Wideband mm-Wave and sub-THz Chip-to-Package Interfaces in Low Cost Organic Substrates

Rami Hijab

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2025-40
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2025/EECS-2025-40.html

May 2, 2025

Wideband mm-Wave and sub-THz Chip-to-Package Interfaces in Low Cost Organic Substrates

by

Rami Raif Mehdi Hijab

A thesis submitted in partial satisfaction of the

requirements for the degree of

Masters of Science, Plan II

in

Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ali M. Niknejad, Chair
Professor Jun-Chau Chien

Fall 2024

The thesis of Rami Raif Mehdi Hijab, titled Wideband mm-Wave and sub-THz Chip-to-Package Interfaces in Low Cost Organic Substrates, is approved:

Chair	Ali Niknejad	Date	12/20/2024
	Jun-Chau Chien	Date	12/20/2024
		Date	

University of California, Berkeley

Wideband mm-Wave and sub-THz Chip-to-Package Interfaces in Low Cost Organic Substrates

Abstract

Wideband mm-Wave and sub-THz Chip-to-Package Interfaces in Low Cost Organic Substrates

by

Rami Raif Mehdi Hijab

Masters of Science, Plan II in Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Ali M. Niknejad, Chair

The transition from transceiver, on integrated-circuit (IC), to antenna, on package or printed circuit board (PCB), is critical for future broadband communication systems that plan to operate in the mm-Wave or sub-THz frequency bands. To date, reported chip-to-package transitions either have high loss, typically $3-4\,\mathrm{dB}$, or require high cost packages to support low loss materials or very fine bump pitches. This work analyzes the impact of transitions on a high frequency, wide bandwidth communication system. The theory, analysis and design of a chip-to-package transition in a commerical CMOS and organic substrate technology is presented. The developed theory and design principles are validated through the measurement of another chip-to-package transition, demonstrating $\leq 1\,\mathrm{dB}$ loss over a large bandwidth. Despite using large dimensions typical of low cost packaging or PCB solutions, the proposed and demonstrated transitions offer competitive performance.

To my parents,

I hope I make you proud.

To the people of Gaza,

your strength astounds me.

Contents

List of Figures

List of Tables

Acknowledgments

While this is only an intermediate milestone, the people that have contributed to my progress thus far certainly deserve acknowledgment.

First and foremost I would like to thank my advisor, Professor Ali Niknejad. It has been a real pleasure working with him. The breadth of his technical mastery is inspiring and motivating, and I look forward to continuing to learn from him. I would also like to thank Professor Jun-Chau Chien for being on the committee and providing feedback on this work.

A special thanks goes to Nima Baniasadi, for getting the project behind this work started and providing guidance and discussions even after graduating from Berkeley. This work does not exist without him. I would also like to thank all the faculty, staff and students at the BWRC. You make the ongoing experience enjoyable.

Finally, I would like to thank my friends and family, all of whom I am very grateful for. Thank you to Rachel Gunawan, who, for better or worse, has seen the brightest and the darkest moments, and supports me regardless. Thank you to both of my parents who are my foundations, challenging me to be better and chase my dreams. They are the reason I have made it this far.

In writing this work I have heavily referenced J.D. Jackson and R.E. Collin. I find it amusing that well over forty years later, I regularly use the copies from my dad's library collection, stamped with his name and scribbled with his notes!

Chapter 1

Introduction

1.1 Motivation

The demand for faster wireless data rates has continued to grow at an increasing rate, with the number of connected devices exceeding 18 billion—far surpassing the human population. Beyond personal connectivity and entertainment, wireless communication is now indispensable for applications such as transportation, healthcare, infrastructure, and others yet to be imagined [1].

To meet these escalating demands, more spectrum and bandwidth are required. However, the Federal Communications Commission (FCC) spectrum below 10 GHz is overcrowded. Research efforts are thus directed toward relatively underutilized millimeter-wave (mm-Wave) bands from 30 GHz-300 GHz[1] [47, 7, 8]. The latest 5G cellular standard introduced mm-Wave bands from 28 GHz to 60 GHz, which are now undergoing widespread deployment, and a new standard for the sub-terahertz (sub-THz) bands has been released in 802.15.3d by the IEEE [35]. In addition to communications, sensing applications such as radar also benefit from the finer spatial resolution enabled by operation in the mm-Wave bands.

While promising, operation at these frequencies poses significant challenges. Integrated circuit (IC) technologies exhibit reduced performance due to limited device speeds or an inability to integrate with digital logic. Additionally, free-space path losses increase dramatically at higher frequencies due to reduced wavelengths.

To address these issues, transceivers need integration into massive arrays, but this presents challenges in integration design. The focus of this work is a key component required for the integration of many transceiver ICs on a single package, the signal transition from chip to package, shown in Fig. 1.1

[1]Sub-terahertz (sub-THz) is another commonly used term that generally refers to 100 GHz-300 GHz, the upper portion of the mm-Wave band, which this work focuses on.

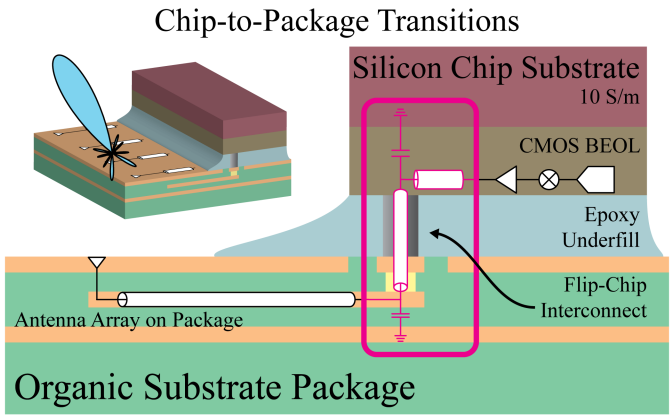

Figure 1.1: A flip-chip CMOS chip-to-package transition on a low-cost organic substrate interposer for mm-Wave communication systems.

1.2 Technology Choices

1.2.1 IC Technology

The IC technology used for a transceiver at sub-THz frequencies should be able to provide the performance requirements while allowing ease of integration with digital logic, and most importantly being compatible with high-volume manufacturing in terms of yield. For the latter two points, CMOS has clear advantages as the technology of choice for very high-volume computer processors. Furthermore, advanced nodes, such as 28 nm Bulk CMOS or even scaled FinFET nodes have been able to reach $f_{\max}$ values exceeding 400 GHz [25, 27]. $f_{\max}$ is a figure of merit for the speed of transistors that will be used to compare technologies, as it represents the frequency beyond which the active device can no longer provide power gain[2]. A chart of the $f_{\max}$ and f_T, or frequency of maximum current gain, is shown in Fig. 1.2.

While CMOS has clear advantages over other technologies in terms of cost (in volume) and yield, it has limited output power due to the low breakdown voltages and modest speed. III-V technologies such as indium phosphide (InP) or gallium nitride (GaN) offer superior output power as compared to CMOS [9]. Nevertheless, implementations in CMOS are preferred for large-scale systems targeted in this work.

[2]The theoretical basis for this merit originates from Mason's invariant unilateral gain and details can be found in [18].

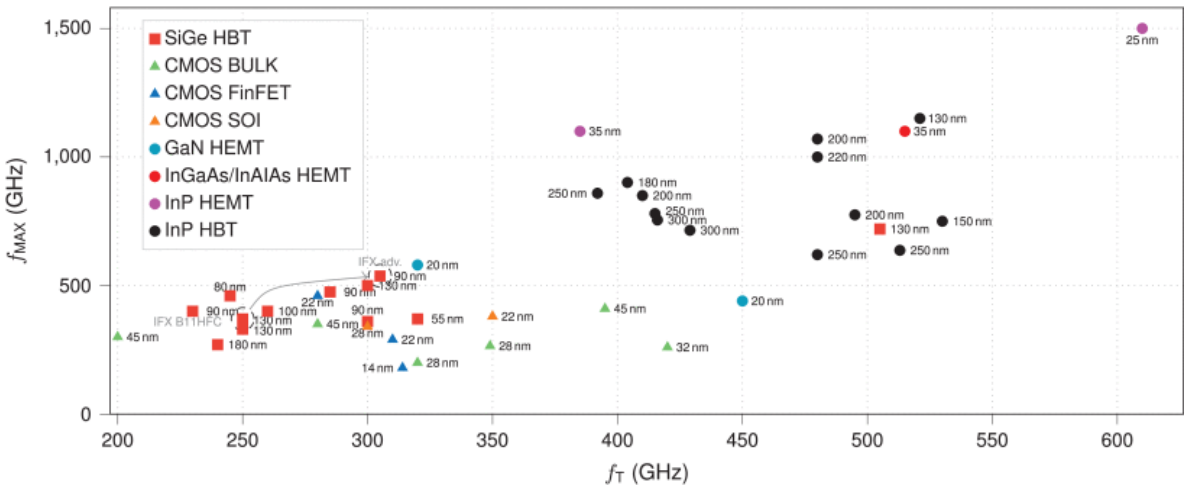

Figure 1.2: A survey of transistor speeds by technology reproduced from [27]. f_T is the frequency at which the current gain becomes unity and f_{max} is the frequency at which the power gain becomes unity. In a CMOS technology dominated by the intrinsic parasitics $f_{\max} \approx \frac{f_T}{2}\sqrt{5 g_m r_o}$, where $g_m r_o$ is the intrinsic gain [32].

1.2.2 Antenna Integration

Regardless of technology, arrays of transceivers are required to compensate for additional path losses at mm-Wave frequencies. This poses an integration challenge for massive arrays of transceivers and antennas. The traditional approach is to include the antenna on PCB, but at these frequencies the losses in connecting to PCB are too high. As the antenna size becomes comparable to, or smaller than, the size of the IC, one possible solution is to integrate the antenna on-chip, shown in Fig. 1.3a. This mitigates some of the integration challenges, but substrate surface waves and ohmic losses of on-chip antennas lead to low radiation efficiency, which limits their performance [33, 21, 34]. Furthermore, the antenna area occupies a significant portion of the chip area. Since most antenna arrays have an element spacing of $\lambda/2$, the available transceiver area can become a limiting factor.

Instead, the antenna can be integrated on an intermediate package between the IC and PCB, as shown in Fig. 1.3b. Antennas on the package have higher gain and radiation efficiency since the substrate is not as lossy as a doped substrate on chip and the metals are generally thicker. Furthermore, the large integration needs of massive arrays can be met since the losses in transitioning to the package are much lower than that of going directly to the PCB [27].

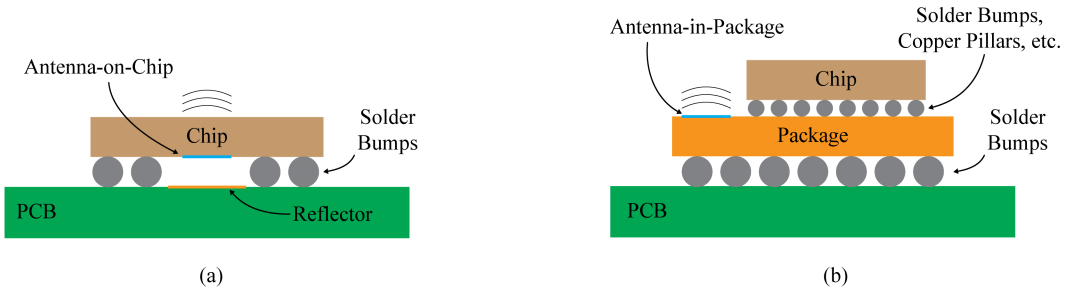

Figure 1.3: Two of many possible schemes for antenna integration in the sub-THz bands: (a) antenna-on-chip with flip-chip interconnect and a PCB reflector for backside radiation and (b) an integrated package between the chip and PCB for lower loss connection to a high efficiency antenna. More possible integration schemes are discussed in [27].

1.2.3 Packaging Interconnect

Even outside of the wireless communications domain, high performance packaging is becoming critical. For example, a move to multi-chip systems-on-a-package has led to the development of the Universal Chiplet Interconnect Express (UCIe) chip-to-chip communications standard. This standard facilitates inter-chip communication when chips co-exist on a single package, and there are two supported packages, standard and advanced supporting bump pitches of 100 µm-130 µm and 25 µm-55 µm respectively [11].

Bond wires are a traditional approach that involves using thin wires (typically gold, copper, or aluminum) to connect the chip pads to the substrate or package. While this method is low-cost and provides flexibility in design, its high parasitic inductance and resistance limit performance at high frequencies [48].

An alternative is flip-chip interconnects, where the die is flipped upside-down and solder bumps are used to directly connect it to the substrate. This approach significantly reduces parasitics and improves thermal dissipation [20]. Despite being more expensive and requiring a more complex assembly process, flip-chip interconnects have become a standard for high-volume manufacturing.

Instead of solder bumps, copper pillars can be used in flip-chip interconnects, where cylindrical copper structures capped with solder provide interconnections between the die and substrate. These pillars offer improved electrical and thermal performance and support higher current-carrying capacity compared to solder bumps at a further increased fabrication complexity and cost [16]. Microbumps are an even more advanced technology, and enable fine pitches, such as those used by the advanced package in UCIe. By utilizing extremely small solder bumps, sub-50 µm interconnections can be made, however they have yet to find high volume usage due to stringent manufacturing requirements and unproven reliability

[28][3].

1.2.4 Package Technology

While low-temperature co-fired ceramic (LTCC), laminate, or even silicon-based package technologies can offer lower loss or higher integration density, this comes at added cost and thermal/mechanical considerations due to thermal expansion mismatch between materials. Meanwhile, organic substrates provide multilayered packages that utilize similar manufacturing techniques as traditional PCB processes, while still providing sufficient pitch for use with ICs and lower loss than direct connection to the PCB. This makes them an attractive alternative to other packaging options [50].

1.3 Link Budget

The motivation for bandwidth comes primarily from Shannon's capacity theorem. It states that the maximum rate of information transmission through an additive white Gaussian noise channel (AWGN) with signal-to-noise ratio given, SNR, is given by $C = B\log_2(1 + \text{SNR})$ [40]. B is the bandwidth limit of the channel and $\log_2(1 + \text{SNR})$ is a spectral efficiency, relating how many bits can be transmitted per unit hertz. While data rates can rapidly increase with bandwidth—linearly as opposed to logarithmically with SNR—the sensitivity to SNR degradation also increases rapidly. This sensitivity increases in MIMO arrays.

The SNR for a MIMO system with N_{ant} antennas can be written as

$$\text{SNR} = \frac{P_t G_t N_{ant}^2}{4\pi d^2} \frac{\lambda^2 G_r N_{ant}^2}{4\pi} \frac{1}{k_B T B F N_{ant}} \tag{1.1}$$

where the first term represents the transmitted power density a distance d away from the source that transmits P_t power per element with an antenna gain of G_t; the second term represents the effective area of the receiver at wavelength λ with antenna gain G_r; and the third term is the input-referred thermal noise floor of the receiver that has noise factor F and the input noise power $P_n = k_B T B$. The signal on the transmitter and receiver side are both assumed to coherently combine from each element, whereas the noise is uncorrelated and its variance adds, leading to a total SNR boost from the array of N_{ant}^3.[4]

In a system with antennas on package, the impact of signal transition losses can be estimated by modifying Eq. (1.1) to explicitly consider IL_{trans}, the insertion loss between the transceiver IC and antennas, which is assumed to occur equally on the transmitter and

[3] In many ways the fabrication capabilities of the packaging world is playing catch-up to the advanced lithographic capabilities of the IC world. Nevertheless, systems-on-package are driving rapid development, and it is reasonable to expect technologies like microbumps to be reliably available in the near future.

[4] There are some subtleties to this analysis of the relation of SNR to N_{ant}, but this result is sufficient for the purpose of this analysis. See [14] for more details.

receiver. The SNR becomes

$$\text{SNR} = \frac{P_t G_t N_{ant}^2}{4\pi d^2 I L_{trans}} \frac{\lambda^2 G_r N_{ant}^2}{4\pi I L_{trans}} \frac{1}{k_B T B F N_{ant}} \propto \frac{1}{I L_{trans}^2}. \tag{1.2}$$

Suppose a base station has an N_{ant}-antenna phased array with N_{pol} polarizations operating at a bandwidth B with N_{beams} beams, where one beam is equivalent to one user. Assuming the beams are independent and have equal channel gains, then for SNR $\gg 1$ the aggregate capacity is

$$C_{tot} \approx N_{beams} N_{pol} B \log_2 \left(\frac{\text{SNR}}{N_{beams}} \right) \tag{1.3a}$$

$$\propto N_{beams} N_{pol} B \log_2 \left(1/I L_{trans}^2\right). \tag{1.3b}$$

The sensitivity to changes in transition losses can then be calculated to be

$$\text{S}_{\text{trans}} \approx -\frac{2\log_2(10)}{10} N_{beams} N_{pol} B \text{ bits/s/dB}. \tag{1.4}$$

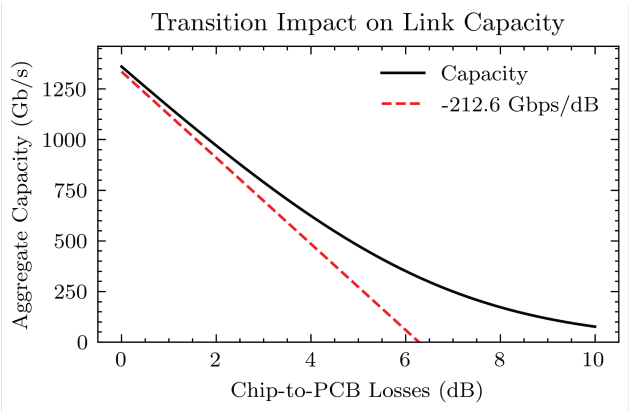

Figure 1.4: The impact of chip-to-package transition losses on link capacity for a potential future wideband wireless link. Here $f_0 = 140\,\text{GHz}$, $B = 20\,\text{GHz}$, $N_{ant} = 16$, $N_{beams} = 8$, $N_{pol} = 2$, $P_t = 4\,\text{dBm}$ after back-off, $G_r = G_t = 5\,\text{dB}$, $d = 5\,\text{m}$, and $F = 10\,\text{dB}$.

