1

Reincorporating Circuit Theory into Information Theory

Amine Mezghani, Mohamed Akrout, Miguel R. Castellanos, Sandy Saab, Bertrand Hochwald, Robert W. Heath Jr., and Josef A. Nossek

Abstract—Multiple-Input Multiple-Output (MIMO) technology has emerged as a critical component in modern wireless communication systems. As antenna systems continue to evolve, the development and refinement of the corresponding information-theoretic bounds and achievability results remain essential in unlocking the full potential of future MIMO technologies in information transmission. To harness the benefits of such informationtheoretic analysis, accurate modeling and analysis of the antenna's behavior are essential. Equivalent circuit models provide a valuable tool for understanding the performance limits of MIMO antennas. By providing a simplified representation of complex antenna structures, these models enable researchers to gain insights into the antenna's behavior, predict performance metrics, and optimize the system for optimal operation. This magazine paper reviews the advancements in the physical consistent modeling techniques of MIMO communications and sheds light on the potential avenues for further research and development in the field of information theory for advanced antenna structures and the role of circuit-based modeling.

I. INTRODUCTION

New antenna and circuit technologies are critical to realizing larger array apertures, higher frequencies, and wider bandwidths that are possible for 6G. Advanced antenna structures such as intelligent metasurfaces and large aperture antenna arrays are no longer well captured in a theory that uses abstract antenna arrays, simplified fading channels, and generic power constraints. We believe that it is time for the broader information theory community to rethink many information theoretic concepts starting from electromagnetic concepts at the physical layer level, such as channel, noise, power, and degrees of freedom, all the way up to the network level.

The link between the actual physical quantities and laws that govern the behavior of the communication system with the random variables and joint probability distributions used in information theory is not straightforward. For instance, the channel in information theory is often considered as the part of the system we are "unwilling or unable to change" [1]. Antennas, which

usually are part of the channel, now have the capability to dynamically adapt to the channel under the control of the transceiver through modern technologies such as metasurfaces. Other important constraints and trade-offs such as antenna size and bandwidth limitations are not traditionally incorporated in the information-theoretic performance analysis. For instance, maximizing the achievable rate while accounting for the physical limitations of antennas (e.g., size and mutual coupling) is a problem that lies at the interface of communication and electromagnetic theories. Bridging this gap between these two theories is therefore becoming increasingly important.

Circuit theory is one approach to building a physically-grounded information theory. While directly linking Maxwell's theory of propagation and Shannon's theory of information is attractive and has been elegantly attempted in prior work [2], there are various reasons for rather connecting them halfway via circuit theory, summarized as follows and illustrated in Fig. 1.

- Kirchhoff's laws are a consequence of electromagnetics laws and therefore enjoy their main key properties such as conservation of power, causality, and reciprocity.
- 2) Equivalent-circuit models are widely used in the design and analysis of antennas/metasurfaces and even photonic systems and provide key insights.
- 3) Circuits are mathematical-graphical abstractions of nature, which makes them very attractive from an information-theoretic perspective. This connection to graphical models has been used in the past to perform inference tasks and channel decoding with analog circuits [3], [4].
- 4) Incorporating other components in the RF chain (i.e., amplifier) becomes straightforward which is important for better model fidelity, particularly at higher frequencies where RF circuitry starts to have severe limitations.
- 5) Circuit theory involves linear algebra and transfer functions, which are familiar and accessible to most electrical and computer engineers. This helps engage

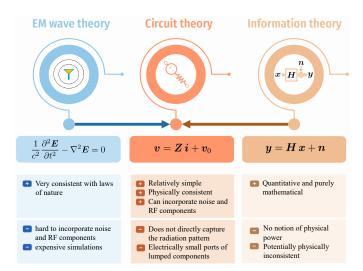


Fig. 1: Information theory meets antenna theory via circuit theory.

more researchers and students to work on this interface.

Moreover, in an era when data-driven and black-box approaches are trending, we do believe it is a perfect time to highlight the importance of circuit-based models in studying communication systems. This is because their white-box nature incorporates additional parameters to define and optimize physical characteristics and constraints of communication components (e.g., antenna spacing, antenna orientation, noise figure).

This paper provides an overview of the fundamentals of circuit-based information theory. It will start by explaining how to use circuits to build a physically consistent model for wireless communication. The implications of such a model will be explored in a single antenna communication setting, through examples that show where traditional Shannon theory and a circuit-based theory diverge in their predictions. The paper then shifts to MIMO communication, where a circuit-based information theory may have many advantages in capturing practical aspects related to antennas, arrays, and radio frequency circuits, as well as challenges such as broadband matching.

There is a developing literature on physics-based and circuit-based information theory, which is built upon earlier efforts in the past seventy years. We will highlight key early results like the bounds by Chu on the theoretical minimum Q-factor as a function of the antenna size [5], and by Fano on fundamental matching bandwidth limits for loads of arbitrary impedance [6]. The fact that these fundamental limits on antennas and circuit matching together with Shannon's limit have been published all in the year 1948 is intriguing. These limits

are all related to the design of communication systems and circuits, even though they refer to different aspects of the system. Besides chronological similarity, there are other interesting analytical parallels. For instance, Shannon's limit on the trade-off between spectral efficiency and energy efficiency is generally known to be mathematically tractable for Gaussian channels whereas Chu's limit on the tradeoff between antenna size and bandwidth is mainly manageable for a spherical size constraint. Shannon is considered the father of digital circuit design theory and information theory, while Fano has significant contributions to both, analog circuit theory and information theory.

Arguments for the necessity to bridge the gap between information theory and physics can be traced back to 1953 [7]. In addition to these historical aspects, we will describe in this paper more recent efforts to bridge information theory and electromagnetics in the context of wireless networks [8] and point-to-point communications [2], as well as holographic MIMO [9]. This line of work focused on quantifying the number of degrees of freedom of radiative communication systems based on wavefield decompositions. Alternatively, and given the advantages of the circuit approach, we review some of our own work on MIMO circuit theory [10] which builds upon the multi-port circuit theory for communication [11], on wideband matching in MIMO systems [12], nearfieldwideband regimes [13]-[15], and achievable rates with antenna size constraints [16]. We also discuss emerging areas of research. Near-field is one example with larger antenna apertures and for certain applications (such as wireless brain-computer interfaces), the communication link is occurring increasingly in the nearfield and even within the reactive zone where the transmitter and receiver pair are interacting bilaterally, the relationship between magnetic and electric fields becomes complex, and the number of degrees of freedom is not limited by the diffraction limit. Metasurface radiating antennas, where the aperture is dynamically controllable with new constraints, are another related example [17].

II. HISTORICAL ASPECTS

Research on fundamental limits imposed by laws of nature and mathematics after World War II experienced significant growth and advancements. Scientists and engineers from various disciplines focused on exploring the fundamental boundaries and constraints of different fields of study. This period witnessed groundbreaking discoveries and theoretical developments in physics, mathematics, engineering, and other scientific branches. The exploration of foundational limits paved the way for

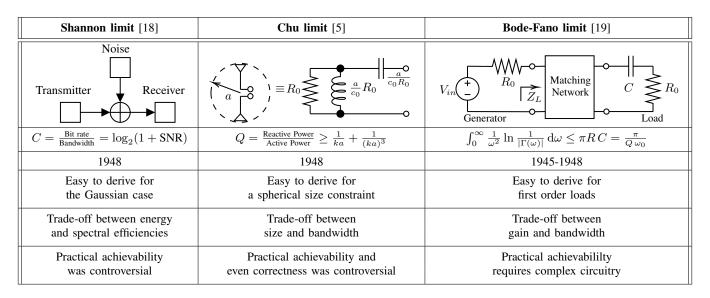


TABLE I: The analogies between the Shannon limit, the Chu limit, and the Fano limit.

a scientific revolution that led to enormous technological advancements in various industries.

