analyses will be given in a future publication. Besides, the intentional arrangement of the open-connected

configuration on Port 3 of the two DIVs in Fig. 4 can aggravate the measurement error by creating first-order couplings. In the future, when every boundary port of all circuit elements is properly terminated with 50Ω resistance, the measurement error will surely be lower. Finally, as for the bandwidth, if we assume that the phase error can be fixed in a future run and be temporarily neglected, the 3%-error bandwidth is found to be about 560 MHz in measurement, slightly wider than the 400 MHz in simulation.

The size of the integrated lumped DIV is about 270 μm × 300 μm, with most of the area occupied by the DIV's two inner inductors. The chip size of the S_{21} detection network is approximately 1.2 mm × 1.2 mm, including the pad area, which will be removed when the detection network is used as embedded circuitry in real applications. Like the DIV, the overall chip area of the S_{21} detection network is dominated by its inner inductors. Fortunately, because the pads will be removed in real applications, the chip area of the S_{21} detection network can become a lot smaller. Furthermore, according to the design theorem of the lumped DIVs, the inductance of inside inductors will continue to reduce as the operating frequencies go higher. This reduction in DIV area means a reduction in chip area of the whole S_{21} detection network. That is, as far as the area overhead is concerned, the S_{21} detection network will become more favorable at higher operating frequencies. Because the proposed detection network, except for the DUT2, contains only three simple lumped integrated passive devices, it does not have the issue of technology dependence. That is, the proposed S_{21} detection network can be a general solution for most semiconductor technologies available today regardless of whether a CMOS, a bipolar, or a BICMOS technology is used.

For ad hoc production tests, it may be feasible not to perform calibration if the expected readings, corresponding to the correct DUT performance, of the S_{21} measurement results are already known. Nevertheless, performing calibration has become a convention in RF measurement and provides many advantages, such as extending the working bandwidth and identifying the erroneous part of the circuit. Hence, we recommend, if possible, to perform calibration procedures on the S_{21} detection network before using it. To accommodate such needs, the calibration procedures for the S_{21} detection network will be given in a future publication.

B. COMPARISONS WITH PAST WORKS

Different from the related works [4]–[9] reviewed in Section II, the goal of our research is to develop a useful core of circuit that can serve as a general solution capable of performing on-chip S_{21} measurement on any DUT. In view of those valuable advantages introduced in Section I, our work especially emphasizes the importance of having embedded measurement for S-parameters. Based on the authors' limited knowledge, very few or no prior works truly focus their attention on the realization on embedded S-parameter measurement.

To perform embedded S_{21} measurement on arbitrary DUTs, we need to make our design as simple and as generally adaptable as possible. This means that the measuring circuit should be small and the measuring purpose should be simple and universal. In other words, the sophisticated and DUT-dependent self-healing techniques [4]–[7] are not adopted in the S_{21} detection circuit presented by this paper. Neither is the stand-alone miniature VNA [8] an option because its dimensions are just too large to be of embedded use. Last, because we set our targeted operating frequency to be around 10 GHz, which is much lower than 120 GHz, it becomes feasible not to use such rigorous theory as the six-port reflectometer [9]. With such relief, the measuring circuit (i.e., the proposed S_{21} detection circuit) can be made much smaller using alternative measuring methodology. Although circuit simplification is a must for embedded purposes, keeping a satisfactory level of measurement precision is also essential. These two contradictory factors constantly fight each other, and we need to find an excellent balance in between. In addition, because the interferences caused by nonideal terms become much less at low frequencies, it may be acceptable not to perform calibration procedures if the process variations are small enough, the requirement for measurement precision is not asked high, or the applications are for ad hoc production tests. For situations in which performing calibration becomes necessary, corresponding calibration methods for the proposed S_{21} detection circuit will be investigated in a future publication.