As an example, consider a MIMO beamforming system with $N_{beams} = 8$ beams and $N_{pol} = 2$ polarizations. At a bandwidth of $B = 20GHz$, tera-bit per second maximum capacity can be achieved, shown in Fig. 1.4. At a carrier frequency of 140 GHz this system is under 15 % fractional bandwidth, showing the promise of such a system. However, the resulting sensitivity to packaging loss is approximately 212 Gbits/s/dB. This high sensitivity highlights the importance of minimizing losses between active devices and the antenna in high bandwidth systems, where fractions of a dB of loss have a large impact on system performance.

1.4 Organization of this Thesis

This research has focused on the analysis, design and measurements of signal transitions from integrated circuit to package. In Chapter 2 key concepts in electromagnetics and transmission line circuit theory will be reviewed and particular emphasis will be placed on circuit abstractions and their utility for design. In Chapter 3 the working principles of several key interconnect types will be discussed and analyzed in detail. The design of a signal transition using low cost, high volume technologies is presented in Chapter 4, achieving record performance in terms of insertion loss and bandwidth. This theory is validated with the measurement of a transition design in Chapter 5, demonstrating that the principles and design methodologies developed are well formed. Chapter 6 concludes the work and provides areas for future development.

Chapter 2

Theoretical Background

Chip and package interfaces are passive electromagnetic structures and can be effectively described using Maxwell's equations. While the coupled, linear, vector differential equations are insightful, their complexity makes them impractical for hand analysis and almost impossible for designing intricate geometries. In practice, advanced modern computer 2.5D and 3D EM simulators, such as Cadence EMX and Ansoft HFSS, are used to model the physics. Nevertheless, a grounding in the field description of structures is required to optimize their design.

To provide design insight, key concepts of both field-based wave equations and circuit-based transmission line wave equations are reviewed and related. The intuition behind discontinuities and the modeling of energy storage in electric and magnetic fields is discussed. Other unique properties of transmission lines and waveguides are discussed. This serves as a basis for connecting physical geometries in EM simulators to designer-friendly circuit models. Derivations are avoided in favor of relating field and circuit equations, and providing intuition. See [22, 10, 36, 32] for further details.

2.1 Electromagnetics and the Relation Between Field and Circuit Transmission Lines

2.1.1 Electromagnetic Foundations

Maxwell's equations approximate electromagnetic classical phenomena well [22], and can be written in the time harmonic case in point, vector differential form as Eq. (2.1). At the field intensities and frequencies of interest, the materials used in integrated circuit and package technologies are linear and isotropic. Therefore, the magnetic flux density and electric flux density are related to the magnetic field and electric field by the linear, scalar constitutive relations, $\mathbf{B} = \mu\mathbf{H}$ and $\mathbf{D} = \epsilon\mathbf{E}$, respectively, where μ is the magnetic permeability and ϵ is the electric permittivity. In vacuum $\epsilon = \epsilon_0 = 8.85\,\mathrm{pF/m}$ and $\mu = \mu_0 = 1.26\,\mu\mathrm{H/m}$.

When losses need to be considered, ϵ can be made complex as $\epsilon' - j\epsilon''$, and the same for $\mu = \mu' - j\mu''$. Conductive losses can be included through Ohm's law, $\mathbf{J} = \sigma\mathbf{E}$.

$$\nabla \cdot \mathbf{B} = 0 \tag{2.1a}$$
$$\nabla \cdot \mathbf{D} = \rho \tag{2.1b}$$
$$\nabla \times \mathbf{E} + j\omega\mathbf{B} = 0 \tag{2.1c}$$
$$\nabla \times \mathbf{H} - j\omega\mathbf{D} = \mathbf{J} \tag{2.1d}$$

As written, it is clear that ρ, the charge volumetric density, and $\mathbf{J}$, the current density, are the sources of the electric and magnetic fields. Furthermore, taking the divergence of Eq. (2.1d) and substituting Eq. (2.1b)[1] (in time domain) results in Eq. (2.2), the continuity equation. This equation states the conservation of electric charge, and therefore the charge density is the ultimate source of the fields.

$$\mathcal{J} + \frac{\partial \rho}{\partial t} = 0 \tag{2.2}$$

The fundamental nature of the electric and magnetic field can be understood further by examining them in the context of special relativity. Historically, Maxwell's equations led to Einstein's development and discovery of special relativity as a way to explain the speed of light[2]. If the charge density and current density are replaced with a 4-vector, and the electric and magnetic field are written in a 2-tensor, then all four of Maxwell's equations can be written compactly as one equation. The benefit is that the equations transform naturally according to Lorentz transformations as a single vector or tensor, which implies that the electric and magnetic field are really the same quantity, and vary depending on the frame of reference – the same goes for charge and current density [22].

2.1.2 Modal Wave Equations

Generally when solving problems we are interested in the solution of Maxwell's equations in a region absent of free charge (e.g. vacuum or a dielectric), and we use conductors as boundary conditions for the solution. Furthermore, we are primarily interested in wave propagation, since this is how signals will be translated from one location to another. Then the problem can be further simplified by solving in the plane normal to the direction of propagation. This

[1] Assuming $\mathcal{D}$ has continuous second derivatives for exchange of divergence and partial derivative with respect to time.

[2] "The precise formulation of the time-space laws was the work of Maxwell. Imagine his feelings when the differential equations he had formulated proved to him that electromagnetic fields spread in the form of polarized waves, and at the speed of light! To few men in the world has such an experience been vouchsafed ... it took physicists some decades to grasp the full significance of Maxwell's discovery, so bold was the leap that his genius forced upon the conceptions of his fellow workers." Einstein (Science, May 24, 1940)

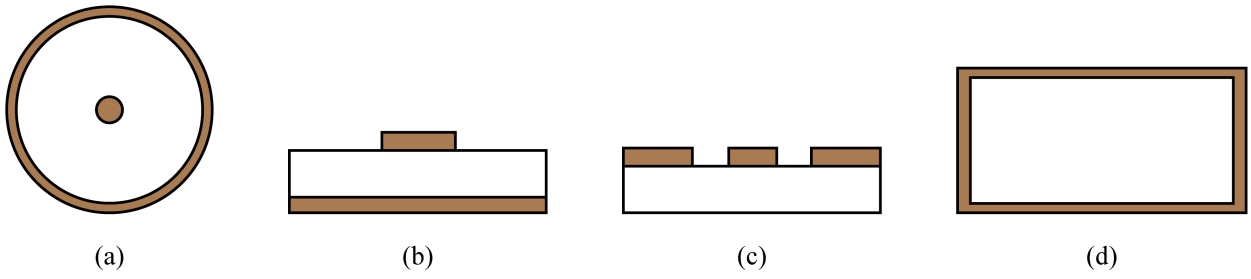

Figure 2.1: Some example of common transmission lines. (a) The coaxial line made of an inner signal conductor and outer ground conductor. (b) The microstrip transmission line made of a signal line and a ground plane that sandwich a dielectric material. (c) The coplanar waveguide that has a signal line in the center, and two ground lines on the outer conductors. (d) The rectangular waveguide which is a single conductor with dielectric or air inside. These examples support TEM, quasi-TEM, and TE/TM mode propagations.

is the case of guided waves, abstracted as transmission lines, and cross-sections of example geometries are shown in Fig. 2.1. If the conductors surrounding the volume are perfect electric conductors, then the fields do not penetrate the conductors and this is the only solution needed. However, even if the conductors are lossy, a low-loss approximation can be used to determine the resulting perturbation to the solution [22, 10]. Therefore, the solution of the equations in a source-free region is of primary interest. These equations are shown in Eq. (2.3)

$$\nabla^2\mathbf{E} + k^2\mathbf{E} = 0 \tag{2.3a}$$
$$\nabla^2\mathbf{H} + k^2\mathbf{H} = 0 \tag{2.3b}$$

where $k = \omega\sqrt{\mu\epsilon}$ is the propagation constant.

Eq. (2.3a) represents a vector wave equation. The simplest solution is a plane wave, written as

$$E_x = E^+e^{-jkz} + E^-e^{jkz} \tag{2.4a}$$
$$H_y = \frac{1}{\eta}\left(E^+e^{-jkz} - E^-e^{jkz}\right) \tag{2.4b}$$

where $\eta = \sqrt{\mu/\epsilon}$ is the intrinsic impedance of the medium. The first exponential represents a $+z$-traveling wave because, when multiplied by $e^{j\omega t}$ the exponent is $j(\omega t - kz)$, which has a positive phase velocity[3] of $dz/dt = \omega/k = 1/\sqrt{\mu\epsilon}$. The second exponent is a $-z$-traveling wave.

[3]Phase velocity is termed as such because it is the velocity of a fixed point (phase) along the wave.

	TEM	TE	TM
Axial Field Components	$E_z = H_z = 0$	$E_z = 0,\ H_z \neq 0$	$H_z = 0,\ E_z \neq 0$
Propagation Constant, β	$k = \omega\sqrt{\mu\epsilon}$	$\sqrt{k^2 - k_c^2}$	$\sqrt{k^2 - k_c^2}$
Wave Impedance, $Z_w = \frac{E_x}{H_y} = \frac{-E_y}{H_x}$	$Z_{TEM} = \frac{\omega\mu}{\beta} = \eta$	$Z_{TE} = \frac{\omega\mu}{\beta} = \frac{k}{\beta}\eta$	$Z_{TM} = \frac{\beta}{\omega\epsilon} = \frac{\beta}{k}\eta$

Table 2.1: A summary of the key properties of TEM, TE, and TM wave propagation.

The general solution of Eq. (2.3) is more complicated,

$$\mathbf{E}(x,y,z) = (\mathbf{e}(x,y) + \hat{z}e_z(x,y))e^{-j\beta z} + (\mathbf{e}(x,y) - \hat{z}e_z(x,y))e^{j\beta z} \tag{2.5a}$$

$$\mathbf{H}(x,y,z) = (\mathbf{h}(x,y) + \hat{z}h_z(x,y))e^{-j\beta z} - (\mathbf{h}(x,y) - \hat{z}h_z(x,y))e^{j\beta z} \tag{2.5b}$$

where $\mathbf{e}(x,y)$ and $\mathbf{h}(x,y)$ represent the transverse field in the x-y plane, and $e_z(x,y)$ and $h_z(x,y)$ are the field in the direction of propagation. The minus signs in Eq. (2.5) account for the physical reversal of the reflected wave for the terms carrying $e^{j\beta z}$, where β is the propagation constant.

The solutions in Eq. (2.5) can be further categorized based on the axial field components, E_z and H_z, as transverse electromagnetic (TEM), transverse electric (TE) and transverse magnetic (TM) waves. The wave properties are summarized in Table 2.1. For TEM waves, since $E_z = H_z = 0$[4], the solution to Eq. (2.3) simplifies to solving Laplace's equation in the transverse plane, $\nabla_t \Phi = 0$[5]. Therefore, the field components behave like static fields, and can be described by an inductance and capacitance per unit length. TE and TM modes are primarily characterized by the nonlinear relationship between the propagation constant and frequency, given by $\beta = \sqrt{k^2 - k_c^2}$. k_c is the cutoff wave number (directly related to a cutoff frequency), and describes the frequency below which the propagation constant becomes imaginary, and the wave exponentially decays. If a TE and TM wave have the same cutoff frequency, then $Z_{TE}Z_{TM} = \eta^2$, since the impedance varies inversely with the propagation constant for the two modes.

2.1.3 Transmission Line Model

Inspired by the description of an inductance and capacitance per unit length for TEM waves, an infinite L-C ladder, shown in Fig. 2.2, can be used as a circuit model for wave propagation. The voltage and current on the line as a function of position is given by

[4]Only lossless conductors can support TEM waves since the addition of loss leads to fields that penetrate the conductor, and the boundary conditions lead to an axial component. However, this effect is typically negligible and can be mostly ignored.

[5]Since Laplace's equation admits only unique solutions, a TEM mode requires more than two conductors. Otherwise, $\Phi = 0$ everywhere would be the solution, so the fields would be zero everywhere [32].

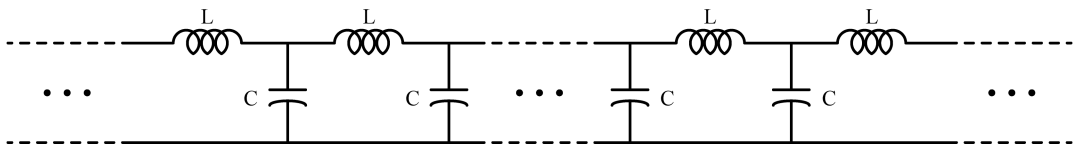

Figure 2.2: The circuit model for a transmission line.

$$v(z) = V^+ e^{-j\beta z} + V^- e^{j\beta z} \tag{2.6a}$$

$$i(z) = \frac{1}{Z_0}(V^+ e^{-j\beta z} - V^- e^{j\beta z}) \tag{2.6b}$$

where $j\beta = \sqrt{(j\omega L')(j\omega C')} = j\omega\sqrt{L'C'}$ and $Z_0 = \sqrt{(j\omega L')/(j\omega C')} = \sqrt{L'/C'}$. These equations are very similar to the plane wave solution for the fields in Eq. (2.4), where the inductance and capacitance per unit length of the volume is given μ and ϵ, respectively.

If the transmission line is terminated with an impedance Z_L, then at $z = 0$ the reflection coefficient is given by

$$\rho_L = \frac{V^-}{V^+} = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{2.7}$$

and the voltage and current on the line can be written as

$$v(z) = V^+ e^{-j\beta z}(1 + \rho_L e^{j2\beta z}) \tag{2.8}$$

$$= V^+ e^{-j\beta z}(1 + \rho_L(z)) \tag{2.9}$$

$$i(z) = V^+ e^{-j\beta z}(1 - \rho_L e^{j2\beta z}) \tag{2.10}$$

where the reflection coefficient has been defined as a function of distance. The input impedance a distance l away from the load can be written as

$$Z_{in}(-l) = Z_0 \frac{1 + \rho_L e^{-j2\beta l}}{1 - \rho_L e^{-j2\beta l}} \tag{2.11a}$$

$$= Z_0 \frac{Z_L + jZ_0 \tan(\beta l)}{Z_0 + jZ_L \tan(\beta l)}. \tag{2.11b}$$

Since the wave $v(z)$ consists of both a forward traveling wave and a backward traveling reflected wave, standing waves may form. The ratio of the peak of the standing wave to the the minimum of the standing wave is given when $2\beta z$ results in constructive and destructive interference, respectively, leading to a voltage standing wave ratio of $V_{max}/V_{min} = (1 + \rho_L)/(1 - \rho_L)$. The current standing wave ratio is the same, but the maximum occurs a quarter wavelength away from the voltage maximum.

2.1.4 Relation Between Field and Circuit Models

From the similarity of Eq. (2.4) and Eq. (2.6) it is clear a relation can be made between field wave propagation and the circuit wave propagation. This is easy when considering TEM waves, since the circuit model should simply follow the solutions of Laplace's equation for the inductance and capacitance (and resistance and conductance) per unit length. If the voltage and current on the line are V and I, respectively, the relationships are given by

$$L' = \frac{\mu}{|I|^2} \int_S \mathbf{H} \cdot \mathbf{H}^* dS \tag{2.12}$$

$$C' = \frac{\epsilon}{|V|^2} \int_S \mathbf{E} \cdot \mathbf{E}^* dS. \tag{2.13}$$

However, for TE and TM waves, voltages and currents are not well defined. Furthermore, the notion of impedance for the waves has multiple different meanings. The key definitions of impedance are [36]

- $\eta = \sqrt{\mu/\epsilon}$: the intrinsic impedance of a medium that does not depend on the geometry, and is the wave impedance for plane waves.
- $Z_w = E_t/H_t$: the wave impedance that is characteristic to a particular TEM, TE or TM wave and may depend on the geometry and frequency.
- $Z_0 = V^+/I^+$: the characteristic impedance of a circuit wave, as the ratio of the voltage to the current. For TEM waves, the voltage and current are uniquely defined on the line, so the characteristic impedance is unique, but this is not the case for TE and TM waves.

For TE and TM waves, the ambiguity of the relationship between voltage and current requires a modal relation to the circuits. While, this is arbitrary, the following considerations are useful [36]:

- Each mode has a distinct definition of voltage and current, and the values are chosen to be proportional to the transverse electric and magnetic fields, respectively.
- The voltage and current are defined so that the product gives the power flow of the waveguide mode.
- The ratio of the voltage and current should give the characteristic impedance, which is usually chosen to be the wave impedance.