In the context of information transmission, the Chu limit [5], the Bode/Fano limit [19], [20] and the Shannon limit [18] are three fundamental limits that play important roles in different aspects of communication systems. A brief summary of the different limits is given in Table I where several parallels can be observed despite their distinct origins and contexts. The Shannon limit gives the maximum achievable data rate that can be transmitted without errors. The Chu limit, also known as the Chu-Harrington limit, is a theoretical characterization of the design and performance of electrically small antennas. The Bode-Fano limit describes how impedance matching affects circuit bandwidth. In the remainder of this section, we give a detailed description of the latter two limits.

The Chu limit was first formulated by Chu and Harrington in the 1940s to analyze the tradeoff between the antenna size and bandwidth. This limit provides an upper bound on the radiation quality factor (Q) that can be achieved by an electrically small antenna for a given operating frequency and antenna size. Q is a measure of the antenna's ability to efficiently radiate or absorb electromagnetic energy over a frequency band. Electrically small antennas, which are significantly smaller than the wavelength of the operating frequency, often face challenges in achieving high radiation efficiency due to their size limitations. The Chu limit specifies the maximal efficiency that can be achieved by small antennas.

The Chu limit sets a fundamental constraint on the achievable Q value based on the physical size of the

antenna as
$$Q \geq \frac{1}{ka} + \frac{1}{(ka)^3}, \tag{1}$$

where k is the wave number and a is the radius of the smallest sphere entirely enclosing the antenna. For large Q-factors, the bandwidth in relation to the carrier frequency is inversely proportional to the Q-factor. In remarkable coincidence to the Shannon limit, this pioneering work triggered vast research efforts to design small antennas approaching this fundamental limit [21]. A large body of research studies within the antenna community has proposed over the years multiple antenna designs approaching the Chu limit. Fig. 2 depicts more than 100 antenna designs for either single or dual radiation modes published in the open literature. More recent advances on the topic of antenna limits under general size constraint other than spherical shape can be found in [22]. Achieving the Chu limit is basically equivalent to the problem of creating the radiation field of an idealized infinitesimal Hertzian dipole, particularly the near-field, with a practical antenna that is confined within a sphere of fixed size. This goal turns out to be very challenging particularly for smaller size antennas because of the low amount of stored energy in the near-field outside the sphere enclosing the antenna.

The Q factor quantifies the ratio of the energy stored in the system to the energy dissipated per cycle. A higher Q factor indicates a more selective resonant response, with a narrower bandwidth. While for resonant circuits a higher Q-factor is desired, wideband antennas should preferably have a low Q-factor.

An accurate characterization of the bandwidth limitation due to the reactive/resonant behavior (and thus the Q factor) of a load such as an antenna has been

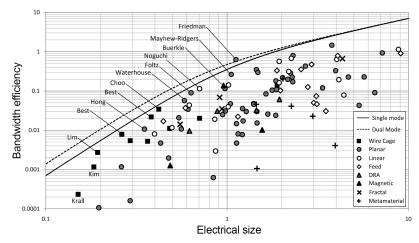


Fig. 2: The maximum bandwidth efficiency product for 110 antenna designs published in the IEEE Transactions on Antennas and Propagation by the end of the year 2010. Figure taken from [21] with permission.

found by Bode [20] for first order circuits and later generalized by Fano [19] for general order circuits again in 1948¹. The Bode-Fano limit sets a fundamental constraint on the achievable bandwidth from impedance matching between, for instance, a generator and a certain load impedance. Matching is the process of ensuring that the output impedance of a source matches the input impedance of a load. Matching can be as simple as tuning a resistor but often involves the help of a matching network as shown in Fig. 3. This is done to maximize power transfer as

$$P_{\text{out}} = |V_{\text{in}}|^2 \frac{\text{Re}\{Z_L\}}{|Z_L + R_0|^2} = \frac{|V_{\text{in}}|^2}{4R_0} \left(1 - \left|\frac{Z_L - R_0}{Z_L + R_0}\right|^2\right)$$
$$= \frac{|V_{\text{in}}|^2}{4R_0} \left(1 - |\Gamma(\omega)|^2\right),$$
(2)

where $\Gamma(\omega) = \frac{Z_L - R_0}{Z_L + R_0}$ is the reflection coefficient. Alternatively one can aim at maximizing the signal-to-noise ratio at the load. The Bode/Fano limit is a consequence of Foster's theorem that states that all lossless, passive circuits must always exhibit a positive slope reactance with respect to frequency. This sets a limit on the achievable transformed reflection coefficient $\Gamma(\omega)$. For the first order circuit as in Fig. 3, the Bode/Fano limit is [20]

$$\int_0^\infty \frac{1}{\omega^2} \ln \frac{1}{|\Gamma(\omega)|} d\omega \le \pi R C = \frac{\pi}{Q \omega_0}, \quad (3)$$

where ω_0 is some center frequency. Since the inverse relationship between Q-factor and bandwidth holds only for high Q-factors, these antenna theoretical results are,

¹While being principally known for their work in information theory, both Shannon and Fano did major contributions to analog and digital circuit theory, which is very remarkable.

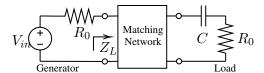


Fig. 3: Matching a generator to a capacitive load

however, mainly applicable to narrowband systems. In addition, the antenna efficiency is calculated assuming that thermal noise is dominant. Therefore, a capacity-based performance metric, as presented in the next section, seems to be much more appropriate for such cases compared to those used in the antenna theory literature.

III. INCORPORATING CIRCUITS INTO INFORMATION THEORY: THE SISO CASE

In this section, we describe how a SISO communication system is represented as a circuit-theoretic model. We highlight how circuit-based models provide a systematic approach to incorporate the physically consistent definition of power, the antenna size constraint, and the realizability conditions of matching networks. We then present the achievable rate optimization that unifies Shannon's theory with the different physical limits discussed in Section II.