Setting the targeted operating frequency as low as approximately 10 GHz is helpful to the expansion of the allowable types of measuring circuits. Lower operating frequencies also help alleviate the level of robustness that is required of calibration algorithms. However, at such low frequencies, passive components in their distributive forms become alarmingly large and hence inappropriate to be fabricated in ICs. For example, the transmission line used in the work of Philippe [9] will become too large to be of any use. Fortunately, passive components in their lumped forms come into play by providing such advantages as being simple in circuit structure, reasonable in layout size, and generally applicable within the frequency range of interest. According to the formulas of the lumped passive devices, the recommended usable frequency bandwidth is between 5 GHz and 30 GHz, in which lumped passive devices can have both reasonable layout sizes and acceptable level of precision. The detailed analyses will be further demonstrated in Subsection V.D. With deliberate arrangements of these simple passive components, the success of measuring S_{21} of DUT is well demonstrated. By utilizing only simple passive components and by placing them in the right places, we broaden the applicable range of the S_{21} detection circuit proposed by this paper. That is, by making the S_{21} detection circuit small and by focusing the measuring purpose on the on-chip S_{21} measurement, our work is not only applicable to RF LNAs or mixers but is also applicable to almost all types of DUTs.

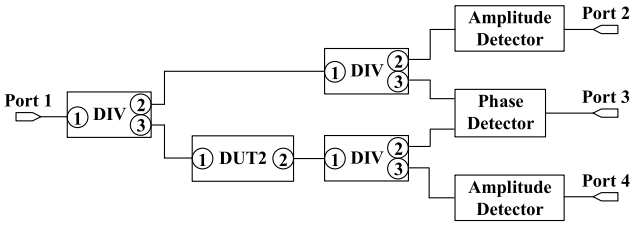

FIGURE 16. Possible integration of cascaded detectors for the S_{21} detection network.

To sum up, being able to properly manipulate the front-end RF signals makes the proposed S_{21} detection circuit useful for on-chip S_{21} measurement mainly for the following three reasons. First, having reasonable layout sizes makes the S_{21} detection circuit appropriate for embedded S_{21} measurement. Second, having simple circuit structures makes the S_{21} detection circuit technology-independent, which means that the proposed S_{21} detection circuit is applicable to a wide variety of semiconductor processes. Third, the solidity of circuit design makes the S_{21} detection circuit readily extensible in the future. That is, with the information of the measured S_{21} soundly retrieved, additional functionalities can be easily added to the S_{21} detection circuit in the future by integrating more useful peripherals, such as amplitude detectors, phase detectors, RF switches, or signal generators.

C. INTEGRATION OF DETECTORS IN THE FUTURE

This paper only designs and verifies the proposed S_{21} detection network. However, in the future, additional high-speed amplitude and phase detectors can be cascaded to Port 2 and Port 3 of the S_{21} detection network such that the processed RF signals can be sensed directly. In this way, the S_{21} detection network will become completer and more integrated.

More specifically, to make the S_{21} detection network more integrated as a measuring system, it is possible and beneficial to add some phase and amplitude detectors to the right part of the detection network in Fig. 4. Fig. 16 shows a possible arrangement of how the detectors can be integrated into the proposed system in the future. With the configuration shown in Fig. 16, the magnitude of the S_{21} of the DUT can be known by dividing the output from Port 4 by that from Port 2, and the phase can be derived by the output from Port 3.

D. POSSIBLE LIMITATIONS AND NICHES

Fundamentally speaking, the success of the proposed S_{21} detection network depends a great deal on the favorable characteristics of its inner integrated DIVs. The lumped integrated DIV plays a significant role in two main aspects. First, the detection network possesses a quite simple structure circuit-wise. Second, it is capable of generating two identical output RF signals from a single RF input source, a signal replication functionality that is essential in almost every RF measurement system. Although the simple structure of the proposed network has the advantage of being technology independent, a property guaranteeing its easy usage in the future, some innate limitations on its applicable operating frequency do exist due to the original derivation equations of the passive lumped components inside the DIV.

TABLE 2. Design values of lumped inductance and capacitance, and their estimated areas when DIV operates at different frequencies.

Frequency (GHz)	L1 (pH)	Area of L1 (μm^2)	C1 (fF)	Area of C1 (μm^2)
1	11254	$3.44x10^4$	2250	$1.21x10^4$
5	2250	$1.68x10^4$	450	$2.41x10^3$
10	1166	$1.43x10^4$	229	$1.23x10^3$
15	750	$1.33x10^4$	150	804
20	563	$1.28x10^4$	112	600
25	450	$1.25x10^4$	90	482
30	375	$1.24x10^4$	75	402
35	321	$1.22x10^4$	64	341