2.1.4.1 Energy and Power

The conservation of energy can be stated in terms of the electric and magnetic fields, and used to make general statements about impedances. The time-average energy stored in the electric and magnetic fields is given by[6]

$$W_e = \frac{1}{2}\mathrm{Re}\int_V \mathbf{E}\cdot\mathbf{D}^* dV \tag{2.14}$$

$$W_m = \frac{1}{2}\mathrm{Re}\int_V \mathbf{H}\cdot\mathbf{B}^* dV \tag{2.15}$$

respectively. Furthermore, the power flow through a closed surface, S, is given by the integration of the complex Poynting vector, $\mathbf{E}\times\mathbf{H}^*$

$$P = \frac{1}{2}\mathrm{Re}\oint_S \mathbf{E}\times\mathbf{H}^*\cdot d\mathbf{S}. \tag{2.16}$$

With these results and Maxwell's equations, the Poynting theorem for conservation of energy can be written as

$$\nabla\cdot\mathbf{E}\times\mathbf{H}^* = -j\omega\mathbf{B}\cdot\mathbf{H}^* + j\omega\mathbf{D}^*\cdot\mathbf{E} - \mathbf{E}\cdot\mathbf{J}^* \tag{2.17a}$$

$$\frac{1}{2}\mathrm{Re}\oint_S \mathbf{E}\times\mathbf{H}^*\cdot(-d\mathbf{S}) = \frac{\omega}{2}\int_V(\mu''\mathbf{H}\cdot\mathbf{H}^* + \epsilon''\mathbf{E}\cdot\mathbf{E}^*)dV + \frac{1}{2}\int_V \sigma\mathbf{E}\cdot\mathbf{E}^* dV \tag{2.17b}$$

$$\frac{1}{2}\mathrm{Im}\oint_S \mathbf{E}\times\mathbf{H}^*\cdot(-d\mathbf{S}) = 2\omega\int_V\left(\mu'\frac{\mathbf{H}\cdot\mathbf{H}^*}{4} - \epsilon'\frac{\mathbf{E}\cdot\mathbf{E}^*}{4}\right)dV \tag{2.17c}$$

where dielectric loss, magnetic loss, and conductive loss have been considered. The real part represents the power loss through each source of loss, due to polarization damping forces in the first term and conduction current in the second term. The imaginary part represents the reactive energy stored in the volume.

A comparison with circuit theory can be made by considering a series RLC circuit. The complex power is given by $(1/2)|I|^2(R + j\omega L - j/(\omega C))$, and the time-average power loss, energy stored in the magnetic field, and energy stored in the electric field are given by

$$P_l = \frac{1}{2}R|I|^2 \tag{2.18}$$

$$W_m = \frac{1}{4}L|I|^2 \tag{2.19}$$

$$W_e = \frac{|I|^2}{4\omega^2 C} \tag{2.20}$$

[6]There are some subtleties with regard to dispersion in the losses of ϵ and μ that are discussed in [10].

respectively. Therefore, the complex power in the circuit is given by

$$\frac{1}{2}Z|I|^2 = P_l + 2j\omega(W_m - W_e) \tag{2.21}$$

and the impedance can be written as

$$Z = \frac{P_l + 2j\omega(W_m - W_e)}{\frac{1}{2}|I|^2}. \tag{2.22}$$

This shows that the real part of any impedance represents power loss, while the imaginary part represents energy stored [22, 10].

2.1.4.2 Perturbative Loss and Dispersion in Transmission Lines

For both field and circuit solutions, loss can be included by modifying the propagation constant $j\beta$ to take the form $\gamma = \alpha + j\beta$, where α is an attenuation coefficient. For the field solutions, the attenuation coefficient can be estimated from the solution of the lossless fields as follows. The lossless field solution for a TEM wave is given by

$$\mathbf{E} = -\nabla_t \Phi e^{-jkz} \tag{2.23}$$

$$\mathbf{H} = Y_w \hat{z} \times \mathbf{E}. \tag{2.24}$$

with $Y_w = 1/Z_w$, the wave impedance. With loss present, the power propagated along the line decreases according to $e^{-2\alpha z}$, and the rate of this decrease must be equal to the power loss, so

$$-\frac{\partial P}{\partial z} = P_l = 2\alpha P_0 e^{-2\alpha z} = 2\alpha P \tag{2.25}$$

and the power loss at any point, z is proportional to the total power at z. In the dielectric, the Maxwell-Ampere law can be written as

$$\nabla \times \mathbf{H} = \mathbf{J} + j\omega\mathbf{D} = j\omega\epsilon'\mathbf{E} + (\sigma + \omega\epsilon'')\mathbf{E} \tag{2.26}$$

and the loss tangent is given by $\tan(\delta) = (\sigma + \omega\epsilon'')/\omega\epsilon'$. The power loss is given by

$$P_l = \frac{1}{2}\mathrm{Re}\int_S \mathbf{J} \cdot \mathbf{E}^* dS = \frac{\sigma + \omega\epsilon''}{2}\int_S \mathbf{E} \cdot \mathbf{E}^* dS \tag{2.27}$$

where the combined conductive and dielectric loss current is substituted. The total power is given by

$$P = \frac{1}{2}\mathrm{Re}\int_S \mathbf{E} \times \mathbf{H}^* \cdot \hat{z} dS = \frac{Y_w}{2}\mathrm{Re}\int_S \mathbf{E} \cdot \mathbf{E}^* dS \tag{2.28}$$

where Eq. (2.24) is substituted to simplify the result. Combining this with Eq. (2.25) the attenuation coefficient due to dielectric losses is found to be

$$\alpha = \frac{P_l}{2P} = \frac{\sigma + \omega\epsilon''}{2Y_w} \approx \frac{\omega\epsilon''}{2Y_w} \tag{2.29}$$

and it is assumed that the dielectric is non-conductive. In the conductors, the surface current is given by

$$Z_m = \frac{1+j}{\sigma\delta_s} \tag{2.30}$$

$$\delta_s = \frac{2}{\omega\mu\sigma} \tag{2.31}$$

$$\mathbf{E} = Z_m\mathbf{J_s} \tag{2.32}$$

which defines the effective surface impedance for a skin depth of δ_s in a good conductor. Following a similar approach to the dielectric losses, the conductive attenuation coefficient can be calculated

$$P_l = \frac{1}{2}\mathrm{Re}\int_S \mathbf{J}\cdot\mathbf{E}^* dS \tag{2.33}$$

$$= \frac{R_m}{2}\mathrm{Re}\int_S \mathbf{J_s}\cdot\mathbf{J_s}^* dS \tag{2.34}$$

$$= \frac{R_m}{2}\mathrm{Re}\int_S (\hat{n}\times\mathbf{H})\cdot(\hat{n}\times\mathbf{H}^*) dS \tag{2.35}$$

$$= \frac{R_m}{2}\mathrm{Re}\int_S \mathbf{H}\cdot\mathbf{H}^* dS \tag{2.36}$$

$$P = \frac{1}{2}\mathrm{Re}\int_S \mathbf{E}\times\mathbf{H}^*\cdot\hat{z} dS \tag{2.37}$$

$$= \frac{Z_w}{2}\mathrm{Re}\int_S (-\hat{z}\times\mathbf{E}^*\times\mathbf{H}^*\cdot\hat{z} dS \tag{2.38}$$

$$= \frac{Z_w}{2}\mathrm{Re}\int_S \mathbf{H}\cdot\mathbf{H}^* dS \tag{2.39}$$

$$\alpha = \frac{P_l}{2P} = \frac{R_m}{2Z_w} = \frac{1}{2Z_w\sigma\delta_s} \tag{2.40}$$

where the boundary condition for the surface current on the conductor has been used to relate $\mathbf{J}$ and $\mathbf{H}$. Combining the dielectric and conductive attenuation coefficients, the total attenuation coefficient is

$$\alpha \approx \sqrt{\frac{\omega\mu}{2\sigma}}\frac{1}{2Z_w} + \frac{\omega\epsilon''}{2Y_w} \tag{2.41}$$

For the circuit transmission line model, if a series loss, R, and a shunt loss, G, are added, then the propagation constant and characteristic impedance become

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{2.42a}$$

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}. \tag{2.42b}$$

After appropriate Taylor expansion under the assumption of low loss, the attenuation coefficient is

$$\alpha \approx \frac{R}{2Z_0} + \frac{G}{2Y_0} \tag{2.43}$$

This is an equivalent equation to Eq. (2.41), where $R = \sqrt{\omega\mu/2\sigma}$ and $G = \omega\epsilon''$. More importantly, it shows that for conductors the losses are proportional to the square root of the frequency, while for dielectrics the directly proportional to frequency. Furthermore, this illustrates that if conductor losses dominate, the attenuation coefficient is inversely proportional to characteristic impedance.

2.1.5 Causality of Dielectric Losses

Since the fields $\mathbf{E}$ and $\mathbf{D}$ are real, ϵ must also be real. Therefore $\epsilon(\omega) = \epsilon(-\omega)^*$ from the properties of the Fourier transform. It can further be shown that causality demands that $\epsilon(\omega)$ is analytic in the upper half ω plane. Therefore the Kramers-Kronig relations can be used to relate the real and imaginary parts of $\epsilon(\omega)$

$$\mathrm{Re}(\epsilon(\omega)) = 1 + \frac{1}{\pi}\mathcal{P}\int \frac{\mathrm{Im}(\epsilon(\omega'))}{\omega' - \omega} d\omega' \tag{2.44}$$

$$\mathrm{Im}(\epsilon(\omega)) = -\frac{1}{\pi}\mathcal{P}\int \frac{\mathrm{Re}(\epsilon(\omega') - 1)}{\omega' - \omega} d\omega' \tag{2.45}$$

where $\mathcal{P}$ denotes the Cauchy principal value. Therefore the loss and the permittivity are coupled together. An identical analysis applies to magnetic permeability and losses.

2.1.6 Discontinuities

At discontinuities in transmission, lines energy is stored that can be modeled with a capacitor, inductor, or resonant structure in cascade, where the equivalent model depends on whether the energy storing mechanism predominantly stores electric, magnetic, or a combination of electric and magnetic energies. From a fields perspective, consider two transmission lines of different geometries connected together. In the first transmission line a mode is excited that

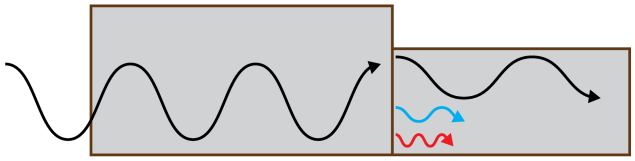

Figure 2.3: An illustration of the excitation of many modes at a geometry discontinuity in transmission lines. The other modes are excited in order to meet the boundary conditions at the interface, and in this illustration are evanescent (below cutoff) and do not propagate into the second waveguide.

propagates towards the second, which has a different set of fundamental orthogonal mode solutions. When the excited wave arrives at the discontinuity, an infinite linear combination of the modes of the second line are excited to satisfy the boundary conditions at the interface. A crude illustration is given in Fig. 2.3. Some of these modes may be evanescent (below cutoff) and store energy in the electric or magnetic field. Therefore discontinuities can be modeled as a discrete lumped circuit connecting two transmission lines together.

2.2 Properties and Special Cases of Field and Circuit Waves

There are several special cases and properties of the equations presented so far. A few interesting cases are shown here.

2.2.1 Distortionless Line

If the transmission line in Fig. 2.2 is modified to have series resistance, R, and shunt conductance, G, then the characteristic impedance and propagation constant are modified as Eq. (2.42). In general both γ and Z_0 will depend on frequency and introduce distortion due to dispersion. However if the elements are modified to have the property that $L/R = C/G$, then the equations can be simplified

$$Z_0 = \sqrt{\frac{R(1+j\omega L/R)}{G(1+j\omega C/G)}} \tag{2.46}$$

$$= \sqrt{\frac{R}{G}} = \sqrt{\frac{L}{C}} \tag{2.47}$$

$$\gamma = \sqrt{RG}(1+j\omega L/R) \tag{2.48}$$

$$= \sqrt{RG} + j\omega L\sqrt{\frac{G}{R}} \tag{2.49}$$

$$= \sqrt{RG} + j\omega\sqrt{LC} \tag{2.50}$$

so that both γ and Z_0 no longer depend on frequency.

2.2.2 Inductance of a Short Transmission Line

Typically, a short transmission line between a feed and its load is thought of as adding inductance. Furthermore, for estimating this inductance, a typical rule of thumb is based on the input impedance of a short-circuit stub. The input impedance of any transmission line is given in Eq. (2.11). When $Z_L = 0$ the input impedance is simply $jZ_0\tan(\beta l)$. For a lossless $50\,\Omega$ line that has a 5° electrical length, this corresponds to $j4.37\,\Omega$, or equivalently 140 pH at 5 GHz and 3 pH at 200 GHz. Why should this be related to the inductance seen when the transmission line is in series, since the equation models the circuit as a stub? Also, can this series stub ever be capacitive[7]?

To consider this case in more detail, the transmission line input impedance equation can be expanded to a power series in terms of $\beta l \ll 1$ in Eq. (2.54)

$$f(x) = \frac{a + jb\tan(x)}{b + ja\tan(x)} \tag{2.51}$$

$$f(0) = \frac{a}{b} \tag{2.52}$$

$$f'(0) = 1 - \frac{a^2}{b^2} \tag{2.53}$$

$$Z_{in}(\gamma l) \approx Z_L + jZ_0\beta l\left(1 - \frac{Z_L^2}{Z_0^2}\right) \tag{2.54}$$

As the length goes to zero, the equation correctly predicts that the input impedance is given by the load impedance. If $Z_L \ll Z_0$, then this equation is approximated further by $Z_L + jZ_0\beta l$, which indicates that the inductance is equal to $Z_0\beta l/\omega = Z_0 l/c$, proportional to

[7]The difference between capacitive and inductive is whether the transmission line stores more electric or magnetic reactive energy, respectively.

the characteristic impedance and length. These results are already predicted by the short-circuit stub. However, now we can determine when the line is inductive or capacitive. If $Z_L < Z_0$ the the line primarily stores inductive energy, since the imaginary component of the input impedance will be greater than zero, whereas when $Z_L > Z_0$ the line will instead store capacitive energy, having a negative imaginary component. When $Z_L = Z_0$ the line is matched and no reactive energy is stored.

2.2.3 Input Impedance of a Waveguide Below Cutoff Frequency

Consider a TE and TM waveguide with cutoff wave number k_c. The wave impedance is given by $Z_{TE} = k\eta/\beta$ and $Z_{TM} = \beta\eta/k$, while the propagation constant for both is $\beta = \sqrt{k^2 - k_c^2}$. Below the cutoff, $k < k_c$, so $\beta = j\sqrt{|k^2 - k_c^2|}$. Using the reflection coefficient form of Eq. (2.11), the exponent $-j2\beta l = 2l\sqrt{|k^2 - k_c^2|}$, which is increasing with l. For very large l, meaning the location the input impedance is evaluated at is very far from the load termination, the exponential dominates the factor of 1 and the input impedance can be simplified

$$Z_{in} = Z_0 \frac{1 + \rho_L e^{2l\sqrt{k^2-k_c^2}}}{1 - \rho_L e^{2l\sqrt{k^2-k_c^2}}} \approx Z_0 \frac{\rho_L e^{2l\sqrt{k^2-k_c^2}}}{-\rho_L e^{-j2\beta l}} = -Z_0. \tag{2.55}$$

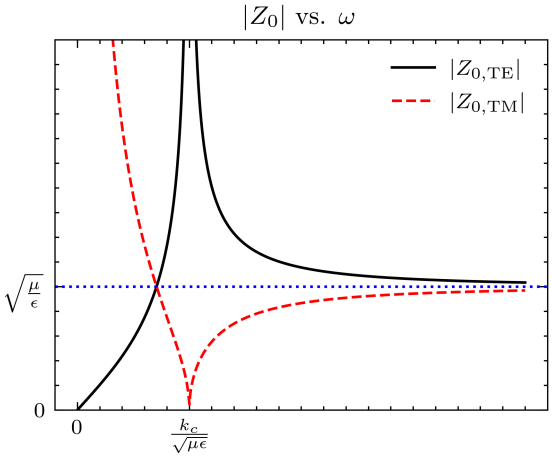

Figure 2.4: The magnitude of the wave impedance for TE and TM modes below and above cutoff.

Therefore, the input impedance below cutoff is the negative of the wave impedance. For TE and TM waves this is given by

$$Z_{TE} = \frac{k}{\beta}\eta = \begin{cases} \frac{-j}{\sqrt{k_c^2/k^2-1}} & k^2 < k_c^2 \\ \frac{1}{\sqrt{1-k_c^2/k^2}} & k^2 > k_c^2 \end{cases} \tag{2.56}$$

$$Z_{TM} = \frac{\beta}{k}\eta = \begin{cases} j\sqrt{k_c^2/k^2-1} & k^2 < k_c^2 \\ \sqrt{1-k_c^2/k^2} & k^2 > k_c^2 \end{cases} \tag{2.57}$$

and after taking the negative, the result is that TE modes are inductive below cutoff, while TM modes are capacitive below cutoff. The magnitude of the wave impedance is plotted in Fig. 2.4. At cutoff the impedance is an open circuit for TE modes and a short circuit for TM modes, signifying parallel and series resonances, respectively.

Chapter 3

Analysis of Chip-to-Package Interconnects

The standard chip-to-package interface is a coplanar ground-signal-ground (GSG) transition. However these transitions become very lossy above 100 GHz. To address the challenges in chip-to-package interface design the loss mechanisms in the traditional GSG transition are analyzed and possible alternatives are discussed. Much of this chapter follows from the work presented in [4, 3].

3.1 Limitations of GSG-type Transitions

The GSG transition structure is shown in Fig. 3.1a. In this structure, a grounded coplanar waveguide (GCPW) on the chip transitions to a microstrip line on the PCB or interposer through three flip-chip bumps. Due to the fanout required for bump pitch limitations, the return current travels a longer path than the signal current in this transition, which can be modeled with transmission lines. The cross section and transmission line model of the structure are shown in Fig. 3.1b and Fig. 3.1c, respectively. In this model, the bumps that carry current in the vertical direction are intentional transmission lines, shown in red, while the horizontal paths that the return currents must follow on the PCB and chip are parasitic transmission lines, shown in green.