A. Circuit theory for SISO communications

A SISO communications system can be represented by a two-port network as shown in Fig. 4. The transmit signal is generated by the voltage source $V_{\rm G}(f)$ with internal resistance R in series. The signal then passes through the transmit matching network with impedance matrix $\mathbf{Z}_{\rm MN_T}$. The noise voltage source $\widetilde{V}_{\rm T}(f)$ and $\widetilde{V}_{\rm R}(f)$

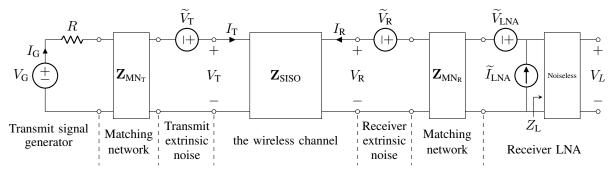


Fig. 4: SISO communication model including the signal generator, the transmit and receive extrinsic noises, the communication channel with antenna mutual coupling, the matching networks, and the LNA model with its two intrinsic noise sources. The output voltage V_L is connected to the load impedance of a receiving device. The frequency argument was dropped to lighten the notation in the figure.

are connected to the ports of the transmit and receive antenna impedances, $Z_{\rm T}$ and $Z_{\rm R}$, respectively. They model the extrinsic noise of the transmit/receive antennas at temperature T originating from background radiation having a zero mean and a variance given by

$$\mathbb{E}\left[\left|\widetilde{V}_{\mathsf{T}}\right|^{2}\right] = 4 k_{\mathsf{b}} T \Re\{Z_{\mathsf{T}}\},\tag{4a}$$

$$\mathbb{E}\left[\left|\widetilde{V}_{\mathsf{T}}\right|^{2}\right] = 4 \, k_{\mathsf{b}} \, T \, \Re\{Z_{\mathsf{T}}\},\tag{4a}$$

$$\mathbb{E}\left[\left|\widetilde{V}_{\mathsf{R}}\right|^{2}\right] = 4 \, k_{\mathsf{b}} \, T \, \Re\{Z_{\mathsf{R}}\}.\tag{4b}$$

The antenna port at the receiver is connected through the matching network \mathbf{Z}_{MN_R} to the circuit model of the lownoise amplifier (LNA). The amplifier is modeled with input resistance $Z_{\rm L}$, a voltage noise source $V_{\rm LNA}$, and a current noise source I_{LNA} . These two noise sources characterize the intrinsic noise of the amplifier and are modeled by the following statistical properties [11]:

$$\mathbb{E}\left[\left|\widetilde{I}_{\text{LNA}}\right|^2\right] = \beta,\tag{5a}$$

$$\mathbb{E}\left[\left|\widetilde{V}_{\mathrm{LNA}}\right|^{2}\right] = \beta\,R_{\mathrm{N}}^{2},\tag{5b}$$

$$\mathbb{E}\left[\widetilde{V}_{\text{LNA}}\,\widetilde{I}_{\text{LNA}}^*\right] = \beta\,\rho\,R_{\text{N}},\tag{5c}$$

where β is the spectral density of the noise current, ρ is the complex noise correlation coefficient between $V_{\rm LNA}$ and I_{LNA} , and R_N is the so-called noise resistance. In Fig. 4, the two-port network \mathbf{Z}_{SISO} represents the antenna-aware wireless channel as it abstracts the relationship between the port variables (I_T, V_T) and (I_R, V_R) of the transmit and receive antennas as follows:

$$\begin{bmatrix} V_{T}(f) \\ V_{R}(f) \end{bmatrix} = \underbrace{\begin{bmatrix} Z_{T}(f) & Z_{TR}(f) \\ Z_{RT}(f) & Z_{R}(f) \end{bmatrix}}_{\mathbf{Z}_{SISO}(f)} \begin{bmatrix} I_{T}(f) \\ I_{R}(f) \end{bmatrix}. (6)$$

In (6), the diagonal elements, Z_T and Z_R , represent the self-impedances of transmit and receive antennas. The

off-diagonal entries, Z_{RT} and Z_{TR} , represent the mutual transmit-receive and receive-transmit impedances, respectively, between transmit and receive antennas. The mutual impedance elements are a practical measure of the mutual coupling effects [23].

The circuit model in Fig. 4 is appropriate to characterize the performance of both near-field and far-field communication. For near-field communication, one can determine the mutual impedance $Z_{RT}(f)$, from which it follows $Z_{TR}(f) = Z_{RT}(f)$ since antennas are reciprocal devices. Optimal near-field impedance matching, however, does not have closed-form solutions and only numerical approximations are achievable [13]. For farfield communication, the unilateral approximation is used by setting the receive-transmit mutual impedance $Z_{\rm TR}(f)$ to zero and resorting to the Friis' equation to compute the transmit-receive impedance $Z_{RT}(f)$. The far-field circuit model also ignores the extrinsic/intrinsic noises at the transmitter as they are attenuated by the large-scale path-loss.

Furthermore, the two-port networks \mathbf{Z}_{SISO} and \mathbf{Z}_{MN_R} can be combined into a single two-port network \mathbf{Z}'_{SISO} as a cascaded connection of two two-port networks. Applying these simplifications to the circuit model in Fig. 4 yields the circuit model for far-field communication depicted in Fig. 5. There, similarly to (6), the relationship between the transmit/receive voltages and currents is given by Ohm's law:

$$\begin{bmatrix} V_{\mathsf{T}}'(f) \\ V_{\mathsf{R}}'(f) \end{bmatrix} = \underbrace{\begin{bmatrix} Z_{\mathsf{T}}'(f) & Z_{\mathsf{TR}}'(f) \\ Z_{\mathsf{RT}}'(f) & Z_{\mathsf{R}}'(f) \end{bmatrix}}_{\mathbf{Z}_{\mathsf{SISO}}'(f)} \begin{bmatrix} I_{\mathsf{T}}'(f) \\ I_{\mathsf{R}}'(f) \end{bmatrix}. (7)$$

In the rest of this section, we drop the frequency argument to lighten the notation.

(11b)

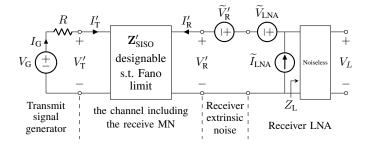


Fig. 5: SISO communication model for far-field communication including the signal generator, the receive extrinsic noise, the communication channel with antenna mutual coupling and the receive matching network, and the LNA model with its two intrinsic noise sources.

B. The physically consistent definition of power

When the communication system model is not physically consistent, the notion of power only refers to the total transmit power. It does not recognize the difference between the total power at the transmitter and the radiated power by the antennas. For this reason, studying the energy efficiency of communication systems must incorporate the physical quantities via EM and circuit theory. It is therefore important to distinguish between the following two notions of power:

ullet the available power $P_{
m T}^{
m avail}$ representing the power budget at the transmitter and is given by:

$$P_{\rm T}^{\rm avail} = \int P_{\rm T}(f) \, \mathrm{d}f = \frac{1}{4 \, R} \int \mathbb{E} \big[|V_{\rm G}(f)|^2 \big] \, \mathrm{d}f, \quad (8)$$

In (8), $P_{\rm T}(f)$ represents the generator power spectral density (PSD). The available power $P_{\rm T}^{\rm avail}$ is an upper bound on the power radiated $P_{\mathrm{T}}^{\mathrm{rad}}$ in the radiated nearfield and far-field regions of the antenna. The available power has multiple hardware constraints such as battery limitations and operational regimes for amplifiers.

• the radiated power corresponding to the fraction of the total power $P_{\rm T}^{\rm rad}$ used to propagate the radio signal from the antenna aperture to the propagation medium. Therefore, the radiated power depends on the EM properties of the antenna and is given by:

$$P_{T}^{\text{rad}} = \mathbb{E}\left[I_{G}^{*} \Re\{Z_{T}'\} I_{G}\right]$$

= $\mathbb{E}\left[V_{G}^{*} (Z_{T}^{'*} + R)^{-1} \Re\{Z_{T}'\} V_{G} (Z_{T}' + R)^{-1}\right].$ (9)

Here, the radiated power also depends on matching condition between the generator internal resistance and the antenna impedance. Moreover, the radiated power is generally regulated by law regulations through radiation exposure limitations.