According to the formulas of the lump devices' design values in the literature [17], the values of L1 and C1 in Fig. 2 at different operating frequencies are shown in Table 2. In addition, based on the IBM 7RF technology, several estimated on-chip areas corresponding to different inductance and capacitance values are also given in Table 2. Since the IBM 7RF is used as the reference to derive the estimated components' areas, the areas in Table 2 are listed just to provide a general idea about how large a passive lumped component may be. That is, the real areas will vary with the fabrication process as well as the layout styles that a designer chooses to use. Because the capacitors in Fig. 2 are simple MIM capacitors, the estimation of capacitor areas in Table 2 is done based on the simple assumption that capacitance is proportional to area. Compared to the area estimation of capacitors, the area estimation of inductors is much more complicated because the inductance of a spiral inductor is not linearly proportional to its area. Besides, different technologies normally have different definitions of the minimal inner diameter (or equivalently, the minimal area) that a spiral inductor can have. Hence, the area of a spiral inductor is rather technology dependent, and the information in Table 2 provides just a glimpse of the possible scale. By considering the rule for the minimal inner diameter of IBM 7RF, the inductor areas in Table 2 are estimated by using the Mohan's formula [18], which utilizes a data-fitted monomial expression to estimate inductance values. Because the Mohan's formula uses five variables to estimate a planar spiral inductor's inductance, the areas in Table 2 are derived by arbitrarily using a reasonable set of variable values. Being one among several other possible arrangements, the estimated inductor areas in Table 2 provide only indicative values even for the IBM 7RF technology. As shown in Table 2, we estimate that the presented S_{21} detection network can best operate from 5 to 30 GHz.

Below 5 GHz, the capacitance and inductance values in Fig. 2 will become alarmingly high, making the sizes of the DIV grow unreasonably large. In this way, the size of the whole S_{21} detection network will increase simultaneously, and the resultant huge area overhead will deprive

the detection network of its original appeal. However, this is only a rule of thumb that we recommend. Whether the area overhead is reasonable or acceptable depends on budgets, purposes, the size of the DUTs, the degree of reusability, and so on, of each design. In other words, the lower bound of the operating frequency, which we suggest to be 5 GHz, is rather flexible since the only concern is not whether the circuit is realizable but whether it is worthwhile to be realized.

The upper bound of the operating frequency is advised to be 30 GHz. This frequency is just an approximate estimation based on the process used in this research. Hence, the following statements are qualitative instead of quantitative. The reader is strongly advised to judge the upper bound of the operating frequency of a circuit by the semiconductor process and layouts used. Above the estimated 30 GHz, the lumped DIV will require much smaller capacitance and inductance values, which will cause a much larger fabrication error percentage because of the smaller denominator. Moreover, at such high operating frequencies, tiny segments of wires can potentially create considerable unwanted inductance, and nearby metals or even semiconductor layers may accidentally form significant parasitic capacitance, both of which can seriously jeopardize the performance of the DIV. What is more, the severity of the effects can escalate with increasing operating frequencies and decreasing inductance and capacitance values. Besides, even if the fabrication precision keeps improving as technologies progress, the S_{21} detection network will still face challenges posed by parasitic capacitors and inductors as the operating frequency increases. Hence, it can be concluded that the upper bound of the operating frequency is objective and rigid, whereas the lower bound of the operating frequency is subjective, flexible, and limited mainly by the chip area costs. That is, it is harder to push the upper bound higher than to push the lower bound lower even if the manufacturing precision of the semiconductor technologies continues to increase and the circuit designers continue to pay attention to the isolation of the passive elements in the layouts. However, there is another way that may help extend the upper bound of the operating frequency, namely calibration, which can effectively inhibit the nonidealities from growing simultaneously as the operating frequency increases. With the calibration procedures that will be given in a future publication, the nonidealities can be suppressed and the functionality of the S_{21} detection network can be kept within a certain range. The effect of suppressing the nonidealities is equivalent to that of extending of the upper bound of the operating frequency.

As elucidated, the proposed network may encounter its upper bound as the operating frequency continues to grow higher. To deal with this potential bottleneck more effectively, we describe two important points of view as guidelines. First, as the frequency becomes higher, the inductance and capacitance values within the DIV will decrease. This means that there is more room for implementing elements' separation on chip. Hence, designers are strongly encouraged to take full advantage of these extra areas to reduce unwanted parasitic inductance and capacitance as much as possible. Second, as far as the area overhead is concerned, one may wonder whether the proposed network is still practicable as today's IC technologies continue to evolve. These doubts may arise especially when the semiconductor technologies are currently migrating into the 3-nm era, with the operating frequencies of transistors becoming higher and higher. However, this does not mean that the proposed S_{21} detection network has become obsolete, for three reasons.