Using this model, the input current into the intended transmission line must equal the input current into the parasitic line. Therefore the currents at either ends of the parasitic transmission line must be the same,

$$I_{in} = \frac{V_{2f} - V_{2r}}{Z_2} = \frac{V_{2f}e^{-j\theta_2} - V_{2r}e^{j\theta_2}}{Z_2} = I_{out} \tag{3.1}$$

where V_{xf} and V_{xr} are the voltage of the propagating waves in the forward and reverse directions respectively, Z_x is the characteristic impedance, τ_x is the time delay, and θ_x is the

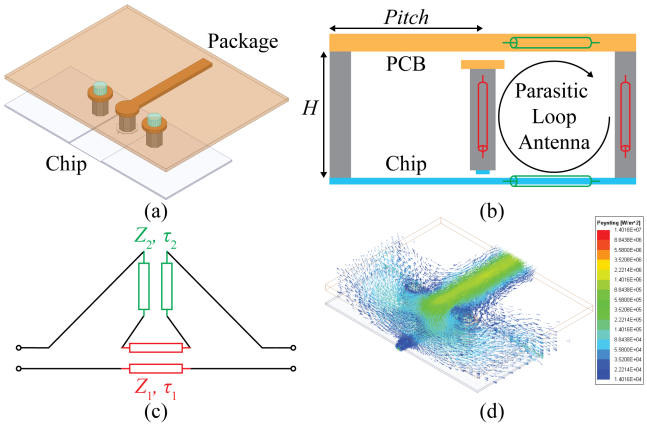

Figure 3.1: Conventional microstrip GSG transition (a) structure, (b) cross section, (c) schematic model and (d) simulated Poynting vector at 400 GHz.

electrical length of each transmission line. To satisfy this equation,

$$e^{j\theta_2} = -\frac{V_{2f}}{V_{2r}}. \tag{3.2}$$

From this, the standing wave ratio on the second transmission line does not depend on the load impedance. Moreover, at the frequency where $\theta_2 = \pi$,

$$I_{in} = \frac{V_{2f} - V_{2r}}{Z_2} \tag{3.3}$$

$$= \frac{V_{2f} + V_{2f}e^{-j\theta_2}}{Z_2} = 0 \tag{3.4}$$

which suggest that at the frequency

$$f_{\text{notch}} = \frac{1}{2\tau_2} \tag{3.5}$$

or an odd integer multiple of that, a notch in the transmission characteristic is expected. In other words, the timing mismatch between current and reverse current results in deep notches in the transition. The other transmission line may also exhibit similar notch behavior; however, for most practical transitions $\tau_1 < \tau_2$. This deep notch is easily seen when the length of the horizontal line is much greater than that of the vertical line, which is usually the case when small bumps are used on low manufacturing resolution PCBs.

The derivation above shows a parallel resonance in the circuit due to this propagation delay mismatch. However, since the circuit in Fig. 3.1c has no loss, it is reasonable to assume this resonance can be tuned out with ideal reactive components. In practice, however, this

strong resonance couples energy to lossy modes, which cannot be recovered by simple reactive tuning. Physically there are two dominant modes. First, since the parasitic transmission line is a 2-D parallel-plate transmission line, the signal can escape by coupling to the parallel-plate propagation mode at the metal-dielectric-metal stack in the transition region, as shown by Poynting vector simulations with Ansys HFSS in Fig. 3.1d. Second, a parasitic loop antenna is excited at the transition, as shown in Fig. 3.1b. This loop antenna is in resonance when the circumferential length of the loop is equal to the wavelength, assuming a short circuit on the chip. In terms of delays in the transmission line model

$$f_{rad} = \frac{1}{2\left(\tau_1 + \tau_2\right)}. \tag{3.6}$$

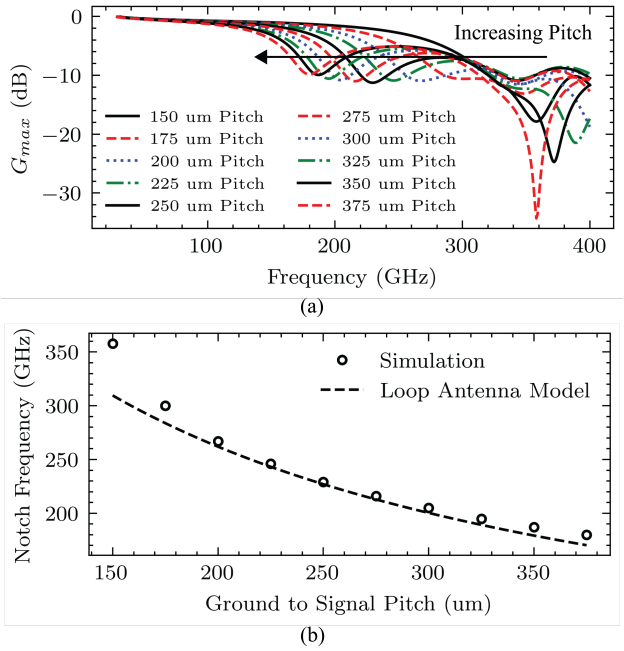

Figure 3.2: GSG simulation showing (a) notches in $G_{\max}$ for different pitches and (b) the loop antenna radiation model for the notch frequency.

The incoming signal near the radiation frequency is dissipated by coupling with parasitic surface wave modes and parallel plate modes, which can be seen by considering $G_{\max}$, the two-port maximum available power gain, in Fig. 3.2a. It can be observed that as the bump pitch increases, the first notch moves closer to the origin. Here, a bump pitch of 150 µm is considered, which is the minimum bump pitch offered by the technology used. In the simulation structure, the total distance between two footprints (H in Fig. 3.1b) is 125 µm. With a dielectric constant of 3.1 for the underfill material, Eq. (3.6) estimates the first notch

to be

$$f_{\text{notch}} \approx \frac{1}{2} \frac{\frac{3\times10^8\,\text{m/s}}{\sqrt{3.1}}}{125\,\mu\text{m} + \text{Pitch}} \tag{3.7}$$

where Pitch is shown in Fig. 3.1b. As evident by Fig. 3.2b, the radiation frequency of the loop antenna model agrees well with the simulated notch frequency.

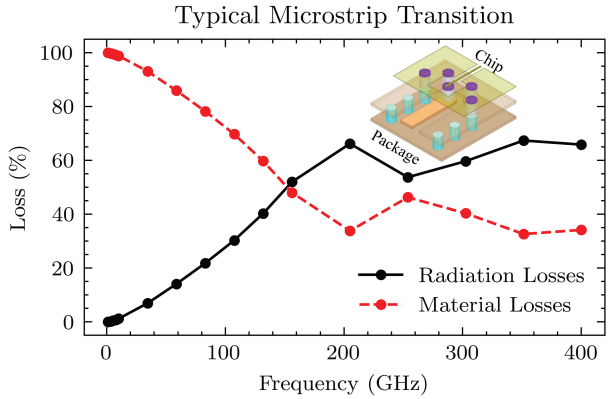

Figure 3.3: Comparison between metal and dielectric materials losses and radiation losses in a back-shielded microstrip to chip transition.

At higher frequencies, and lower wavelengths, these radiation losses become more critical when compared to metal and dielectric material losses. As an example, Fig. 3.3 demonstrates that even in a back-shielded microstrip-to-chip transition, radiation losses dominate beyond 140 GHz. Therefore alternative chip-to-package transition structures need to be considered.

3.2 Alternative Transition Structures

Since GSG transitions are not suitable for high frequencies, alternative structures are needed, as presented in Fig. 3.4. Additional ground bumps are added to the structure in Fig. 3.4a. This can partially reflect surface waves, however the reflected wave will still reach the chip boundary and be dissipated either by radiation or excitation of surface waves, shown in Fig. 3.4b. To combat this effect, two sets of ground bumps with positive and negative offsets in a rectangular shape can be used to make the transition shown in Fig. 3.4c. This is a much better approach in the lower frequency range because it can effectively reject forward and backward surface waves. However, as the distance between two ground bumps or the frequency increases, higher leakage is expected, as shown in Fig. 3.4d. Moreover, as the length of the PCB microstrip line increases over the chip region, this structure suffers from a higher degree of de-tuning and coupling with the silicon substrate.

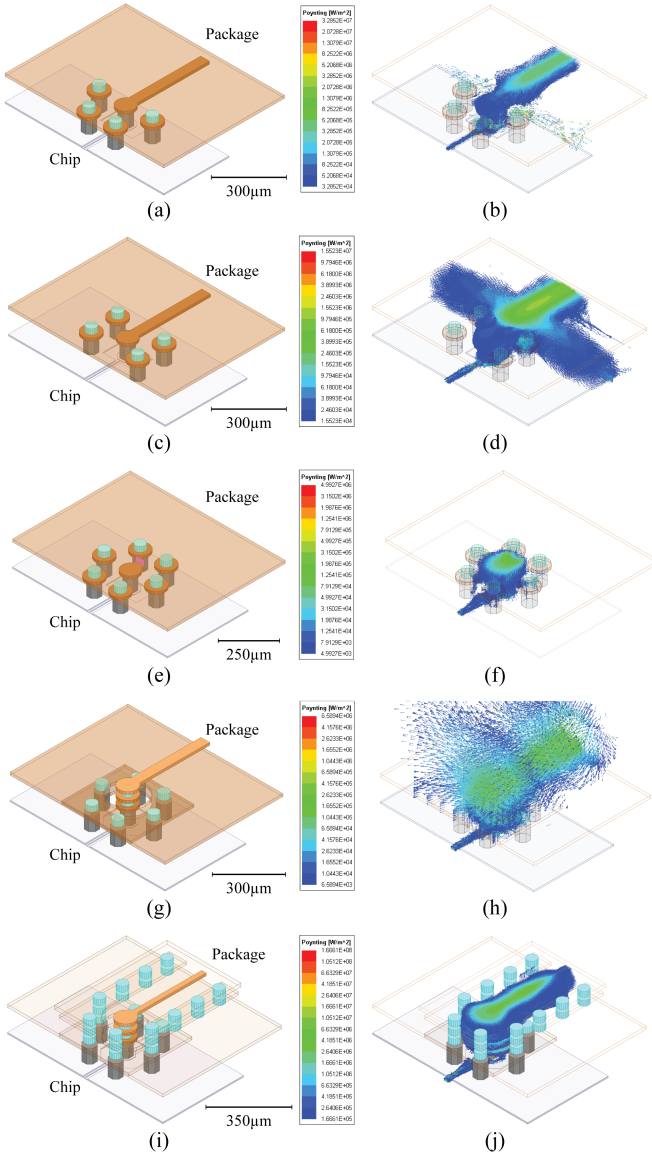

Figure 3.4: Possible modified transition structures: (a) half-shield, (c) rectangular shield, (e) full shield, (g) reverse microstrip, and (i) stripline. The Poynting vector for each corresponding transition at 400 GHz is shown in (b), (d), (f), (h), and (j).

The previous two structures indicate that the least leakage is expected when a full bump cage is formed with minimal spacing. To achieve this, the ground bumps are placed in a hexagonal pattern surrounding the signal, shown in Fig. 3.4e. The simulation results shown in Fig. 3.5 indicate that this structure achieves the best performance in terms of transition loss and notch frequency. Unfortunately, depending on the capabilities of the PCB manufacturer, this design may be impractical since the microstrip signal must be squeezed out of two ground bumps and their associated pads.

If the previous structure with a full shield is not feasible, a reverse microstrip could be used, as in Fig. 3.4g. In this structure, the ground plane is implemented on the second metal layer, while the signal resides on the third metal layer, and the stack-up is shown in Fig. 5.2d. This shields the signal line from coupling with other chip signals or the substrate. Furthermore the transition can be placed in the middle of the chip without interruption of other signals. While this transition structure minimizes leakage at the chip interface, signal losses occur at the inner via. Additionally, at higher frequencies the reverse microstrip can become a radiating element, as shown in Fig. 3.4h. This necessitates a large keep-out region above the signal line to reduce the parasitic coupling, making it less attractive.

To solve the problem with the previous structures, the microstrip line can be replaced by a stripline, as shown in Fig. 3.4i. The advantage of the stripline is that the signal is completely shielded from the external environment, mitigating the impact of variations in the shape of the underfill or the expansion of the silicon. The Poynting vector shown in Fig. 3.4j shows that the fields are contained by this structure.

The performance of the structures discussed is summarized in Fig. 3.5. While the microstrip with a full shield performs the best, the stripline design provides a practical signal escape and has superior performance among practical options.

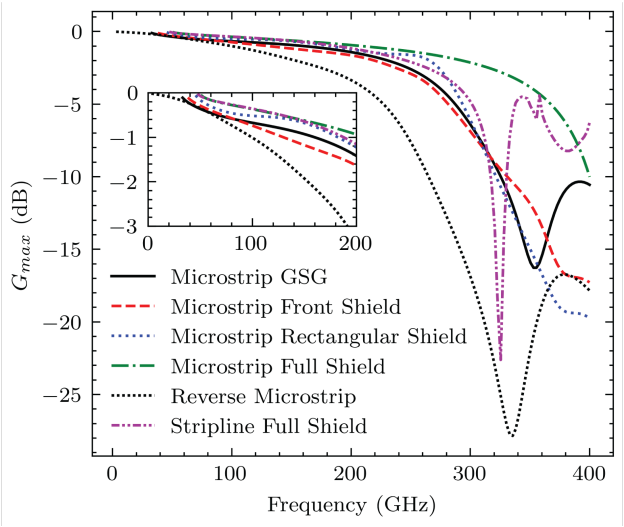

Figure 3.5: $G_{\max}$ of the transition structures presented in Fig. 3.4 from $0 - 400\,\text{GHz}$.

3.3 Limitations of the Stripline Structure

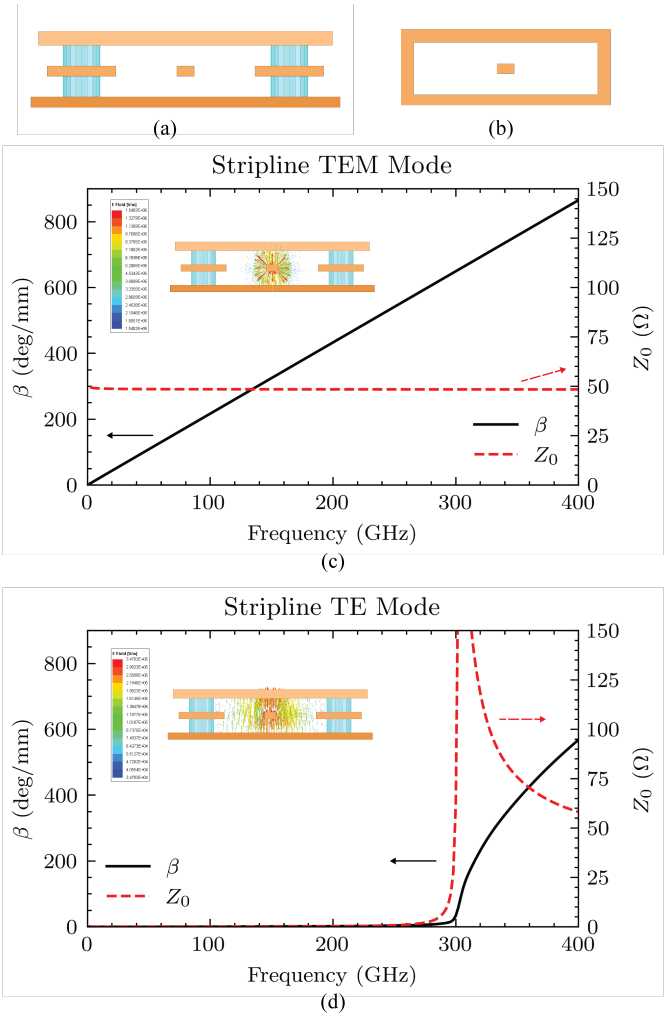

Figure 3.6: The (a) actual and (b) simplified cross section of a stripline substrate-integrated waveguide, as well as the (c) first TEM and (d) second TE propagation modes. The TE mode has a cutoff frequency of 300 GHz

The stripline design of Fig. 3.4i is the most promising solution for high frequencies. Therefore, it is desirable to explore this structure and investigate its possible limitations. The cross section of the stripline is shown in Fig. 3.6a. The first propagation mode of this structure in Fig. 3.6c is the intended TEM mode, which has no cut-off frequency. However, as the frequency increases, the metal cage around the line forms an effective waveguide, commonly called a substrate-integrated waveguide. Note that the discrete nature of microvias allows only TE propagation modes in the waveguide [13, 54]. The effective width of the waveguide

can be approximated by [6]

$$W_{eff} \approx W - \frac{D^2}{0.95P} \tag{3.8}$$

where W is the center-to-center spacing of the microvias on two sides, D is the diameter of the vias, and P is the spacing of the vias on the same side. The first TE mode of this effective waveguide is shown in Fig. 3.6d, with a cut-off frequency of 300 GHz. While an ideal straight stripline will perform smoothly in a simulation platform, any other structure may exhibit unpredictable performance above cut-off if the exact length of the transmission lines is not known at the design stage. Therefore, it should be ensured that the cut-off frequency of the TE mode is well above the highest frequency range of interest.

Considering Fig. 3.5, since lossy resonant modes result in notches in $G_{\max}$, an eigenmode solver of Ansys HFSS was used to study these loss mechanisms. The structure was modified to remove the access transmission lines. Among the numerous resonant modes, one of the modes corresponds to the cavity where the signal goes down through microvias in the shielded cage, which couples to the TE mode of the parasitic stripline waveguide. Depending on the reflection phase of the coupled wave, the resonant frequency of the loaded structure changes slightly. The actual reflection phase is unknown because this parasitic mode is not necessarily terminated with an actual load. Therefore, the waveguide is short-circuited at the end of the stripline, and several different lengths of the stripline are simulated.