From a design perspective, these different definitions a) the antenna size constraint by incorporating the anof powers are relevant and can be incorporated as two

separate constraints in a potential transmission rate optimization. This highlights again the benefits of a circuitbased modeling.

C. Physically realizable impedance matching

The impedance matching problem can be stated with reference to the receive side of Fig. 4 as follows. Find the conditions of physical realizability of the impedance $Z'_{\rm R}$ or equivalently the reflection coefficient $\Gamma(f)$ between the noise resistance R_N of the LNA and the matched receive antenna Z'_{R} which is given by:

$$\Gamma(f) = \frac{Z'_{\rm R} - R_{\rm N}}{Z'_{\rm R} + R_{\rm N}}.$$
 (10)

The square magnitude of the reflection coefficient $|\Gamma(f)|^2$ characterizes the fraction of the received power $P_{\rm R}(f)$ rejected by the LNA. Fano has characterized in [19] the conditions of physical realizability of the reflection coefficient $\Gamma(f)$ which takes the form of the following two integral constraints:

$$\frac{1}{2\pi^2} \int_0^\infty f^{-2} \ln\left(\frac{1}{|\Gamma(f)|^2}\right) df = \left(\frac{2a}{c} - 2\gamma^{-1}\right),$$
(11a)
$$\frac{1}{8\pi^4} \int_0^\infty f^{-4} \ln\left(\frac{1}{|\Gamma(f)|^2}\right) df = \left(\frac{4a^3}{3c^3} + \frac{2}{3}\gamma^{-3}\right),$$

where a is the radius of the Chu's sphere of the antenna, and γ is the positive real-valued zero of the reflection coefficient.

In the rest of this section, we will characterize the achievable rate of the SISO communication circuit model for far-field communication. For simplicity, we assume that the transmit antenna port is matched to the resistance of the voltage generator, i.e., $Z_T = R$, and focus on the optimization at receiver side.

D. Achievable rate criterion: Shannon meets Chu Bode and Fano

Although antennas are fundamental devices for wireless transmissions, they are commonly designed based on non-communication metrics (e.g., scattering parameters and impedance matching) while the figure of merit for communication systems is the achievable rate and spectral efficiency criteria. By employing circuit theory, it is possible to maximize the SISO achievable rate while jointly accounting for the following constraints:

tenna circuit model from Chu's theory. This is contrary

is assumed to be infinite.

- b) the physical realizability conditions of matching networks between the antenna and the LNA from Bode-Fano's matching theory.
- c) the impact of the antenna on environmental noise, intrinsic noise, as well as interference.

The achievable rate of the SISO communication circuit model in Fig. 5 is established through the input-output relationship between the transmit voltage $V_G(f)$ and the load voltage $V_L(f)$. Using basic circuit analysis, it follows that:

$$V_{\rm L}(f) = C(f) V_{\rm G}(f) + W(f).$$
 (12)

In (12), $V_G(f)$ is the frequency-domain representation of the signal to be transmitted over the channel. C(f)represents the transfer function of the overall channel and W(f) is the Fourier transform of the noise w(t) given

$$C(f) = \frac{Z_{\rm L} Z'_{\rm RT}}{2 R (Z_{\rm L} + Z'_{\rm R})},$$
 (13a)

$$W(f) = \frac{Z_{\rm L}}{Z_{\rm L} + Z_{\rm R}'} \left(Z_{\rm R}' \, \widetilde{I}_{\rm LNA} - \widetilde{V}_{\rm LNA} + \widetilde{V}_{\rm R}' \right). \quad (13b)$$

Using the noise properties in (4) and (5), we write the noise PSD of W(f) as

$$N(f) = \left| \frac{Z_{L}}{Z_{L} + Z'_{R}} \right|^{2} \left(\beta \left| Z'_{R} + R_{N} \right|^{2} + 4 k_{b} T \Re \{ Z'_{R} \} \right), \tag{14}$$

where we set ρ to -1 to capture the worst-case noise correlation scenario between the two intrinsic noise sources I_{LNA} and V_{LNA} of the LNA.

It is important to highlight the key features of the aforementioned circuit-based modeling approach:

- a) A controllable channel: The transfer function of the channel in (13a) depends on both self- and mutual impedances. As a result, the channel can be designed as a function of the transmit/receive antenna properties. This illustrate the partial controllability of the channel.
- b) A richer description of the noise: Unlike the conventional additive Gaussian noise, the noise in (13b) involves the impedances at receive side. This means that the frequency-domain noise representation depends on the circuitry. The noise accounts for multiple sources as opposed to the previous prior art which only considers the coarse superposition of noise sources modeled as a single additive Gaussian random variable. When active devices are present, it is also possible to incorporate the model of their noise beyond the fully passive AWGN noise.

to classical Shannon's theory where the antenna size c) **Inclusion of RF chain components**: The end-to-end circuit-based modeling of the communication system in Fig. 4 incorporates the circuit model of the amplifier. It can also be extended to incorporate filters and more detailed circuits of matching networks and amplifiers.

> Using the channel and the noise PSD in (13a) and (14), we write the signal-to-noise ratio (SNR) for a given transmit PSD $P_{\rm T}(f)$ defined in (8) as follows:

$$SNR(f) = \frac{4RP_{T}(f)|C(f)|^{2}}{N(f)}$$

$$= \frac{P_{T}(f)|Z'_{RT}|^{2}}{R\left(\beta \left|Z'_{R} + R_{N}\right|^{2} + 4k_{b}T\Re\{Z'_{R}\}\right)}.$$
(15)

Here, the mutual impedance Z'_{RT} represents the spatially correlated communication channel and is given by:

$$Z'_{RT}(f) = 2 \Re \left\{ Z'_{R} \right\}^{\frac{1}{2}} H(f) \Re \left\{ Z'_{T} \right\}^{\frac{1}{2}},$$
 (16)

where H(f) is the communication channel². In (16), the channel $Z'_{RT}(f)$ is spatially correlated as opposed to H(f) due the transmit/receive impedance factors stemming from the the mutual coupling within the arrays. After recalling the fact that the transmitter is matched (i.e., $Z'_T = R$), we inject (16) back into (15) to get:

$$SNR(f) = \frac{4 P_{T}(f) |H(f)|^{2} \Re\{Z'_{R}\}}{\beta |Z'_{R} + R_{N}|^{2} + 4 k_{b} T \Re\{Z'_{R}\}}.$$
 (17)

After dividing both the nominator and denominator of (17) by $|Z'_{R} + R_{N}|^{2}$, we use the equality

$$\frac{\Re\{Z_{\rm R}'\}}{|Z_{\rm R}' + R_{\rm N}|^2} = \frac{1}{4R_{\rm N}} \left(1 - |\Gamma(f)|^2\right),\tag{18}$$

to obtain

$$SNR(f) = \frac{P_{T}(f) |H(f)|^{2} (1 - |\Gamma(f)|^{2})}{\beta R_{N} + k_{b} T (1 - |\Gamma(f)|^{2})}.$$
 (19)