First, although the process scales continue to shrink, the shrinkage generally applies only to the transistors in the digital circuits. That is, due to the design preferences of analog and RF circuits, the downsizing trend in active transistors does not extend to analog and RF circuits. In other words, analog and RF circuit designers would normally not use state-of-the-art active components. They would choose state-of-the-art technologies only for special cases: for example, if a tiny portion of an IC needs to have extremely high-speed transistors or when the foundries want to promote new processes by offering a special price. Since the proposed detection network is used to measure RFICs, it will not discord with the DUT because both will reduce in size as the frequency increases and because the speed of downsizing the DUT is, in general, lower than the speed of downsizing semiconductor technologies. For instance, by comparing a power amplifier that is fabricated by 90-nm CMOS technology and operating at 5.8 GHz with an amplifier that is fabricated by 65-nm CMOS technology and operating at 60 GHz, we find that the former [19] occupies a die area of 0.9 mm $\times$ 0.9 mm and the latter [20] a die area of 0.825 mm $\times$ 0.825 mm. As can be seen, the area reduction of the power amplifiers, about 16%, is not as large as the reduction of the smallest transistors' sizes, about 28%.

Second, the costs of extra chip area of the proposed detection network can be further reduced if the network can be reused to measure different DUTs through the utilization of modern RF switches. As the RF switch performance continues to improve [21], [22], designers can place these switches between the DUTs of interest and the proposed detection network. With this arrangement, the same detection network can be selectively used to measure different DUTs depending on which DUT's RF switches are turned on. That is, the area overhead created by the testing circuit can be reduced to a minimum since only one detection network is needed to measure many DUTs.

Third, as the operating frequencies of RFICs continue to rise, the external testing instruments, such as VNAs, RF cables, and RF probes, will become more expensive. Therefore, whether it is worthwhile to have embedded S-parameter measurement cannot and should not be judged solely by the extra costs caused by the embedded testing circuits. Instead, the expenses due to measuring RFICs with external instruments need to be taken into consideration as well. Besides, the higher the measuring frequencies become, the more delicate and vulnerable these external measurement instruments will tend to be. Therefore, the hidden expenses

of maintenance and replacement of the external measurement instruments will get higher as well. This is also a possible source of major costs that adopting external S-parameter measurement may bring about and that circuit designers must factor in when considering measuring S-parameters externally.

Although the proposed S_{21} detection network may cause extra costs by occupying extra chip area, it has a crucial niche in some special applications. Of all the possible applications, the most important and valuable one is to measure devices within 3D ICs. Buried deeply inside 3D ICs, the inner devices become very difficult to measure through external VNAs. In fact, because the RF stimulus signals transmitted from and received by the VNAs need to go through so many stacks, micro bumps, and through-silicon-vias (TSVs), as well as the DUT itself, and because the normalization plane is usually set at the probe tips, it is almost impossible for external VNAs to accurately measure the inner devices of 3D ICs. In other words, the VNA fails because the measurement results contain the total information of all the components on the measuring path, including the DUT and other unwanted terms, but the VNA is unable to retrieve the information about the DUT by de-embedding the information of other components.

In addition to the above reasons that 3D ICs prefer embedded, rather than external, S-parameter measurement, 3D ICs do have other serious issues that need to be carefully monitored. For example, joule heating, vertical interconnection, electro-migration, and stress-migration [23] can all have serious and unknown effects on the performance of the inner devices of 3D ICs. And the urgent need to monitor these serious issues is where the proposed S_{21} detection network comes into play. With the help of the detection network, measurements can be taken from inside the chips, so the results will be more trustworthy. Some golden DUTs, as well as RF switches, can be placed inside 3D ICs to help the S_{21} detection network execute the calibration procedures. These golden DUTs can be located at places where no components with changeable process variations, such as TSVs, exist. In this way, these golden DUTs are expected to bear a close resemblance to their expected behavior from simulation and to provide better calibration quality for the S_{21} detection network.

VI. CONCLUSION

This paper proposes an integrated S_{21} detection circuit that can help realize on-chip S_{21} measurements. The proposed concept has been shown to be feasible based on the successful implementation of real on-chip lumped passive devices with reasonable sizes. The proposed S_{21} detection circuit is a three-port network that is constructed mainly by three lumped passive integrated DIVs. The S_{21} detection network receives input RF signals at its Port 1 and sends the processed signals to its Port 2 and Port 3. When placed in the right locations, the three inner DIVs are able to manipulate the input RF signals properly at the targeted operating frequencies and, therefore, help measure the S_{21} of the device under test.