The electric fields are shown in Fig. 3.7a,b. Note that the phase constant of the TE mode approaches zero near the cut-off frequency of the waveguide. Once the resonant frequency of the cavity is shifted down towards the cut-off frequency of the waveguide, the phase shift of the reflected wave becomes independent of the length, and therefore the notch in $G_{\max}$ will not cross the TE cut-off frequency, as shown in Fig. 3.7c. While other waveguide modes may be excited, they will occur above 300 GHz, and therefore have no effect on the performance of the transition below 200 GHz.

Below 300 GHz, or below the TE mode cut-off, the transition can be modeled with two capacitors and a series transmission line. This represents the pad capacitance, the effective delay, and the characteristic impedance of the microvias from the stripline opening to the chip, as shown in Fig. 3.8.

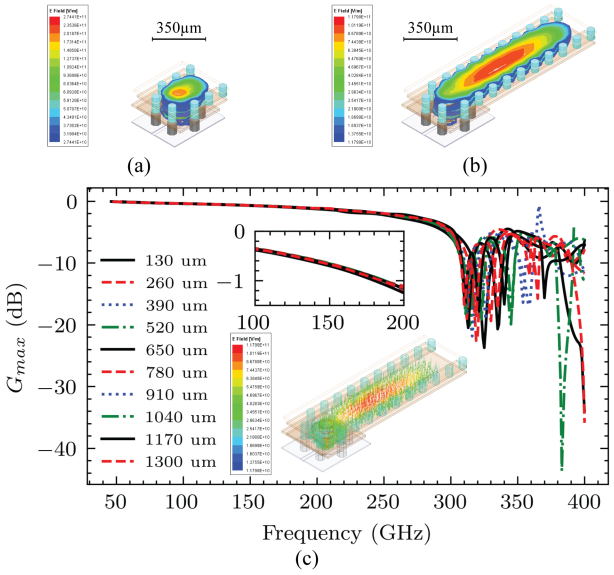

Figure 3.7: The results of eigenmode analysis of the stripline structure. The electric field magnitude for a (a) short line and (b) long line are shown. The resulting notch frequency of the stripline transition with varying stripline length is shown in (c). The variation below the TE mode cutoff frequency due to line length is negligible.

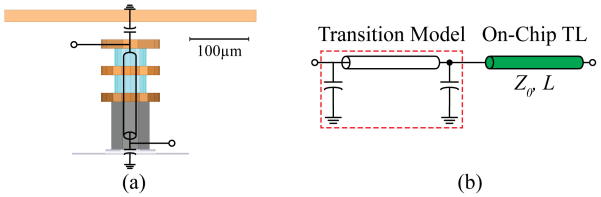

Figure 3.8: The (a) side view of the stripline transition and the (b) corresponding circuit model.

Chapter 4

Design of a 200 GHz Chip-to-Package Interconnect

In this chapter the design of a chip-to-package transition operating at 200 GHz is presented. The design procedure follows directly from the theory in Chapter 2 and the analysis in Chapter 3.

4.1 Technology and Stackup

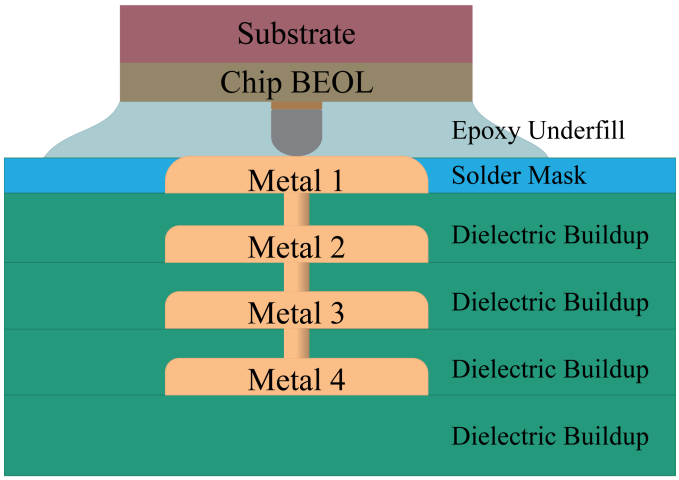

Figure 4.1: The material stackup used for the flip-chip-to-package transition.

For the transition, the chip is designed in a commercial 16 nm FinFET CMOS technology, and the package is made from an organic substrate material. Copper pillars are provided by the manufacturer for the interconnect. A diagram of the electrical stackup is shown in

Fig. 4.1. Material dispersion properties, such as dielectric permittivity and loss tangent are provided by the manufacturer from 0.1 GHz to 80 GHz. In HFSS this data is entered to form a causal, frequency-dependent dispersion model. The package metals are made from copper, and have been simulated with a 5 µm surface roughness nodule radius. This is implemented in HFSS with the Groiss surface roughness model as a finite conductivity impedance boundary condition.

4.2 Package Stripline Performance

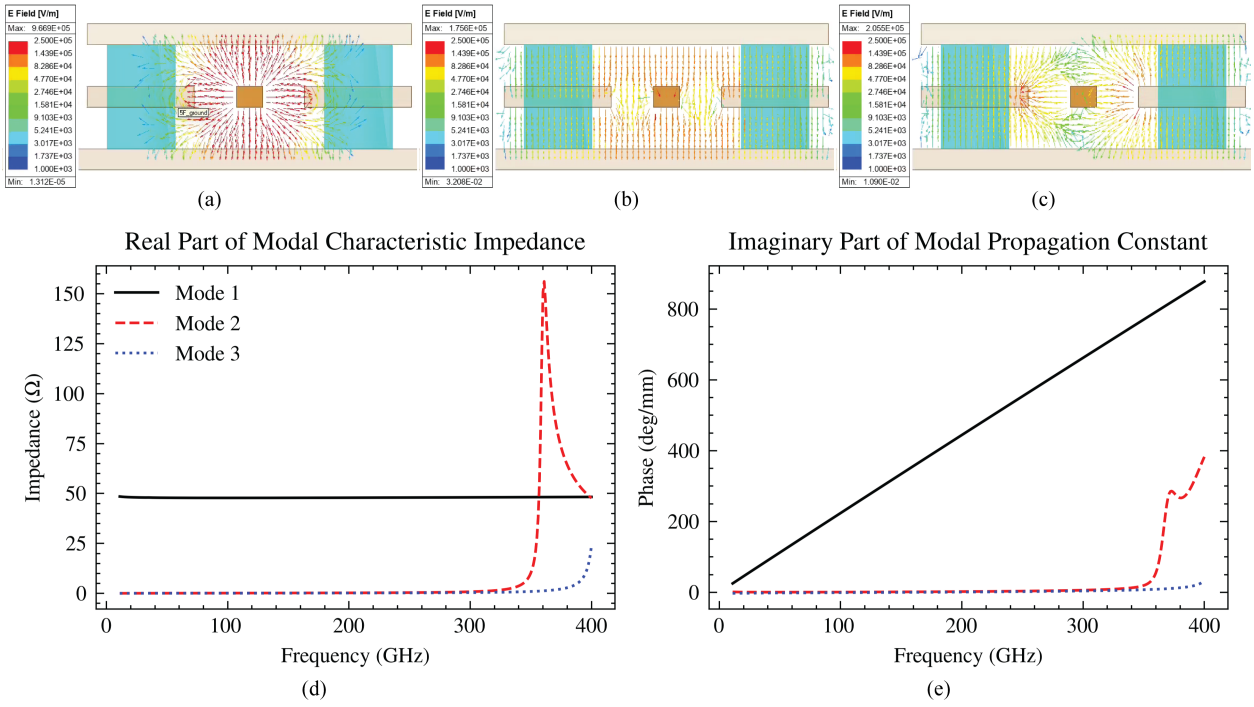

Figure 4.2: The modes of the on-package stripline transmission line. The electric field distribution is shown for (a) mode 1, the TEM mode, (b) mode 2, at TE mode, and (c) mode 3 a TE mode. (d) The real part of the characteristic impedance and (e) the imaginary part of the propagation constant.

Since the targeted frequency is 200 GHz, the transition is implemented using a stripline caged transition, as presented in Chapter 3. First, a 50 Ω stripline on the package is designed. The signal conductor resides on metal 2, and has a width of 19 µm, and the spacing from the conductor on the same layer is 30.12 µm. Stichting vias are placed to ensure a well-connected ground with a pitch of 112 µm.

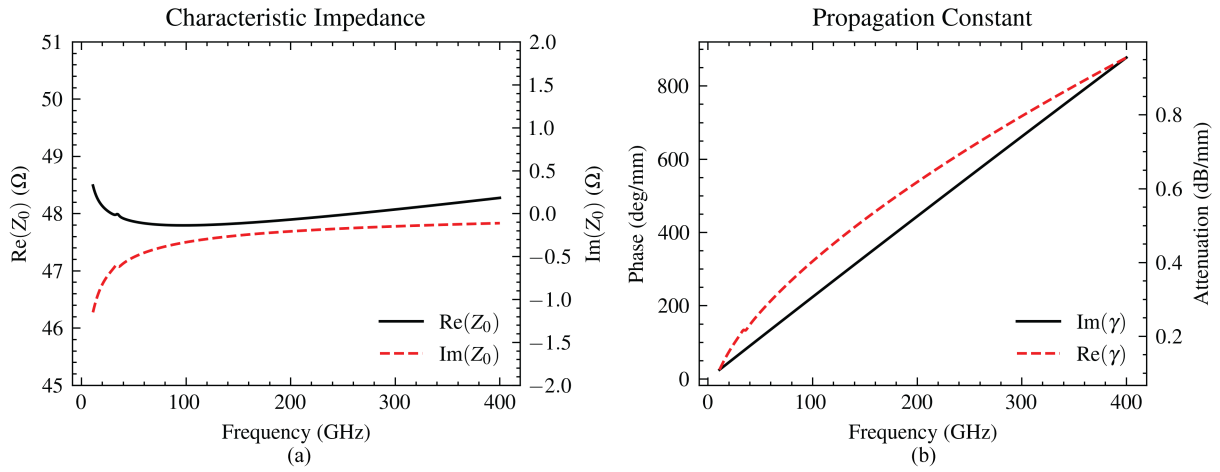

Figure 4.3: The performance of the TEM mode 1 of the designed stripline, including loss. (a) The real and imaginary part of the characteristic impedance and (b) the real and imaginary part of the propagation constant.

In order to ensure that the limitations of the second, TE mode do not occur, the first three modes of the stripline are analyzed. The electric field patterns are shown in Fig. 4.2a,b,c, where it can be seen that the first mode is a TEM mode, while the second and third are TE, which can be determined from comparing the characteristic impedance with the general result in Fig. 2.4. The characteristic impedance and propagation constant are shown in Fig. 4.2d and Fig. 4.2e, respectively. The second mode has a cut-off frequency of 361 GHz, well above the intended design frequency range.

The performance, including loss, of the first mode is shown in Fig. 4.3. This mode has a characteristic impedance of 48 Ω and an attenuation constant of 0.62 dB/mm. A stripline of 1.12 mm length is simulated, and the S-parameters are shown in Fig. 4.4.

4.3 Bump Transition and Impedance Matching

The bump locations are quantized by the manufacturer, so the design of the hexagonal shield pattern is limited. The geometry and dimensions of the bump pattern are shown in Fig. 4.5. The performance prior to implementing on-chip matching is shown in Fig. 4.6. As predicted in Chapter 3, a notch in G_{max} corresponds to the cutoff frequency of the TE mode of the stripline, at 361 GHz. Despite this, the transition shows potential for broadband operation with very low loss.

Matching at 200 GHz is performed with a series low impedance transmission line, as shown in Fig. 4.7. The low impedance line naturally transforms the impedance back to 50 Ω by rotating around a VSWR circle centered at the line characteristic impedance. It is

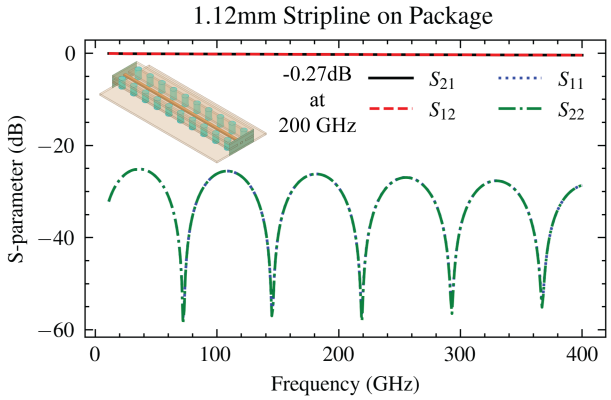

Figure 4.4: The S-parameters for a 1.12 mm stripline on the package.

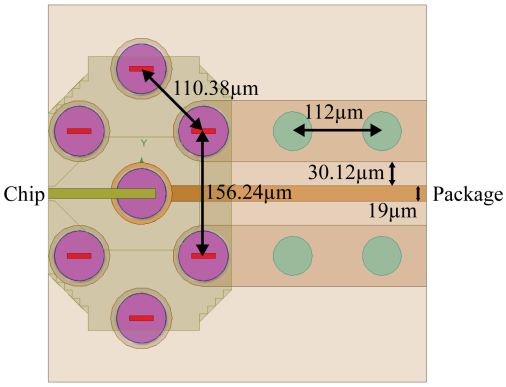

Figure 4.5: The geometry of the hexagonal shield transition, including key dimensions.

worth noting that in many practical cases the input impedance of the transition on-chip is very close to the desired impedance for matching to an on-chip amplifier, such as a PA. In this case further losses can be avoided by engineering the transition to present the optimum impedance to the amplifier. Nevertheless, a match to 50 Ω on-chip is implemented with a 28 Ω series transmission line that is electrically 32° at 200 GHz.

The low impedance line is implemented on chip as a grounded coplanar waveguide (GCPW). The signal conductor is 12 µm wide, with a horizontal gap of 3 µm. The electric field distribution of the GCPW is shown in Fig. 4.8a and the characteristic impedance and propagation constant are shown in Fig. 4.8b and Fig. 4.8c, respectively. The attenuation constant of the line is 1.6 dB/mm.

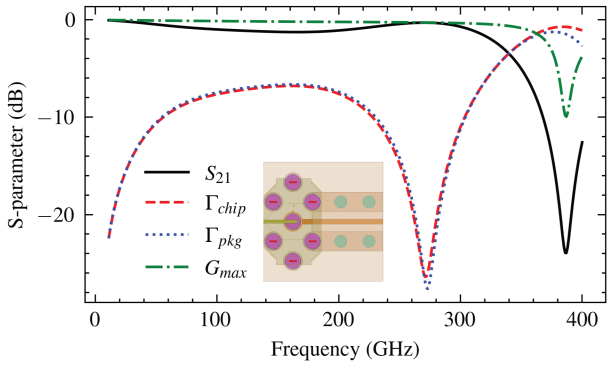

Figure 4.6: The performance of the transition prior to implementing the on-chip matching network. Γ_x indicates the reflection coefficient (S_{xx}) seen from x.

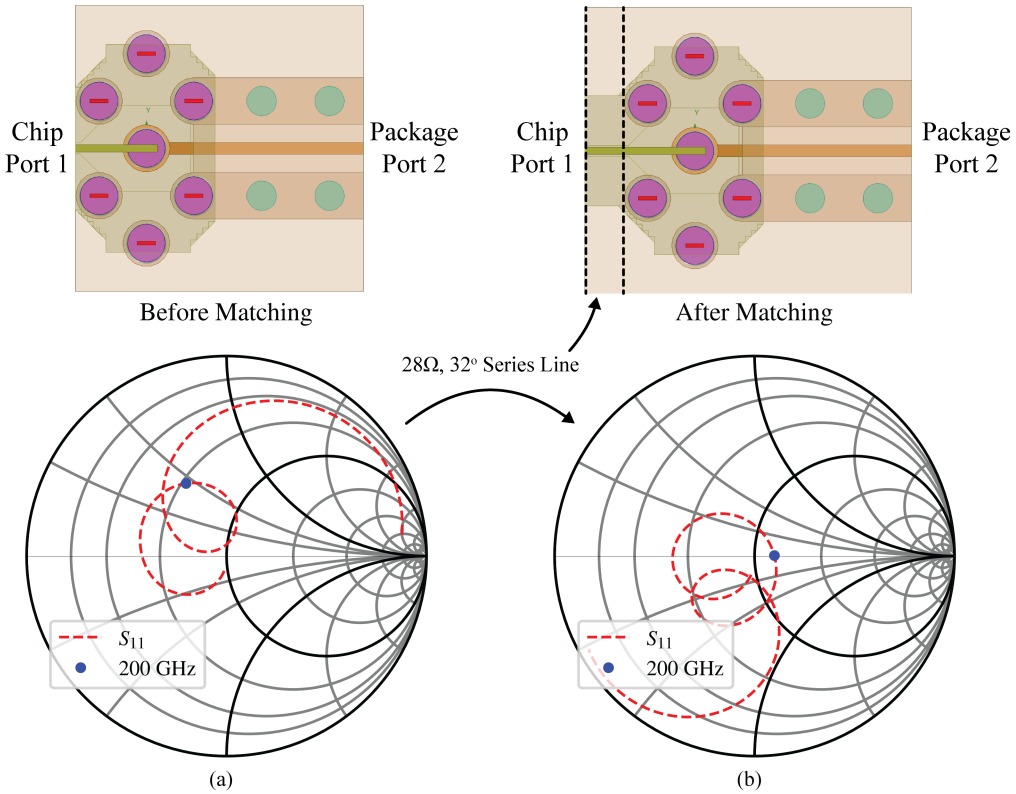

Figure 4.7: The on-chip input impedance (a) before and (b) after matching with a series low impedance line.