Using the input-output relationship of the SISO circuit model given in (13), the achievable rate under the aforementioned physical constraints is the highest mutual information $I(V_G(t); V_L(t))$ per unit time between the two random processes $V_G(t)$ and $V_L(t)$ representing the input and output signals of the communication system in Fig. 5. In other words, [16]:

$$C = \max_{\mathbb{P}_{\mathsf{V}_{\mathsf{G}}}, \Gamma(f)} I(\mathsf{v}_{\mathsf{G}}(t); \mathsf{v}_{\mathsf{L}}(t)) \text{ [bits/s]}. \tag{20}$$

The maximization of the mutual information is performed w.r.t. the probability measure \mathbb{P}_{V_G} on the space

²The factor 2 in (16) stems from the conversion between the impedance and scattering parameters.

of possible generator voltages $V_G(t)$, as well as the reflection coefficient $\Gamma(f)$. Using the expression of the SNR established in (19), we rewrite (20) as

$$C = \max_{\mathbb{P}_{V_{G}}, \Gamma(f)} \int_{0}^{\infty} \log_{2} \left(1 + \frac{P_{T}(f) |H(f)|^{2} \left(1 - |\Gamma(f)|^{2} \right)}{\beta R_{N} + k_{b} T \left(1 - |\Gamma(f)|^{2} \right)} \right) \mathrm{d}f, \tag{21}$$

subject to

$$\begin{cases} \int_0^\infty f^{-2} \ln\left(\frac{1}{|\Gamma(f)|^2}\right) df = 2\pi^2 \left(\frac{2a}{c} - 2\gamma^{-1}\right), \\ \int_0^\infty f^{-4} \ln\left(\frac{1}{|\Gamma(f)|^2}\right) df = 8\pi^4 \left(\frac{4a^3}{3c^3} + \frac{2}{3}\gamma^{-3}\right), \\ 0 \le |\Gamma(f)|^2 \le 1, \ \forall f, \end{cases}$$
(22)

The achievable rate optimization in (21)–(22) combines three fundamental limits at once: i) Shannon's limit presented by the celebrated Shannon capacity formula for band-limited additive white Gaussian noise (AWGN) channels, ii) the Chu's limit when the self- and mutual impedances involved in the two-port network \mathbf{Z}'_{SISO} are those of Chu's antennas confined in a sphere of a finite radius a, and iii) the physical realizability conditions of the matching network from Bode-Fano's theory. This optimization demonstrates how circuit-based models for communications are an effective abstraction to analyze the realistic physical limits on the achievable rate which are not captured by Shannon's theory.

In practice, the antenna design is not performed by jointly accounting for all the aforementioned fundamental limits as done in (21)–(22). Indeed, the antenna community focuses on optimizing the scattering parameters of antennas and improving the performance of matching networks. For this reason, existing antenna designs provide sub-optimal radiation behaviors for communication scenarios. For example, antennas are designed to mitigate the extrinsic noise from the environment only and hence are not optimized to account for the contribution of the interference noise between users. Using circuit-based models, it is possible to incorporate in Fig. 5 an additional current in parallel to one of the receive antenna to model the effect of the homogeneous interference scenario with a total received interference power I. Toward this goal, we consider a downlink multi-cell system. The interferer's locations $x_i \in \mathbb{R}^2$ and small scale powers $p_i \in \mathbb{R}^+$ follow a 2-dimensional Marked Point Process $\Phi = \{x_i, p_i\}$. Moreover, we consider a Poisson Point Process (PPP) with density ρ such that the marks are exponentially distributed according to the small scale power distribution of the Rayleigh faded channel. The mean and variance of the total received interference power I with Rayleigh fading interference channels have well-known closed-form expressions [24]. The characterization of the distribution of the interference power I, $p_{\rm I}(I)$ can then be approximately achieved using the second-order moment matching with the Gamma distribution $q_{\rm I}(I)$ [25]. As a result, given the optimal reflection coefficient $\Gamma^{\star}(f)$, the achievable rate in (21) under the distribution of the interference power I becomes:

$$C = \mathbb{E}_{q_{\mathrm{I}}(I)} \left[\int_{0}^{\infty} \log_{2} \left(1 + \frac{P_{\mathrm{T}}(f) |H(f)|^{2} \left(1 - |\Gamma^{\bigstar}(f)|^{2} \right)}{\beta R_{\mathrm{N}} + (k_{\mathrm{b}} T + I) \left(1 - |\Gamma^{\bigstar}(f)|^{2} \right)} \right) \mathrm{d}f \right]. \tag{23}$$

By doing so, one can investigate how the fundamental physical limitations affect the achievable rate in the presence of interference [16]. Fig. 6 depicts the impact of the antenna size and the optimal matching network on the achieved percentage of Shannon's capacity (i.e., when the antenna size is infinite) as a function of the users' density There, we used a fixed boxcar transmit PSD and we considered two antenna sizes with or without a receive matching network.

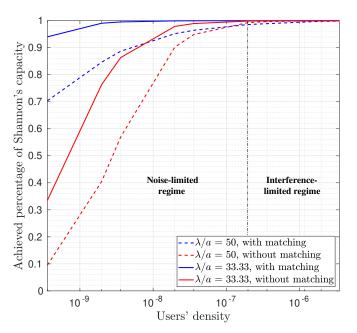


Fig. 6: Achievable rate ratio vs. users' density with a path-loss of $\alpha=2.5$ and fixed transmit power of P=6 [W].

It is seen that as the interference increases with the user density, the achievable rate does not depend on the antenna size. This is because both the interference signal and the intended signal are undergoing the same antenna frequency response regardless of the antenna size. In this *interference-limited* regime (i.e., the achievable rate plateau), the improvements brought by incorporating the antenna size and the matching network constraints are dominated by the interference signal. In the *noise-limited* regime, when the noise contribution becomes more significant at lower user density, matching networks and larger antenna sizes provide a higher achievable rate.

With this observation at hand, the distinction between intrinsic noise from devices and extrinsic noise from the environment is key to investigating the impact of the physical limitations on the achievable rate. This could not be possible without the flexibility of circuit-based models for communication to incorporate both antenna and communication considerations into a single optimization framework. This also suggests that information-theoretic design criteria can be more appropriate than conventional antenna design practices.

IV. INCORPORATING CIRCUITS INTO INFORMATION THEORY: THE MIMO CHALLENGE

Incorporating physical laws into informationtheoretical analyses becomes even more critical when dealing with multiple antennas. To emphasize the significance of circuit theory, let's consider a scenario involving multiple-input-single-output (MISO) communication system composed of a uniformly weighted transmit antenna array with N elements and one single receive antenna. The nth antenna transmits a single scalar signal $x_n(t)$ representing a signal $s_n(t)$ with a different phase ϕ_n and amplitude a_n , i.e., $x_n(t) = s_n(t) \cdot a_n \exp(j\phi_n)$. Let P_T and P_R denote the transmit and receiver power, respectively. The concept of the transmit array gain $A_{\rm T}$ defined as

$$A_{\rm T} = \frac{\max_{a_1, \phi_1, \dots, a_n, \phi_n} P_{\rm R}/P_{\rm T}}{\max_{a_1, \phi_1, \dots, a_n, \phi_n} P_{\rm R}/P_{\rm T}|_{N=1}}$$
(24)

quantifies how much more receive power can be obtained at most for the same transmit power by using all N antennas instead of using only a single antenna at the transmitter. The result is that the transmit array gain equals to the number of transmit antennas. This is a well-known fact in the array signal processing literature for uniformly weighted arrays (i.e., all antenna elements contribute equally) [26].