To verify the design concept, we fabricated the detection network using the 0.18-μm IBM 7RF semiconductor technology. The total die area including the pads is 1.2 mm $\times$ 1.2 mm. The measurement results show the expected behavior at around the designed frequency. The well-verified test chips demonstrate the performance and feasibility of the proposed circuit. In conclusion, the simple circuit structure and the downsizing trend of chip designs as the measuring frequency increases make the proposed network a favorable and practicable solution for embedded S-parameter measurements. In the future, the whole circuit can be further reduced in size when the measuring frequency rises, an advantageous development as far as chip sizes and fabrication costs are concerned. In addition, thanks to the high maturity of semiconductor technologies, peak and phase detectors can easily be cascaded to the proposed network to create a more integrated system for embedded testing.

Based on the results obtained in this paper, we propose two possibilities that can be further investigated in future research. First, to make the S_{21} detection network more integrated as a measuring system, it is possible and beneficial to add some phase and amplitude detectors in the right part of the detection network in Fig. 4, as demonstrated in Subsection V.B. Another possible and important future research direction is to try to build a detection network that can properly manipulate the RF signal to enable the on-chip S_{11} measurement of a DUT. With the S_{21} detection circuit proved to be successful in this study, if an S_{11} detection network can be constructed, the whole set of four S-parameters, S_{11}, S_{12}, S_{21}, and S_{22}, of a 2-port DUT can be measured on chip. This is so because S_{12} and S_{22} are simply the counterpart of S_{21} and S_{11} with a reverse measurement configuration.

ACKNOWLEDGMENT

The authors would like to thank the editor and anonymous reviewers for their invaluable comments on the article. Their comments have been very constructive, helping us to improve the quality of the article.

REFERENCES

[1] G. Gonzalez, "Presentations of two-port networks," in *Microwave Transistor Amplifiers: Analysis and Design*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1997, ch. 1, pp. 1–86.

[2] J. P. Dunsmore, *Handbook of Microwave Component Measurements: With Advanced VNA Techniques*. Hoboken, NJ, USA: Wiley, 2012.

[3] J. L. Hesler, Y. Duan, B. Foley, and T. W. Crowe, "THz vector network analyzer measurements and calibration," in *Proc. 21st Int. Symp. Space THz Technol.*, Oxford, U.K., 2010, pp. 318–320.

[4] K. Jayaraman, Q. Khan, B. Chi, W. Beattie, Z. Wang, and P. Chiang, "A self-healing 2.4 GHz LNA with on-chip S_{11}/S_{21} measurement/calibration for *in-situ* PVT compensation," in *Proc. IEEE Radio Freq. Integr. Circuits Symp.*, Anaheim, CA, USA, May 2010, pp. 311–314.

[5] A. Goyal, M. Swaminathan, and A. Chatterjee, "Self-correcting, self-testing circuits and systems for post-manufacturing yield improvement," in *Proc. IEEE 54th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Seoul, South Korea, Aug. 2011, pp. 1–4.