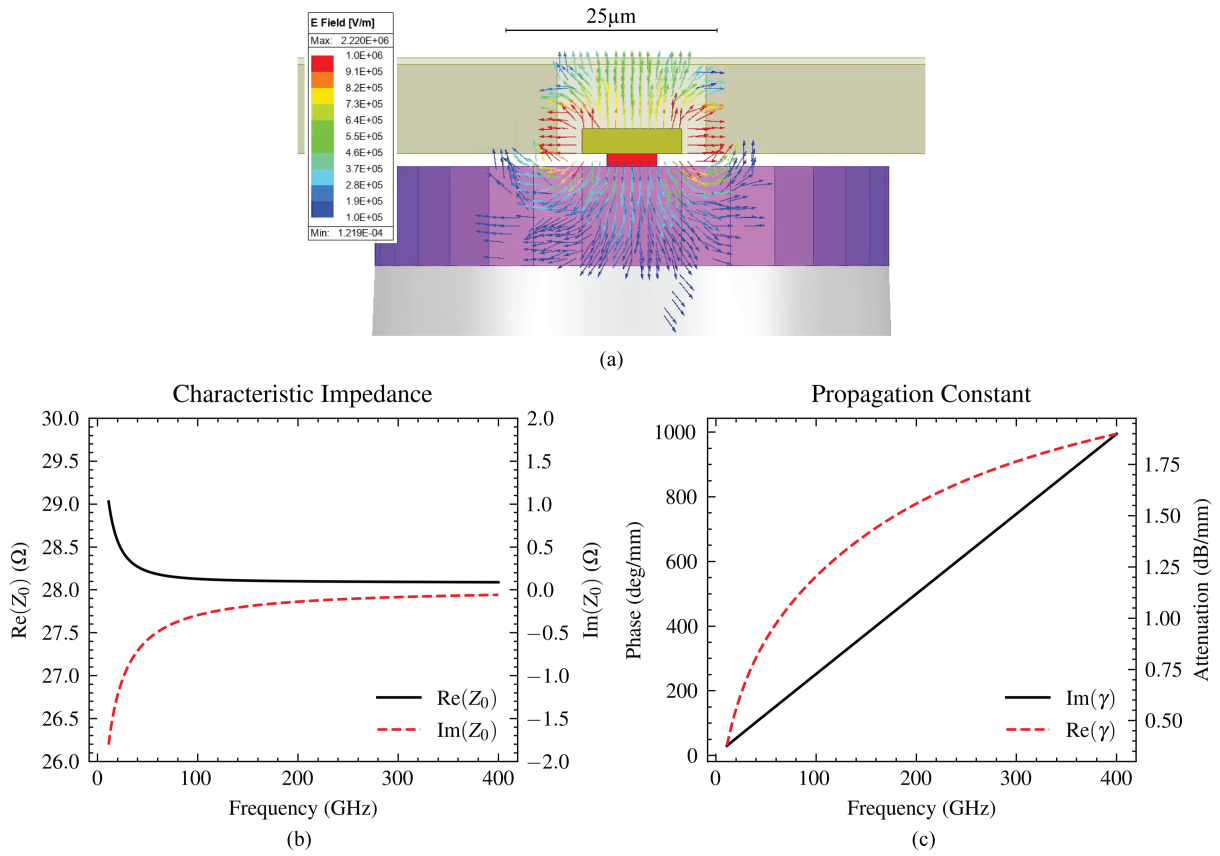

Figure 4.8: The (a) electric field pattern of the GCPW on-chip transmission line used for matching and its (b) characteristic impedance and (c) propagation constant.

4.4 Simulation Results

The final geometry is shown in Fig. 4.9a, and the performance is shown in Fig. 4.9b. Matching is implemented as an approximately 28 Ω line with an electrical length of 32° at 200 GHz. The transition has an insertion loss of 0.4 dB with a 3 dB bandwidth from DC to 339 GHz. The notch in $G_{\max}$ corresponds with the cutoff frequency of the TE mode of the stripline, as discussed in Section 3.3. The performance is summarized in Table 4.1.

4.5 Field Distribution and Substrate Shielding

The electric field distribution of the designed transition is shown in Fig. 4.10. The mode conversion from a TEM stripline mode on package to a GCPW mode on chip is clear from the E-field pattern. In the process significant energy is stored in the electric field at the

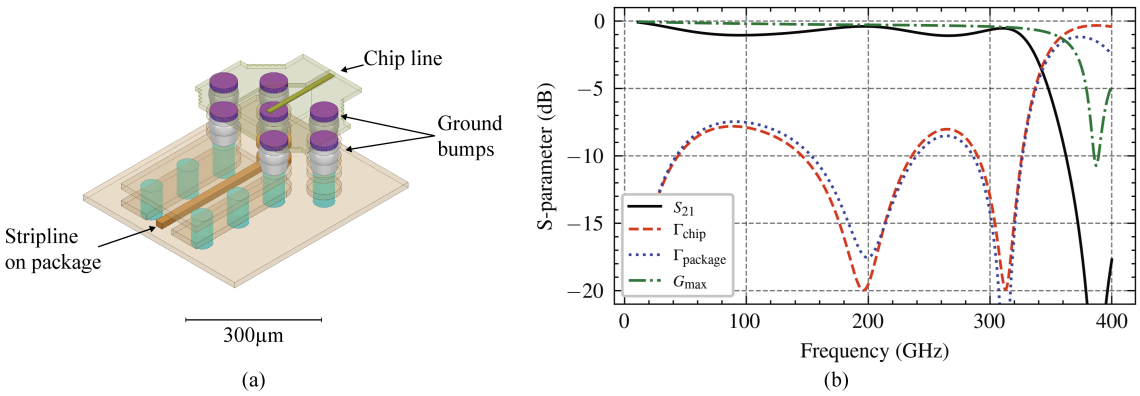

Figure 4.9: The (a) geometry of the transition from 16 nm FinFET CMOS technology to an organic substrate interposer and the (b) corresponding transition performance designed for 200 GHz. Γ_x indicates the reflection coefficient (S_{xx}) seen from x.

interfaces. This is dominant on the chip side of the transition primarily because the length scales are much smaller, so a given voltage drop results in a higher electric field intensity.

The performance shown in Fig. 4.6 and the reflection coefficient in Fig. 4.7a suggest that the broadband nature of the transition is primarily limited by parasitic capacitance in the transition, as modeled in Fig. 3.8b. This is primarily due to the capacitance to the chip's ground plane. It is therefore tempting to introduce cuts in the ground plane on-chip to reduce the coupling to the substrate.

The Poynting vector with and without the addition of ground plane cuts is shown in Fig. 4.11a and Fig. 4.11b, respectively. From the Poynting vectors it is clear that the addition of ground plane cuts leads to a significant coupling to the silicon substrate. This

Package Technology	Organic substrate
Chip Technology	16 nm FinFET CMOS
Interconnect	Copper Pillar
Bump Height(µm)	66
Bump Diameter (µm)	62
Bump Pitch (µm)	156
Center Frequency (GHz)	200
Insertion Loss (dB)	0.4
3 dB Bandwidth (GHz)	339
Matching 10 dB Bandwidth (GHz)	90

Table 4.1: A performance summary of the designed transition, shown in Fig. 4.9.

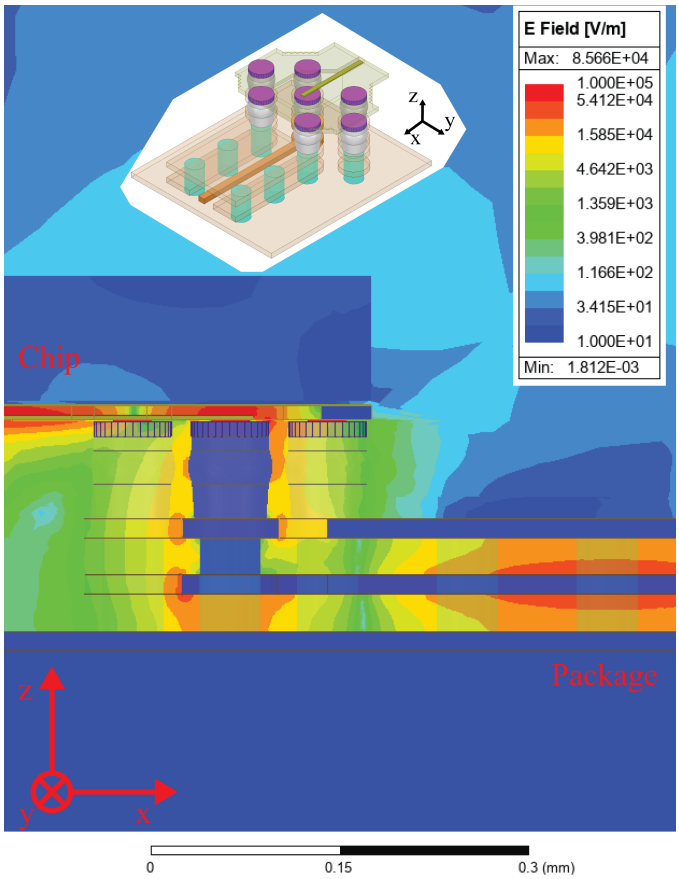

Figure 4.10: A cross-section view of the magnitude of the electric field of the transition.

energy coupling can excite surface waves along the substrate surface, as well as rectangular cavity modes that depend strongly on the size of the substrate. Since the silicon substrate is very lossy, the excitation leads to a significant degradation in G_{max}, shown in Fig. 4.11c. Therefore the ground plane remains solid to effectively shield the coupling to the substrate, at the cost of matching bandwidth from DC.

4.6 Robustness with Respect to Cross-talk

Another advantage to the stripline transition is that the ground shield provides strong isolation of the signals. In an array application it would be desirable to have as small a footprint per transition as possible, since the transition already requires seven bumps. This footprint can be reduced by sharing an adjacent ground bump, but this may come at the cost of transition-to-transition coupling.

This effect is studied in Fig. 4.12. Four transitions are aligned in a row, while sharing a

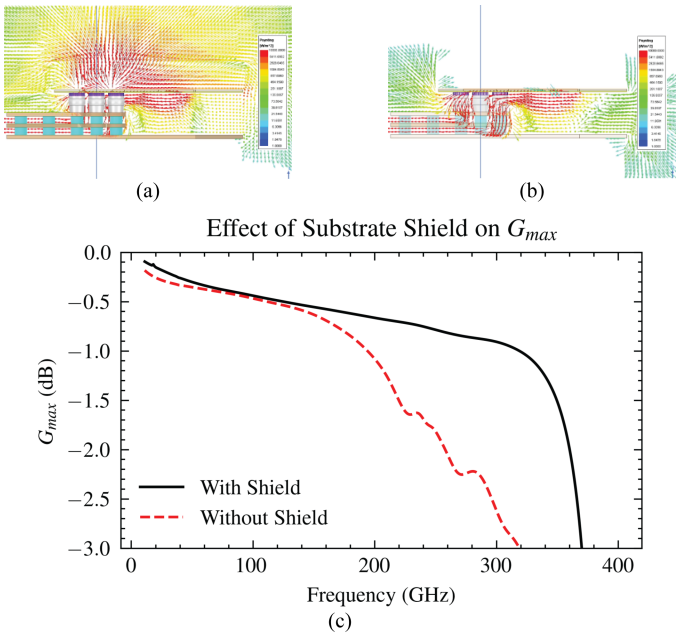

Figure 4.11: The Poynting vector (a) without and (b) with substrate shield on-chip shows that coupling of energy into the Silicon substrate is very high without shielding. This directly translates to losses, shown in the plot of G_{max} for the two cases in (c).

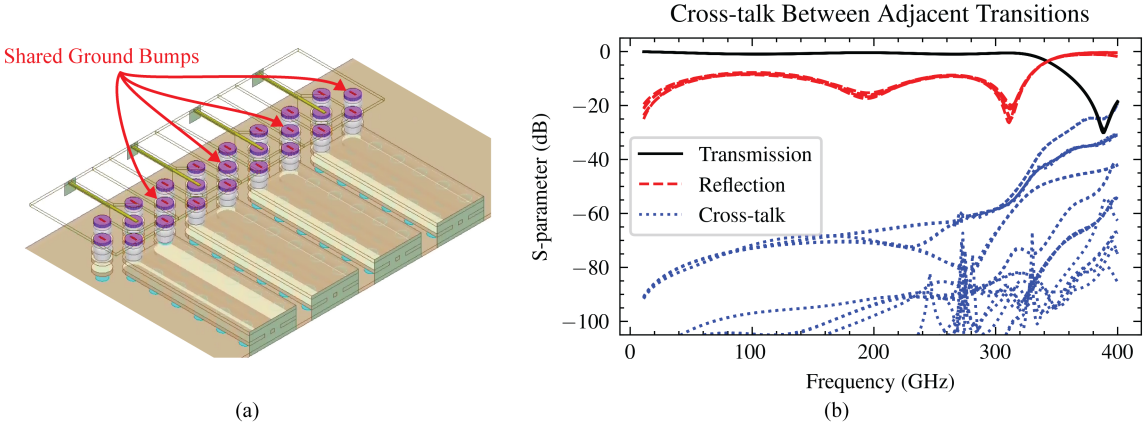

Figure 4.12: (a) The geometry used to test the cross-talk, with four transitions abutted next to each other and sharing a ground bump. (b) The resulting cross-talk between transitions, showing better than −60 dB rejection.

single ground bump between them, shown in Fig. 4.12a. The full 8-port S-parameters are simulated in HFSS, and the results in Fig. 4.12b show that cross-talk is limited to −60 dB up to 300 GHz.

4.7 Loss Mechanisms

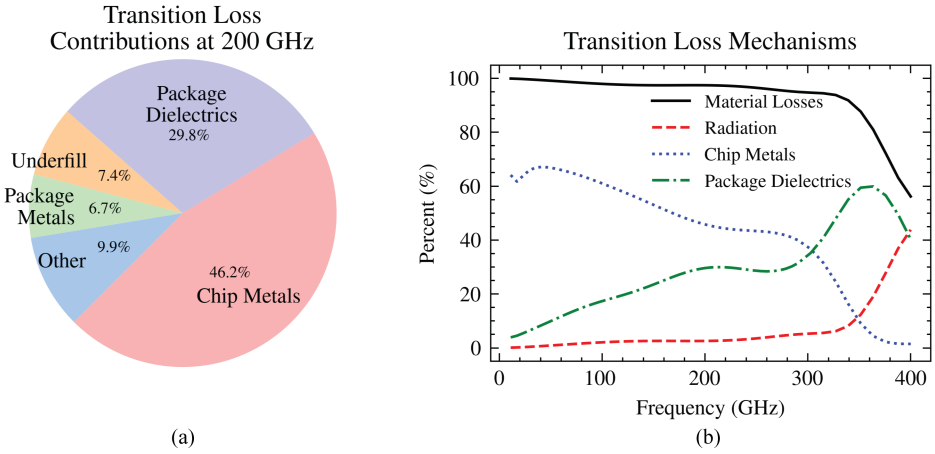

Figure 4.13: (a) A breakdown of the contributions to losses at 200 GHz, where radiation losses are included, but negligible, while reflection losses are not included, corresponding to G_{max}. (b) The key loss contributions across frequency.

While the transition performance is very good, it is still valuable to understand the dominant sources of loss. A summary of the key contributors is shown in Fig. 4.13a at 200 GHz. The summary includes radiation losses, which are negligibly small due to the effective shielding in the stripline structure, unlike the transition described in Fig. 3.3. However the breakdown does not include reflection losses at either port. Therefore, the summary corresponds to the case of G_{max}, where it is assumed that reflections can be matched ideally at a given frequency. It is separately found that the value of G_{max} calculated from these losses matches the value of G_{max} calculated directly from S-parameters.

The chip metals contribute the highest loss to the system. This is primarily due to the high resistance of small feature sizes on chip. Furthermore, the low impedance transmission line is particularly lossy since resistive, conductor losses dominate dielectric losses and the attenuation coefficient is inversely proportional to the characteristic impedance in Eq. (2.42) and Eq. (2.43). The next highest loss contributor is the dielectric material on the package, indicating the importance of including dielectric losses in simulations. Since this loss varies with frequency, it will contribute to dispersion. Fig. 4.13b shows the loss contributions versus frequency.

While the low impedance transmission line matching is compact and simple, the breakdown of loss suggests that if it is a limiting factor in a transition design other matching topologies should be considered.

Chapter 5

Calibration and Measurements of a 140 GHz Chip-to-Package Interconnect

Theory and simulations have been presented in the preceding chapters. In this chapter, the measurements of a chip-to-package transition are presented. The calibration procedure and error mechanisms are discussed, and steps for more accurate measurements are described. Despite the error sources in the measurements, a good correlation with simulations is established.

5.1 A 140 GHz Chip-to-Package Transition

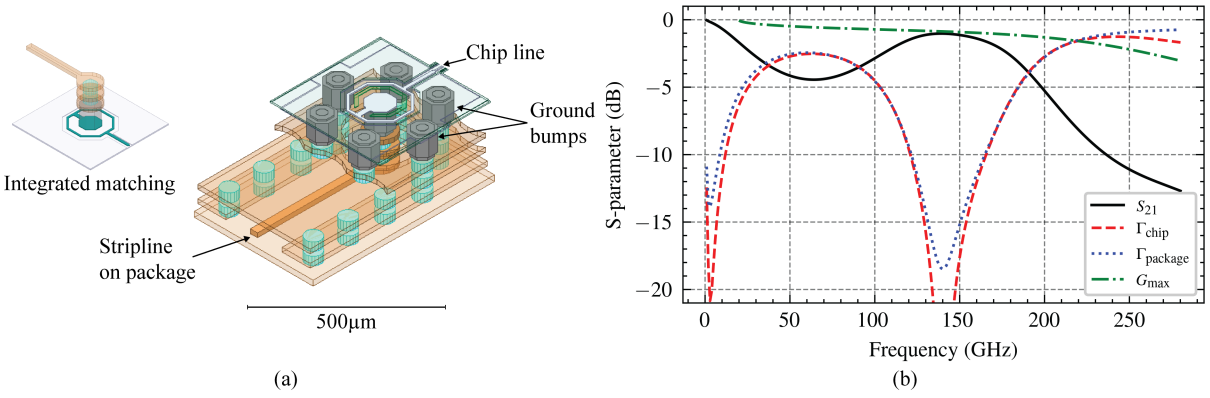

Figure 5.1: The (a) geometry of the transition from 28 nm Bulk CMOS technology to an organic substrate interposer, including integrated symmetrical transmission line matching, and the (b) corresponding transition performance designed for 140 GHz. Γ_x indicates the reflection coefficient (S_{xx}) seen from x.