By adopting circuit-based models using the impedance description presented in Section III, it has been shown in [11] that the transmit array gain depends on the number of antennas and the array impedance, as well as the direction of beamforming when considered. This means that the transmit array gain is a function of the distance between the antenna elements the relative orientation between them. As a result, one can incorporate the transmit array parameters and the physical definition of radiated power which are not captured in information theory to optimize the achievable rate, thereby bridging the gap between information and antenna theories. In the following, we will summarize previous efforts aimed at

addressing this gap. Then, we discuss recent developments and open challenges.

A. Prior attempts to bring together the different results

1) Field-based approaches: Another category of work that has aimed to bridge the gap between physical models and mathematically-abstracted communication models involves a field-based analysis of wireless systems [9], [27], [28]. The main premise of these studies is to analyze key properties of contiguous surfaces, such as the degrees of freedom and propagation models, by leveraging a number of decompositions and approximations of electromagnetic waves. The achievable link gain and spatial degrees-of-freedom of large intelligent surfaces are analyzed in [27] by studying the eigenmodes of current sheets. A near-field model for communication between a large planar array and a single antenna in [28] accounts for near-field propagation and polarization mismatches, and this model is used to analyze the signalto-noise ratio as the array size grows asymptotically. A Fourier plane-wave expansion is used in [9] to develop stochastic electromagnetic models for communication with large holographic arrays. A common feature of these studies is the focus on asymptotically large numbers of antennas, which can mean an infinitely large aperture or a continuous surface comprising denselypacked elements.

While the field-based approaches excel at characterizing over-the-air propagation, the circuit models encompass a larger portion of the system by accounting for the RF components, matching, and even coupling. The field-based models largely assume an existing current distribution on the antennas to carry out the mathematical analysis. In practice, however, the antenna currents and even the feed currents depend on the underlying circuitry. A combination of these two approaches could potentially leverage circuit-based models for finding the antenna currents and then the field-based models for characterizing the channel. In the remainder of this paper, we focus solely on the circuit-based models, and leave the combination of these two approaches for future work.

2) Circuit-based models for narrow-band systems: Circuit-based modeling of MIMO communication systems is another important approach for enabling accurate performance evaluation which was introduced and popularized by [11], [29]. An unexpected outcome of this study is that the laws of nature can actually yield superior performance than the one predicted by conventional oversimplified models. The transition from the general electromagnetic theory to circuit theory is based on the concept of lumped ports and terminals.

An antenna system interacts with its circuit environment through the exchange of electromagnetic energy at specific points known as ports. These ports serve as interfaces between the antenna and the circuitry, allowing for the transfer of signals in both directions. The antenna ports are typically connected to transmission lines or other circuit elements such as amplifiers, enabling the flow of electrical energy to and from the antenna. It is generally assumed that each port can only support one single mode and therefore it can be fully described by two scalar quantities such as incident and reflected wave each expressed as phasor a and b or equivalent voltage vand current i. In contrast to signal and information theory, a circuit port or connection is characterized by two physical quantities which are required to define physical power. The assignment between the physical quantities and the information-theoretic signal establishes the link between the underlying physical phenomena and the corresponding mathematical representations in the realm of information theory [30], [31].

Through the ports, electrical signals can be injected into the antenna for transmission or extracted from it for reception. This bidirectional signal transfer allows for communication and information exchange between the antenna and the circuitry. The ports provide the necessary interface for the electrical energy to flow to and from the antenna, ensuring efficient signal transmission and reception. The concept of lumped ports enables the modeling of an antenna as an equivalent circuit element at its interface. It simplifies the representation of the antenna's behavior in a circuit or system.

It is important to note that while a lumped port provides a simplified representation for circuit analysis and modeling, it does fully capture the radiation behavior and characteristics of the antenna. In addition to the circuit behavior, radiation patterns are required which describes the emission of electromagnetic waves from each excited antenna port into the surrounding space. It represents the primary purpose of the antenna, which is to transmit or receive signals wirelessly. Interestingly, there is an intricate relationship between the circuit behavior of an antenna system and its embedded patterns which will be explained in the following and is based on the following three principles: i) linearity (a.k.a., the superposition principle), ii) conservation of power particularly when the antenna is lossless, and iii) reciprocity between the transmission and reception modes of antennas.

As shown in Fig 7, the electromagnetic properties of an antenna array or a scatterer are described by its radiating/receiving patterns, the space-side scattering pattern, as well as the electrical multi-port behavior of its terminals [32], [33]. For simplicity, we treat the

electromagnetic (EM) field, which is generally a vector complex quantity, as a complex-valued scalar quantity (e.g., considering a single polarization). At each mth accessible port, the forward and backward traveling wave phasors along the antenna feed line are denoted by $a_{\alpha,m}$ and $b_{\alpha,m}$, respectively. At a distant sphere, the angular transmission characteristic of a certain polarization when the antenna port m is connected to a wave-source amplitude $a_{\alpha,m}$ (in $\lceil \sqrt{\mathbf{W}} \rceil$) at port m is defined as (also known as complex pattern)

$$s_m(\theta, \varphi) = \lim_{r \to \infty} -j r e^{jkr} \frac{E^{(m)}(\theta, \varphi, r)}{a_{\alpha, m}} \sqrt{1/\eta_0}, \quad (25)$$

where η_0 is the characteristic impedance of free space and $E^{(m)}(\theta,\varphi,r)$ is the corresponding generated electrical far-field. For simplicity, we will assume that all antennas have the same polarization. These embedded element complex patterns, $s_m(\theta,\varphi)$, are stacked together to form the receiving/transmitting characteristic vector $s_{\alpha\beta}(\theta,\varphi)$. Assuming single polarization hemispherical radiation, the general antenna array can now be characterized by the linear scattering description for the terminal side

$$\boldsymbol{b}_{\alpha} = \boldsymbol{S}_{\alpha\alpha}\boldsymbol{a}_{\alpha} + \int_{0}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} \boldsymbol{s}_{\alpha\beta}(\theta, \varphi) \, a_{\beta}(\theta, \varphi) \sin\theta \, d\varphi \, d\theta,$$
(26)

where $\boldsymbol{a}_{\alpha} \triangleq [a_{\alpha,1},\ldots,a_{\alpha,M}]^{\mathrm{T}}$ and $\boldsymbol{S}_{\alpha\alpha}$ is the scattering matrix . For the space-side, the angular spectra of the outgoing propagating wave phasors $b_{\beta}(\theta,\varphi)$ are expressed as linear functions of the incoming wave $a_{\beta}(\theta,\varphi)$ as well as the port incident phasor

$$b_{\beta}(\theta, \varphi) = \mathbf{s}_{\alpha\beta}(\theta, \varphi)^{\mathsf{T}} \mathbf{a}_{\alpha} + \int_{0}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} \mathbf{s}_{\beta\beta}(\theta, \varphi, \theta', \varphi') a_{\beta}(\theta', \varphi') \sin\theta' d\varphi' d\theta',$$
(27)

where $s_{\beta\beta}(\theta, \varphi, \theta', \varphi')$ is the wave back-scattering characteristic. Ideally, $s_{\beta\beta}(\theta, \varphi, \theta', \varphi')$ is effectively negligible or is just considered as part of the fixed environment.