[6] Y.-R. Wu, Y.-K. Hsieh, P.-C. Ku, and L.-H. Lu, "A built-in gain calibration technique for RF low-noise amplifiers," in *Proc. IEEE 32nd VLSI Test Symp. (VTS)*, Napa, CA, USA, Apr. 2014, pp. 1–6.
[7] D. C. Howard, P. K. Saha, S. Shankar, T. D. England, A. S. Cardoso, R. M. Diestelhorst, S. Jung, and J. D. Cressler, "A SiGe 8-18-GHz receiver with built-in-testing capability for self-healing applications," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 10, pp. 2370–2380, Oct. 2014.
[8] J. Nehring, M. Schütz, M. Dietz, I. Nasr, K. Aufinger, R. Weigel, and D. Kissinger, "Highly integrated 4-32-GHz two-port vector network analyzers for instrumentation and biomedical applications," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 1, pp. 229–244, Jan. 2017.
[9] B. Philippe and P. Reynaert, "An F-band n-probe standing wave detector for complex reflection coefficient measurements in 40-nm CMOS," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 10, pp. 4278–4286, Oct. 2019.
[10] T. Zhang, W. R. Eisenstadt, and R. M. Fox, "A novel 5 GHz RF power detector," in *Proc. IEEE ISCAS*, Vancouver, BC, Canada, vol. 1, May 2004, pp. 897–900.
[11] Q. Yin, W. R. Eisenstadt, and R. M. Fox, "A translinear-based RF RMS detector for embedded test," in *Proc. IEEE ISCAS*, Vancouver, BC, Canada, vol. 1, May 2004, pp. 245–248.
[12] T. Zhang, W. R. Eisenstadt, and R. M. Fox, "20 GHz bipolar RF RMS power detectors," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, Santa Barbara, CA, USA, 2005, pp. 204–207.
[13] J. Yoon, R. M. Fox, and W. R. Eisenstadt, "Integrated BiCMOS 10 GHz S-parameter module," in *Proc. Bipolar/BiCMOS Circuits Technol. Meeting*, Maastricht, The Netherlands, 2006, pp. 1715–1720.
[14] D. M. Pozar, "Power dividers and directional couplers," in *Microwave Engineering*, 3rd ed. Hoboken, NJ, USA: Wiley, 2005, ch. 7, pp. 308–365.
[15] J. A. G. Malherbe, *Microwave Transmission Line Filters*. Norwood, MA, USA: Artech House, 1979.
[16] J. A. G. Malherbe, *Microwave Transmission Line Couplers*. Norwood, MA, USA: Artech House, 1988.
[17] J. Yoon and W. R. Eisenstadt, "Lumped passive circuits for 5 GHz embedded test of RF SoCs," in *Proc. IEEE ISCAS*, Vancouver, BC, Canada, vol. 1, May 2004, pp. 1241–1244.
[18] S. S. Mohan, M. del Mar Hershenson, S. P. Boyd, and T. H. Lee, "Simple accurate expressions for planar spiral inductances," *IEEE J. Solid-State Circuits*, vol. 34, no. 10, pp. 1419–1424, Oct. 1999.
[19] P. Haldi, D. Chowdhury, P. Reynaert, G. Liu, and M. Ali Niknejad, "A 5.8 GHz 1 V linear power amplifier using a novel on-chip transformer power combiner in standard 90 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 43, no. 5, pp. 1054–1063, May 2008.
[20] J. Lai and A. Valdes-Garcia, "A 1V 17.9 dBm 60 GHz power amplifier in standard 65 nm CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, USA, Feb. 2010, pp. 424–425.
[21] T. Y. Lee and S. Lee, "Modeling of SOI FET for RF switch applications," in *Proc. IEEE Radio Freq. Integr. Circuits Symp.*, Anaheim, CA, USA, May 2010, pp. 479–482.
[22] R. Wolf, A. Joseph, A. Botula, and J. Slinkman, "A thin-film SOI 180 nm CMOS RF switch," in *Proc. IEEE Top. Meeting Silicon Monolithic Integr. Circuits RF Syst.*, San Diego, CA, USA, Jan. 2009, pp. 1–4.
[23] K. N. Tu, "Reliability challenges in 3D IC packaging technology," *Microelectron. Rel.*, vol. 51, no. 3, pp. 517–523, Mar. 2011.

MING-CHE LEE received the B.S. and M.S. degrees in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2000, and 2002, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in May 2010.

From 2010 to 2012, he worked as a Circuit Designer with ILITEK and as an Assistant Engineer with the National Synchrotron Radiation Research Center, Hsinchu, Taiwan. From 2012 to 2016, he served as a Flight Operation Officer in the Kaohsiung International Airport, Taiwan. Since 2016, he has been an Associate Technical Specialist with the Civil Aeronautics Administration of Taiwan, and has been an Adjunct Assistant Professor with the Department of Mechanical Engineering, National Taipei University of Technology.

CHI-YO HUANG received the B.S. degree in electrical engineering from National Cheng-Kung University, Taiwan, in 1990, the M.S. degree in computer engineering from Syracuse University, Syracuse, NY, USA, in 1993, and the Ph.D. degree in management of technology from National Chiao-Tung University, Taiwan, in 2006. From 1994 to 2006, he worked in the IC industry as an IC Design Engineer, a Marketing Manager, and the Director. He is currently a Distinguished Professor with the Department of Industrial Education, National Taiwan Normal University, Taiwan. His current research interests include management and applications of technology, multiple attribute decision making, and business data analytics.