At the time of writing, the design presented in Chapter 4 has not been fabricated for measurements. In [3, 4], the design of a chip-to-package transition was implemented in a 28 nm Bulk CMOS technology and an organic substrate interposer. The design is similar to what was presented in Chapter 4. To utilize the full silicon area, a matching network is implemented within the ground cage. It consists of two symmetrical transmission lines shown in Fig. 5.1a, whose characteristic impedance and length are calculated to obtain a matched impedance at 140 GHz. Integrating the matching network within the hexagonal ground cage saves area, but at the cost of bandwidth. The performance is shown in Fig. 5.1b, and the transition achieves 1.03 dB loss with a 85 GHz 3 dB bandwidth.

5.2 Measurement Setup

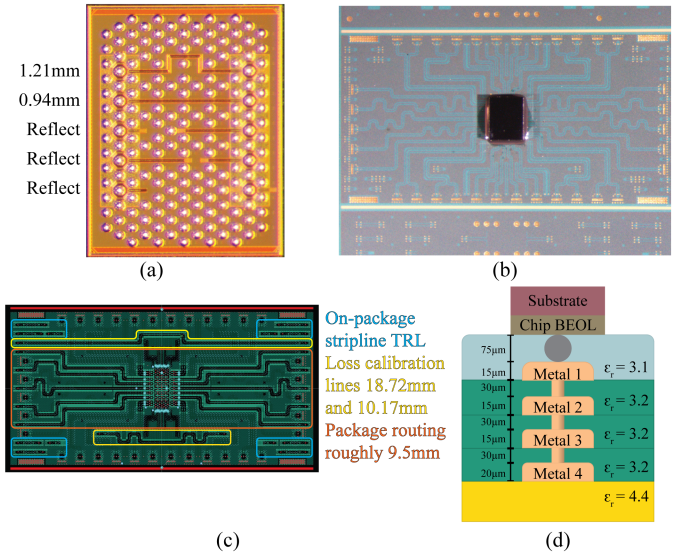

Figure 5.2: The (a) 28 nm Bulk CMOS die photo, (b) corresponding assembled organic substrate package, as well as (c) the layout of metal three, where the stripline signal traces reside. Reflect denotes either a short or open, following TRL naming conventions. The relevant part of the stackup is shown in (d). BEOL stands for "back end of the line."

Measurements were carried out on a test structure consisting of back-to-back transitions of the design in Fig. 5.1. A fabricated die photo is shown in Fig. 5.2a. Five test structures include a 1.21 mm line (that takes three 90° bends); a 0.94 mm through line; and six reflecting structures of various types are arranged in a row to form the remaining three test structures. The top four of the reflect structures are short circuits at the end of the transmission line, while the lowest left reflect is an open circuit and the lowest right is a reversed short circuit.

The assembled organic substrate package is shown in Fig. 5.2b. The package transmission lines are striplines which reside underneath the top metal, so the layout is also included in Fig. 5.2c. The five on-chip test structures in Fig. 5.2a are routed with roughly 9.5 mm meandering lines that fan out to enable on-package probing. This was done primarily to enable compatibility with other test chips in another project. In addition to the fan-out lines from the chip, the package contains custom stripline through-reflect-line (TRL) calibration standards that can be used to remove the effect of the on-package pads. Finally, two long meandering lines of 10.17 mm and 18.72 mm are included to aid the calibration further.

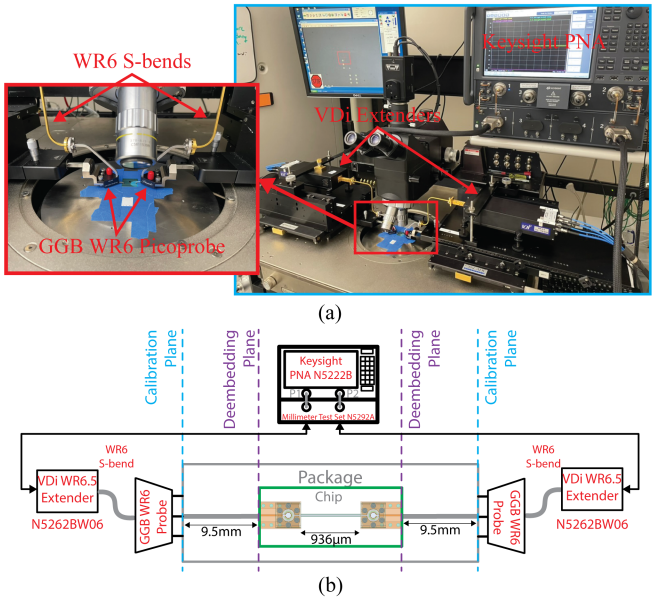

Figure 5.3: The (a) chip-to-package transition measurement setup and (b) the schematic diagram of the setup. The initial calibration plane from on-package TRL is shown, as well as the reference plane after de-embedding the package feed lines.

The measurement setup is shown in Fig. 5.3. S-parameters of the various test structures are measured using a Keysight N5222B PNA, which measures up to 26.5 GHz. Measurement of the *D*-band from 110 GHz to 170 GHz is enabled with the combination of the Keysight Millimeter Test Set N5292A and two VDi WR6.5 N5262BW06 Extenders. The WR6[1] waveguide output of the extender is connected to a WR6 S-bend and GGB WR6 Picoprobe for landing on the package.

[1]WR6 is the rectangular waveguide standard that corresponds to the *D*-band. "WR" stands for "waveguide, rectangular" and the numerical value represents the inner width of the waveguide in hundredths of an inch. WR6 corresponds to an inner width of 0.06 in, or 1.52 mm, so the cutoff frequency for a half-wavelength mode is 3×10^8 m/s/2 $\times$ 1.52 mm = 98.7 GHz.

The desired device under test (DUT) is a single chip-to-package transition, however it is not possible to setup calibration planes at the input and output of a single transition with the described setup, since the chip is flipped and inaccessible. Therefore, a combination of calibration and de-embedding is used.

5.3 Types of Calibration and Algorithms

For a given test setup and frequency, the choice of calibration procedure can significantly impact the accuracy of the measured data. In the measurement setup shown in Fig. 5.3, a complicated set of calibrations must be performed, in addition to further de-embedding. The entire equipment setup, up to and including the probe tips must be calibrated properly, which can be done with manufactured calibration standards. During wafer/package probing, calibration can be sensitive to the repeatability of landing the probe tips with the same positioning and over-travel [41]. The package pad landing must also be calibrated, but this must be done with the custom calibration standards on the package, since the pad is a custom structure. This calibration must be designed carefully to minimize measurement inaccuracies.

There are several types of calibration procedures that can be used to accurately measure S-parameters on a vector network analyzer (VNA). Some common types include short-open-load-through (SOLT), through-reflect-line (TRL), and line-reflect-reflect-match (LRRM), to name a few [37]. SOLT calibration is very accurate when the calibration standards are known. However, as the frequency increases, creating a well characterized short or open standard becomes very difficult. This can be corrected for if a manufactured standard is used, but is difficult to characterize when developing a custom calibration standard.

TRL calibration is one of the most popular calibration standards because it directly measures the traveling waves fundamental to transmission lines [53]. The calibration algorithm takes advantage of the field inversion of quarter-wavelength lines to create a set of independent solutions for the error terms in the model. Since the line is quarter-wavelength at only one frequency, the bandwidth is limited. In practice a rule of thumb is the usable bandwidth is the frequency range where the electrical length of the quarter-wavelength line remains between 20° and 160°[2]. Unlike the SOLT standard, TRL does not have a defined load. After calibration the S-parameters will be referenced to the characteristic impedance of the line, so this value should be known.

One modification to the TRL procedure is to include a defined load resistance so that the characteristic impedance of the line does not need to be known. This is done in LRRM. Some calibration algorithms can also account for parasitics in the load resistance. Another modification is to add multiple lines to increase the bandwidth beyond the quarter-wavelength limit. A different line can be used for different measurement bands. However, the accuracy can be further improved with the NIST multiline TRL algorithm [12]. Instead of using a

[2]This rule of thumb comes from the need to invert $\sin(\theta)$ in the calibration equations. In reality, the usable range is continuous and set by the numerical error that can be tolerated.

fixed through standard as a phase reference, the NIST multiline algorithm, for each frequency, picks the line that gives the lowest phase difference between it and all other lines to use as the common reference line. By reducing the phase differences the numerical accuracy is improved.

5.4 Calibration Procedure and Measurement Results

Different calibration configurations were tried and compared to give the best result and numerical accuracy. After an initial calibration, the short and open reflect standards on the custom package are shown in Fig. 5.4. While the short is a good approximation of a short ($S_{11} \approx -1$), the open has substantial capacitance ($S_{11} \approx -1j$). When the open standard was used in the calibration algorithm, the results were not numerically stable, so it was omitted.

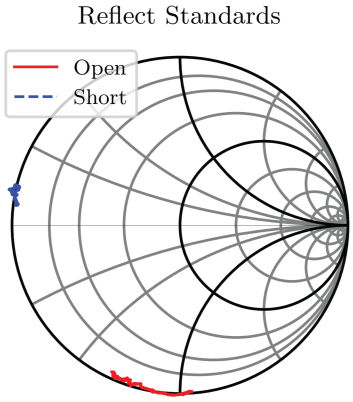

Figure 5.4: The measured S_{11} of the short and open standard on the custom package in the D-band. The short is close to $S_{11} = -1$, but the open is closer to $S_{11} = -1j$. This indicates a substantial parasitic capacitance, so the open is not used in the calibration.

Since TRL calibration relies on transmission line wave equations, the lines used for calibration must be straight. This is because the assumption used in solving for transmission lines is that a constant cross-section is maintained across the whole line. Any bend in the lines corrupts this fundamental assumption and the calibration is no longer possible. For this reason, the 18.72 mm and 10.17 mm lines cannot be used in the calibration.

In each calibration two steps are performed: 1) an initial calibration with the PNA and 2) an offline calibration with scikit-rf [2] to further move the reference plane. A comparison is made between performing initial calibration to the WR6 waveguide with a WR6 calibration kit, the probe tips with a manufactured calibration substrate, and direct calibration with the custom package TRL. Each case is followed by the offline calibration. Directly using the custom TRL package provides the best numerical stability and continuity across frequency.

After TRL calibration is performed with the calibration on the package, the reference plane is at the feeding point for the 9.5 mm lines, as shown in Fig. 5.3b. The resulting S-parameters are referenced to the characteristic impedance of the measured striplines, since there is no well-defined load present on the package. Therefore the propagation constant of the lines can easily be extracted from $S_{21} = e^{-\gamma l}$ of the long 18.72 mm and 10.17 mm lines on the package. The average propagation constant of the two lines is shown in Fig. 5.5. The measured transmission lines have slightly more loss than simulated.

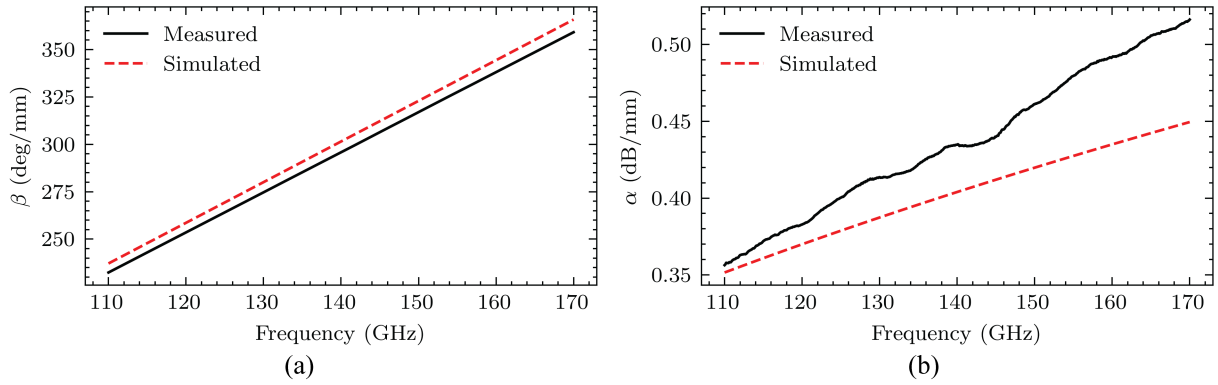

Figure 5.5: The measured propagation constant $\gamma = \alpha + j\beta$, where (a) β is in deg/mm and (b) α is in dB/mm.

With the package transmission line properties determined, 9.5 mm of line is de-embedded from the on-chip back-to-back transition structure. This can be done using the measured propagation constant since the S-parameters are referenced to the line characteristic impedance. The de-embedded S-parameters are shown in Fig. 5.6, where Fig. 5.6a shows the results of the straight 0.94 mm line on the chip, and Fig. 5.6b shows the results of the 1.2 mm line on the chip. The results correspond well with simulations, which model surface roughness with the Huray parameters of a 0.25 µm nodule radius and a surface area ratio of 4 [19]. The deviation in reflection coefficient corresponds to $\lambda/2$ every 9 GHz, which for $\epsilon_r = 3.1$ corresponds to the line length de-embedded from the package measurements. Therefore, the errors in reflection coefficient are primarily attributed to the errors in the de-embedding procedure. The total loss of the back-to-back through structure is 5.5 dB, which includes the 0.94 mm line on-chip. This compares well with the simulated value of 5.2 dB.

Since the 1.21 mm on-chip transmission line contains bends and the feed networks are not identical for each on-chip structure, calibration to the on-chip reference planes is not possible. However, the approximate loss of the on-chip transmission line can be estimated by taking the difference between the losses of the 1.21 mm structure and the 0.94 mm structure. After normalizing by the line lengths, the on-chip transmission line loss is found to be approximately 2.7 dB/mm. Using this loss, the estimated loss of a single transition is found

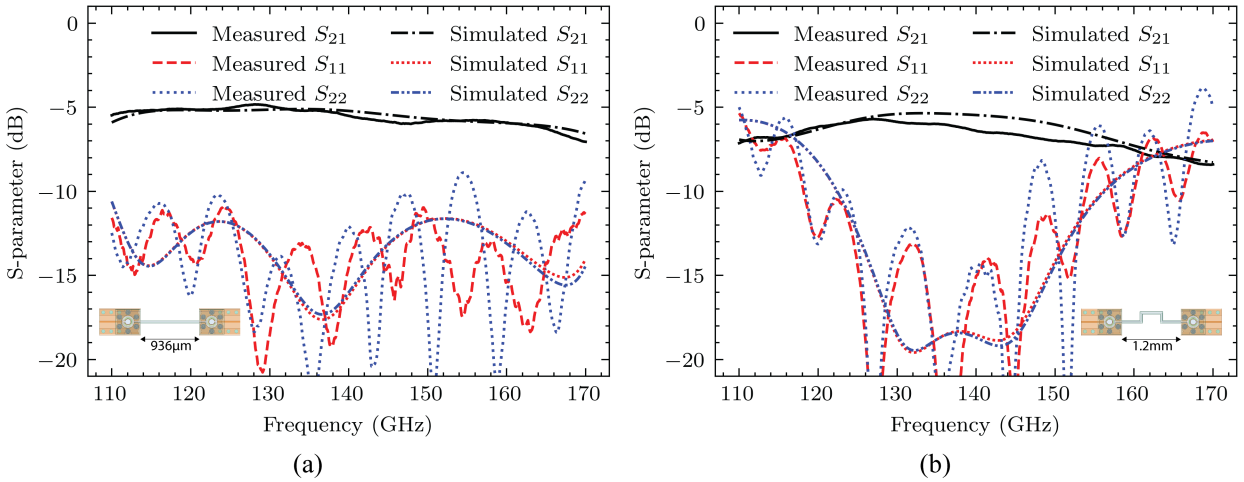

Figure 5.6: The measurement results of the back-to-back transition with a 1.2 mm on-chip transmission line between transitions.

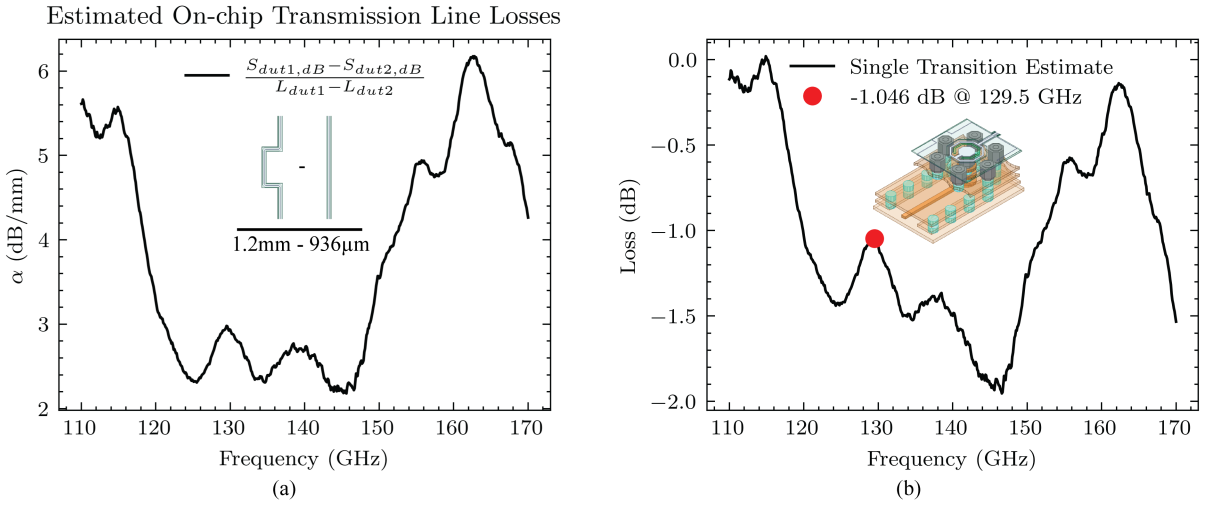

Figure 5.7: (a) The estimated on-chip transmission line loss. (b) The estimated loss of a single chip-to-package transition. These results are only valid within the bandwidth, as the calculation does not consider the impact of reflection losses.

to be 1.04 dB, again very close to simulations. These results are shown in Fig. 5.7, where the analysis is only valid in the bandwidth so that reflection losses do not contribute.