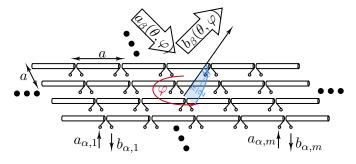


Fig. 7: Antenna scattering with m accessible ports.

Assuming lossless antennas with the property

$$S_{\alpha\alpha}S_{\alpha\alpha}^{\mathsf{H}} + \underbrace{\int_{0}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} s_{\alpha\beta}(\theta, \varphi) s_{\alpha\beta}(\theta, \varphi)^{\mathsf{H}} \sin\theta d\varphi d\theta}_{\triangleq \mathbf{I},$$

$$\underline{\underline{\triangle}}_{B}$$
(28)

then a relationship between the embedded pattern coupling matrix B (also called radiation matrix) and the S-matrix $S_{\alpha\alpha}$ is given by

$$B = I - S_{\alpha\alpha} S_{\alpha\alpha}^{\mathsf{H}} = U \Lambda U^{\mathsf{H}}. \tag{29}$$

Due to reciprocity, i.e., $S_{\alpha\alpha} = S_{\alpha\alpha}^{\mathsf{T}}$, we can obtain when all eigenvalues of B are distinct

$$S_{\alpha\alpha} = U \operatorname{Diag}(e^{j\alpha_1}, \dots, e^{j\alpha_M}) \sqrt{\mathbf{I} - \mathbf{\Lambda}} U^{\mathsf{T}},$$
 (30)

with arbitrary phases $\alpha_1 \cdots \alpha_M$ (which are equal if the antennas are assumed to have identical embedded patterns). The fact that the mutual coupling effects of any antenna structure, being a near-field phenomenon, can be derived solely from the complex far-field pattern is remarkable but it is just a direct consequence of the uniqueness theorem of electromagnetism where the boundary conditions are given at infinity. However, it is important to note that the required patterns are the embedded patterns (irrespective of the multi-port termination) and not the isolated patterns. Alternatively, one can use port current and voltage phasors to describe the antenna instead of wave phasors using the relation

$$a_{\alpha} = \frac{v + R_0 i}{2\sqrt{R_0}},$$

$$b_{\alpha} = \frac{v - R_0 i}{2\sqrt{R_0}}.$$
(31)

Substituting this into (26), we get the alternative impedance-based representation

$$v = Zi + \int_0^{\frac{\pi}{2}} \int_{-\pi}^{\pi} s_{\alpha\beta}^{OC}(\theta, \varphi) a_{\beta}(\theta, \varphi) \sin\theta \, d\varphi \, d\theta, \quad (32)$$

where the impedance matrix is then related to the S-matrix via the relation

$$Z = R_0 (\mathbf{I} - S_{\alpha \alpha})^{-1} (\mathbf{I} + S_{\alpha \alpha}), \tag{33}$$

and the open circuit embedded radiation pattern is related to the terminated radiation pattern as follows

$$\mathbf{s}_{\alpha\beta}^{\mathrm{OC}}(\theta,\varphi) = \frac{1}{\sqrt{R_0}} (\mathbf{Z} + R_0 \mathbf{I}) \, \mathbf{s}_{\alpha\beta}(\theta,\varphi).$$
 (34)

Finding the impedance matrix Z directly from $s_{\alpha\beta}^{\rm OC}$ is also possible [34]. In both cases, however, the embedded patterns are required. Only in the case of canonical minimum scattering antennas [34] is the isolated pattern equivalent to the open-circuit embedded pattern.

The directivity of an antenna is given by the maximum

$$D_{\max} = \max_{\boldsymbol{a}_{\alpha}} \frac{|\boldsymbol{s}_{\alpha\beta}(\boldsymbol{\theta}, \varphi)^{\mathsf{T}} \boldsymbol{a}_{\alpha}|^{2}}{\boldsymbol{a}_{\alpha}^{\mathsf{H}} \boldsymbol{B} \boldsymbol{a}_{\alpha}}.$$
 (35)

This maximization can lead to the case that both the numerator and denominator go to zero but with a substantial value for the ratio, known as the super gain or super directivity. The super gain is in the ideal case unbounded for a given size as it can be increased with denser sub-wavelength sampling. For larger antenna arrays, however, this ratio is limited by the efficiency and ohmic losses of the antenna, and the gain is given by the normal gain, i.e., the physical area of the antenna [35]:

$$D_{\text{normal}} = (ka)^2 + ka \approx \pi a^2 \cdot \frac{4\pi}{\lambda^2}.$$
 (36)

In other words, for large electrical sizes, the effective area of the antenna is given by the physical area which is consistent with the results assumed in the field-based approaches, where $D_{\rm normal}$ is related to the number of degrees of freedom.

This result (36) applies, however, only for the large electrical aperture case (a is a multiple of λ). For smaller electrical apertures (in the order of λ), more degrees of freedom with sub-wavelength spacing could be achieved at the cost of a higher Q-factor. These small-scale compact MIMO systems have been considered in some prior work [36]–[38] and are very relevant for low-frequency systems. Understanding the trade-off between bandwidth, efficiency and the number of independent degrees of freedom is still an open problem and requires the combination of circuit and information theoretical approaches [38].

B. More recent work

1) Radiated power constraints: Circuit models accurately yield the radiated power from antenna arrays and can be used to model electromagnetic exposure constraints. All RF devices emit electromagnetic waves in the non-ionizing spectrum that is absorbed by users. EM absorption induces tissue heating in users, and the level of exposure is highly monitored and regulated by bodies such as the Federal Communications Commission (FCC) [39] and the International Commission on Non-Ionizing Radiation Protection (ICNIRP) [40]. The exposure constraints as a constraint on the radiated power, and the presence of this limit can severely impact spectral efficiency [41], [42]. The exposure constraint differs from the conventional transmit power constraint because of the relationship between the radiated power $P_{\mathrm{T}}^{\mathrm{rad}}$ and the total transmit power $P_{\mathrm{T}}^{\mathrm{total}}$. The radiated power depends on the properties of the antenna array and precoding,

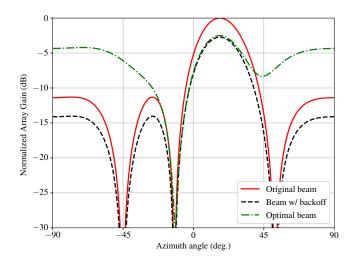


Fig. 8: Example of how the radiated power constraints modeled with the circuit-theoretic approach affect beam patterns. While power back-off reduces the gain of the beam over the entire angular range, the optimal exposure-constrained beam uses a combination of higher sidelobes and a lower main lobe to maximize beamforming gain.

while the total power represents the transmitter power budget. The circuit model demonstrates that constraints on these two quantities are not equivalent and should be considered separately.

The key to designing exposure-aware transmission methods is to incorporate both radiated and total power constraints into system design simultaneously. The easiest method to achieve compliance with regulatory exposure limits is to reduce the total power until the exposure level lies below the limit. This approach, however, neglects the fact that beamforming and precoding can be leveraged to vary the radiated power while keeping the total power constant. A number of studies have shown that including total and radiated power constraints complicates signal processing but results in significant gains over the power backoff method [41], [43], [44]. In Fig. 8, we provide an illustrative example of the effect of radiated power constraints on a system. We show how power backoff reduces the gain pattern of the entire beam. In contrast, the optimal exposure-constrained beam leverages higher sidelobes to increase overall beamforming gain while lying below the exposure limit. Circuit models provide a clear path to modeling both types of power and can be leveraged for advanced signal processing in exposure-constrained systems.