A summary of the measured performance is given in Table 5.1, and measurements match with simulations well.

Package Technology	ABF GL102
Chip Technology	28 nm Bulk CMOS
Interconnect	Solder Bump
Bump Height(µm)	75
Bump Diameter (µm)	80
Bump Pitch (µm)	150
Center Frequency (GHz)	140
Insertion Loss (dB)	1.04
3 dB Bandwidth (GHz)	85
Matching 10 dB Bandwidth (GHz)	43

Table 5.1: A performance summary of the measured transition, shown in Fig. 5.1.

Chapter 6

Conclusion

6.1 Conclusions

A comparison of the presented work with published literature is shown in Table 6.1. Both the designed transition in Chapter 4 and the measured transition presented Chapter 5 demonstrate the best performance among low-cost technologies in terms of insertion loss and bandwidth.

6.2 Future Work

Although the measurement results match well with the simulations, the accuracy of the measurements could be improved. The biggest source of error in the measurements is the de-embedding procedure. A new package could be designed for the sole purpose of measuring these transitions so that the feed lines are negligible length and identical. With the addition of a load resistor on the package, the characteristic impedance of the on-package transmission lines could be determined. Finally, the chip footprint could be adjusted to enable quarter-wavelength difference between the two on-chip lines without any bends. In this case the two-port S-parameters for the transition could be accurately determined through two calibration steps.

Another area for investigation is the design of these transitions to optimally provide impedance matching for the amplifiers on-chip. As discussed in Chapter 4, prior to matching the transition, the impedance it presents to the chip is in a typical region desired for matching an amplifier. Developing a concrete set of design knobs that can be used for integrated matching would save area on-chip and potentially also reduce losses.

Finally, other transition types can be investigated. The design of differential structures on-chip is highly valuable for common-mode rejection of interference. Therefore the study of differential transitions, or transitions that can naturally integrate a balun, would similarly reduce on-chip area and losses. Furthermore, as 3D IC integration develops, the design of transitions in a 3D-integrated system is worth exploring.

	This Work	**This Work (Simulation)**	[15]	[38]	[26]	[43]
Package Technology	**ABF GL102**	**Organic substrate**	LTCC GL771	RO4350	IPD Carrier	Thin-Film Benzocyclobutene on Silicon
Chip Technology	28 nm **Bulk CMOS**	16 nm **FinFET CMOS**	22 nm SOI CMOS	55 nm SiGe HBT	90 nm CMOS	InP HBT
Interconnect	**Solder Bump**	**Copper Pillar**	Copper Pillar	Copper Pillar	Gold Bump	-
Bump Height(µm)	**75**	**66**	30	-	65	2
Bump Diameter (µm)	**80**	**62**	50	-	65	10
Bump Pitch (µm)	**150**	**156**	175	-	170	-
Center Frequency (GHz)	**140**	**200**	135	130	163	DC-500
Insertion Loss (dB)	**1.0**	**0.4**	1.1	3.0	2.8	0.9
3 dB Bandwidth (GHz)	**85**	**339**	180*	-	170†	500
Matching 10 dB Bandwidth (GHz)	**43**	**90**	N/A	-	200	500

Table 6.1: Chip-to-Package Transition Comparison Table.
* 2 dB *bandwidth;* † *Estimated from graph;* — *Not provided*

Bibliography

[1] Muntadher Alsabah et al. "6G Wireless Communications Networks: A Comprehensive Survey". In: *IEEE Access* 9 (2021), pp. 148191–148243.

[2] Alexander Arsenovic et al. "scikit-rf: An Open Source Python Package for Microwave Network Creation, Analysis, and Calibration [Speaker's Corner]". In: *IEEE Microwave Magazine* 23.1 (Jan. 2022), pp. 98–105.

[3] Nima Baniasadi. "Millimeter-Wave Receiver and Package Design Close to the Device Activity Limits". PhD thesis. EECS Department, University of California, Berkeley, 2021.

[4] Nima Baniasadi, Rami Hijab, and Ali Niknejad. *mm-Wave and sub-THz Chip-to-Package Transitions for Communications Systems*. Apr. 2024.

[5] Mark T. Bohr and Ian A. Young. "CMOS Scaling Trends and Beyond". In: *IEEE Micro* 37.6 (Nov. 2017), pp. 20–29.

[6] Y. Cassivi, L. Perregrini, P. Arcioni, M. Bressan, K. Wu, and G. Conciauro. "Dispersion characteristics of substrate integrated rectangular waveguide". In: *IEEE Microwave and Wireless Components Letters* 12.9 (Sept. 2002), pp. 333–335.

[7] Ethan Chou, Nima Baniasadi, Hesham Beshary, Meng Wei, Emily Naviasky, Lorenzo Iotti, and Ali Niknejad. "A Low-Power and Energy-Efficient D-Band CMOS Four-Channel Receiver with Integrated LO Generation for Digital Beamforming Arrays". In: *ESSCIRC 2022- IEEE 48th European Solid State Circuits Conference (ESSCIRC)*. Sept. 2022, pp. 489–492.

[8] Ethan Chou, Nima Baniasadi, and Ali Niknejad. "A 200 GHz Wideband and Compact Differential LNA Leveraging an Active Balun Input Stage in 16nm FinFET Technology". In: *2024 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*. June 2024, pp. 375–378.

[9] Ethan Chou, Hesham Beshary, Meng Wei, Rami Hijab, Farhana Sheikh, Steven Callender, and Ali M. Niknejad. "Comparative Performance of 100–200 GHz Wideband Transceivers: CMOS vs Compound Semiconductors". In: *2023 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)*. Oct. 2023, pp. 292–299.

[10] Robert E. Collin. *Foundations for microwave engineering*. 2nd ed. IEEE Press series on electromagnetic wave theory. New York: IEEE Press, 2001. ISBN: 978-0-7803-6031-0.

[11] Debendra Das Sharma, Gerald Pasdast, Zhiguo Qian, and Kemal Aygun. "Universal Chiplet Interconnect Express (UCIe): An Open Industry Standard for Innovations With Chiplets at Package Level". In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* 12.9 (Sept. 2022), pp. 1423–1431.

[12] D.C. DeGroot, J.A. Jargon, and R.B. Marks. "Multiline TRL revealed". In: *60th ARFTG Conference Digest, Fall 2002.* Washington, DC, USA: IEEE, 2002, pp. 131–155. ISBN: 978-0-7803-8124-7.

[13] D. Deslandes and K. Wu. "Integrated microstrip and rectangular waveguide in planar form". In: *IEEE Microwave and Wireless Components Letters* 11.2 (Feb. 2001), pp. 68–70.

[14] Rachel J. Dewey. "The Effect of Correlated Noise in Phased-Array Observations of Radio Sources". In: *The Astronomical Journal* 108 (July 1994), p. 337.

[15] Ali A. Farid, Ahmed S. H. Ahmed, Arda Simsek, and Mark J. W. Rodwell. "A Packaged 135GHz 22nm FD-SOI Transmitter on an LTCC Carrier". In: *2021 IEEE MTT-S International Microwave Symposium (IMS)*. June 2021, pp. 713–716.

[16] Chong Leong Gan and Chen-Yu Huang. *Interconnect Reliability in Advanced Memory Device Packaging*. Springer Series in Reliability Engineering. Cham: Springer International Publishing, 2023. ISBN: 978-3-031-26708-6.

[17] Ramesh Garg, I. J. Bahl, and Maurizio Bozzi. *Microstrip lines and slotlines*. Third edition. Artech House microwave library. Boston: Artech House, 2013. ISBN: 978-1-60807-535-5.

[18] M.S. Gupta. "Power gain in feedback amplifiers, a classic revisited". In: *IEEE Transactions on Microwave Theory and Techniques* 40.5 (May 1992), pp. 864–879.

[19] Stephen Hall, Steven G. Pytel, Paul G. Huray, Daniel Hua, Anusha Moonshiram, Gary A. Brist, and Edin Sijercic. "Multigigahertz Causal Transmission Line Modeling Methodology Using a 3-D Hemispherical Surface Roughness Approach". In: *IEEE Transactions on Microwave Theory and Techniques* 55.12 (Dec. 2007), pp. 2614–2624.

[20] W. Heinrich. "The flip-chip approach for millimeter wave packaging". In: *IEEE Microwave Magazine* 6.3 (Sept. 2005), pp. 36–45.

[21] Sanming Hu, Yong-Zhong Xiong, Bo Zhang, Lei Wang, Teck-Guan Lim, Minkyu Je, and Mohammad Madihian. "A SiGe BiCMOS Transmitter/Receiver Chipset With On-Chip SIW Antennas for Terahertz Applications". In: *IEEE J. Solid-State Circuits* 47.11 (Nov. 2012), pp. 2654–2664.

[22] John David Jackson. *Classical electrodynamics*. 2d ed. New York: Wiley, 1975. ISBN: 978-0-471-43132-9.

[23] A. Jentzsch and W. Heinrich. "Theory and measurements of flip-chip interconnects for frequencies up to 100 GHz". In: *IEEE Transactions on Microwave Theory and Techniques* 49.5 (May 2001), pp. 871–878.

[24] Umut Kodak, Bhaskara Rupakula, Samet Zihir, and Gabriel M. Rebeiz. "60-GHz 64- and 256-Element Dual-Polarized Dual-Beam Wafer-Scale Phased-Array Transceivers With Reticle-to-Reticle Stitching". In: *IEEE Transactions on Microwave Theory and Techniques* 68.7 (July 2020), pp. 2745–2767.

[25] Hyung-Jin Lee, Steven Callender, Said Rami, Woorim Shin, Qiang Yu, and Jose Mauricio Marulanda. "Intel 22nm Low-Power FinFET (22FFL) Process Technology for 5G and Beyond". In: *2020 IEEE Custom Integrated Circuits Conference (CICC)*. Mar. 2020, pp. 1–7.

[26] Chun-Hsing Li, Wan-Ting Hsieh, and Te-Yen Chiu. "A Flip-Chip-Assembled W-Band Receiver in 90-nm CMOS and IPD Technologies". In: *IEEE Transactions on Microwave Theory and Techniques* 67.4 (Apr. 2019), pp. 1628–1639.

[27] Tim Maiwald et al. "A Review of Integrated Systems and Components for 6G Wireless Communication in the D -Band". In: *Proceedings of the IEEE* (2023), pp. 1–37.

[28] Didem Melek. *Low Power and Highly BW Efficient D2D Link Design for Chiplets*. Nov. 2024.

[29] Sirinpa Monayakul et al. "Flip-Chip Interconnects for 250 GHz Modules". In: *IEEE Microwave and Wireless Components Letters* 25.6 (June 2015), pp. 358–360.

[30] Ali Niknejad. *EE 142/242A Lectures at University of California, Berkeley*. 2021.

[31] Ali Niknejad. *EE 242B Lectures at University of California, Berkeley*. 2024.

[32] Ali M. Niknejad. *Electromagnetics for High-Speed Analog and Digital Communication Circuits*. Cambridge: Cambridge University Press, 2007. ISBN: 978-0-521-85350-7.

[33] Shiji Pan and Filippo Capolino. "Design of a CMOS On-Chip Slot Antenna With Extremely Flat Cavity at 140 GHz". In: *Antennas Wirel. Propag. Lett.* 10 (2011), pp. 827–830.

[34] Shiji Pan, Leland Gilreath, Payam Heydari, and Filippo Capolino. "Investigation of a Wideband BiCMOS Fully On-Chip W-Band Bowtie Slot Antenna". In: *Antennas Wirel. Propag. Lett.* 12 (2013), pp. 706–709.

[35] Vitaly Petrov, Thomas Kurner, and Iwao Hosako. "IEEE 802.15.3d: First Standardization Efforts for Sub-Terahertz Band Communications toward 6G". In: *IEEE Communications Magazine* 58.11 (Nov. 2020), pp. 28–33.

[36] David M. Pozar. *Microwave engineering*. 4th ed. Hoboken, NJ: Wiley, 2012. ISBN: 978-0-470-63155-3.

[37] Andrej Rumiantsev and Nick Ridler. "VNA calibration". In: *IEEE Microwave Magazine* 9.3 (June 2008), pp. 86–99.

[38] Mahmoud Sawaby, Nemat Dolatsha, Baptiste Grave, Cheng Chen, and Amin Arbabian. "A Fully Packaged 130-GHz QPSK Transmitter With an Integrated PRBS Generator". In: *IEEE Solid-State Circuits Lett.* 1.7 (July 2018), pp. 166–169.

[39] Shahriar Shahramian, Michael J. Holyoak, Amit Singh, and Yves Baeyens. "A Fully Integrated 384-Element, 16-Tile, W -Band Phased Array With Self-Alignment and Self-Test". In: *IEEE Journal of Solid-State Circuits* 54.9 (Sept. 2019), pp. 2419–2434.

[40] C. E. Shannon. "A mathematical theory of communication". In: *The Bell System Technical Journal* 27.3 (July 1948), pp. 379–423.

[41] Choon Beng Sia. "Improving Wafer-Level S-parameters Measurement Accuracy and Stability with Probe-Tip Power Calibration up to 110 GHz for 5G Applications". In: Oct. 2019.

[42] Arda Simsek, Ahmed S. H. Ahmed, Ali A. Farid, Utku Soylu, and Mark J. W. Rodwell. "A 140GHz Two-Channel CMOS Transmitter Using Low-Cost Packaging Technologies". In: *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. Apr. 2020, pp. 1–3.

[43] Siddhartha Sinha, Ralf Doerner, Franz-Josef Schmückle, Sirinpa Monayakul, Michael Hrobak, Nils G. Weimann, Viktor Krozer, and Wolfgang Heinrich. "Flip-Chip Approach for 500 GHz Broadband Interconnects". In: *IEEE Transactions on Microwave Theory and Techniques* 65.4 (Apr. 2017), pp. 1215–1225.

[44] Ho-Jin Song and Namyoon Lee. "Terahertz Communications: Challenges in the Next Decade". In: *IEEE Transactions on Terahertz Science and Technology* 12.2 (Mar. 2022), pp. 105–117.

[45] Wouter Steyaert and Patrick Reynaert. "Layout optimizations for THz integrated circuit design in bulk nanometer CMOS". In: *2017 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS)*. Oct. 2017, pp. 1–4.

[46] Andrew Townley. "Broadband mm-Wave Transceivers for Sensing and Communication". PhD thesis. EECS Department, University of California, Berkeley, 2018.

[47] Andrew Townley, Nima Baniasadi, Sashank Krishnamurthy, Constantine Sideris, Ali Hajimiri, Elad Alon, and Ali Niknejad. "A Fully Integrated, Dual Channel, Flip Chip Packaged 113 GHz Transceiver in 28nm CMOS supporting an 80 Gb/s Wireless Link". In: *2020 IEEE Custom Integrated Circuits Conference (CICC)*. Mar. 2020, pp. 1–4.

[48] Václav Valenta, Thomas Spreng, Shuai Yuan, Wolfgang Winkler, Volker Ziegler, Dragos Dancila, Anders Rydberg, and Hermann Schumacher. "Design and experimental evaluation of compensated bondwire interconnects above 100 GHz". In: *International Journal of Microwave and Wireless Technologies* 7.3-4 (June 2015), pp. 261–270.

[49] Luca Valenziano, Joachim Hebeler, Georg Gramlich, Alexander Quint, Thomas Zwick, and Akanksha Bhutani. "Advanced Bond-Wire Interconnect Solution for Ultra-Broadband Applications Covering DC to 210 GHz". In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* (2024), pp. 1–1.

[50] Atom O. Watanabe, Muhammad Ali, Sk Yeahia Been Sayeed, Rao R. Tummala, and Markondeya Raj Pulugurtha. "A Review of 5G Front-End Systems Package Integration". In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* 11.1 (Jan. 2021), pp. 118–133.

[51] Nils G. Weimann et al. "Manufacturable Low-Cost Flip-Chip Mounting Technology for 300–500-GHz Assemblies". In: *IEEE Transactions on Components, Packaging and Manufacturing Technology* 7.4 (Apr. 2017), pp. 494–501.

[52] Martin White. *Physics 209 Lectures at University of California, Berkeley*. 2022.

[53] Dylan Williams. "Traveling Waves and Power Waves: Building a Solid Foundation for Microwave Circuit Theory". In: *IEEE Microwave* 14.7 (Nov. 2013), pp. 38–45.

[54] Ke Wu, D. Deslandes, and Y. Cassivi. "The substrate integrated circuits - a new concept for high-frequency electronics and optoelectronics". In: *6th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Service, 2003. TELSIKS 2003.* Vol. 1. Oct. 2003, P–III.