2) Aperture and bandwidth interactions: Spatial wideband effects are exacerbated by large arrays, leading

to tradeoffs between the bandwidth of the system and the aperture of the array. The spatial wideband effect occurs when the worst-case array delay spread is larger than the symbol rate [45]. When this occurs, delays in the time domain induce different phase shifts in the frequency domain, leading to frequency-selective propagation. Beam squint, for example, is a negative effect in phased-array systems that causes the beamforming direction to shift as the signal frequency moves away from the center. For this reason, wideband signals can appear as though they originate from a range of angles rather than a single direction. Increasing either the array aperture or the signal bandwidth increases wideband effects. The results in [46] show that the product of these two quantities, referred to as the aperture-bandwidth product (ABP) is upper bounded by a term that depends on the beamforming gain at a fixed angle. The results in [46] show that as the bandwidth doubles, the maximum array aperture needed to maintain the same beamforming gain halves. Overall, the ABP reveals that the array size cannot be increased beyond a critical point without readjusting the signal design to account for both nearfield and wideband effects.

3) Analog beamforming constraints: Analog beamforming architectures are diverse and each restricts communication in a different manner. Phased arrays can be employed with active or passive phase shifters, each of which introduces different insertion losses and gains into the system. Traditional phased-shifter-based analog precoding architectures employed above 6-GHz yield frequency-flat beamforming, which can cause beam severe squinting at frequencies far from the center frequency, leading to a trade-off between beamforming gain and bandwidth [47]. Instead, frequency-selective analog beamforming can be enabled by applying a time delay, rather than a phase shift, to the signal. The performance of TTD beamforming is highly dependent on the hardware implementation of TTD elements [48]-[50] in terms of metrics like maximum achievable delay, delay-bandwidth product, power consumption, chip area, etc. For example, implementing delay in analog RF consumes more area and power unlike area- and energy-efficient switched-capacitor-based circuits [48] that apply delay in analog baseband. For all phase-shifter and TDD architectures, the underlying signal-processing constraints are notably different.

The circuit model can provide a unified framework for modeling different kinds of beamforming implementations. Each architecture would place different constraints on the impedance matrices in Section III. As demonstrated in the model, different matching networks, losses, noise sources, and arrays are all encompassed

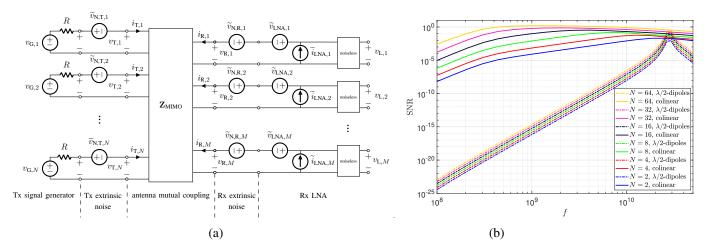


Fig. 9: The MIMO circuit model and the bandwidth gain: (a) Linear multiport model of a radio MIMO communication system showing signal generation, antenna mutual coupling, and noise from both extrinsic (i.e., picked up by the antennas) and intrinsic (i.e., generated by low-noise amplifiers and local circuitry) origins, (b) the MISO SNR with optimum beamforming as a function of the frequency f with $a_R = 100 \, \delta$. Transmit antennas are either Chu's CMS antennas aligned in a colinear configuration (solid lines) with the same fixed spacing $\delta_T = \delta_R = \delta = 0.5 \, [\text{cm}]$ or standard $\lambda/2$ dipole elements, and $P_{\text{max}} = 2 \, [\text{W}]$.

by the model. The common framework would provide a more streamlined approach for signal design and for choosing a beamforming architecture. As an example, an analysis of beamforming using the circuit model could reveal which architecture would be most beneficial for a particular scenario.

4) Broadband MIMO: The performance analysis of MIMO communication links in terms of rate and reliability has historically evolved around the characterization of the conventional multiplexing, antenna, and diversity gains. These three gains, however, disregard the physical constraints on MIMO transceivers. When the EM properties of antenna arrays are incorporated into the circuit-based modeling of MIMO communication systems using multi-port circuit theory (similarly to the SISO case in Section III), a missing gain of MIMO systems, called the bandwidth gain, can be characterized [10]. Fig. 9a depicts a circuit model for MIMO communications with N transmit voltage generators $\mathbf{v}_{G}(f) = \begin{bmatrix} v_{G,1}(f), v_{G,2}(f), \dots, v_{G,N}(f) \end{bmatrix}^{\top}$ and M receive loads $\mathbf{v}_{L}(f) = \begin{bmatrix} v_{L,1}(f), v_{L,2}(f), \dots, v_{L,M}(f) \end{bmatrix}^{\top}$. The joint impedance matrix \mathbf{Z}_{MIMO} models the overall system impedance depending on whether the mutual coupling effects within the transmit and receive arrays are considered or ignored. Here, we rely on the circuit representation based on impedance matrices without loss of generality because impedance, admittance, and scattering matrices are all equivalent representations to any given circuit model [51]. The choice of the circuit representation depends on the type of antenna termination,

namely open circuit, short circuit, or load terminations for the impedance, the admittance, or the scattering parameters, respectively.

After finding the input-output relation of the MIMO circuit model, one maximizes the achievable rate for multiple antenna arrays. As one example, Fig. 9b depicts SNR plots for both Chu's antennas and linear $\lambda/2$ dipole arrays as a practical example of the standard MIMO case. We simulated dipole array impedances in Matlab using the Antenna toolbox package. It is found that the operational bandwidth of colinear Chu's antennas becomes wider as the number of antenna elements increases (or equivalently as the MC becomes higher) leading to an almost constant current over the array aperture. It is also seen that $\lambda/2$ arrays (as in conventional MIMO) do not exhibit any bandwidth gain (i.e., widening) as they resonate around the center frequency (as done by each dipole separately). This is due to their conventional parallel orientation of antenna elements which does not enable strong mutual coupling effects between neighboring elements. It is worth noting that this result challenges the negative perception of mutual coupling, which is typically considered harmful for MIMO communication in the open literature [52]. By carefully modeling mutual coupling using the physically consistent features of circuit-based models, a higher SNR/achievable rate can be achieved.

The benefits of mutual coupling effects stem from the increase of the effective length of the connected antenna array, thereby allowing resonance at larger wavelengths,

i.e., lower frequencies. This is indeed why the bandwidth gain only is observed on the left in Fig. 9b. While this benefit has been demonstrated for linearly polarized and lossless antennas, studying the bandwidth gain for lossy and/or multi-polarization antennas is crucial from an antenna design perspective for future work.

C. Advanced antennas for next generation wireless systems

1) Circuit models for controllable channels and IRS: The development of highly-reconfigurable large apertures will have a significant influence on various areas, including communication, imaging, remote sensing, and space applications. Circuit models for apertures provide a valuable tool for analyzing and understanding the behavior of these complex systems. By representing the aperture as an electrical multiport circuit as shown in Fig. 10, it becomes possible to study its electrical characteristics, performance, and interactions with other components such as termination loads and generators in a systematic manner. A circuit model describes the aperture as a network of interconnected components, such as amplifiers, filters, and switches. It provides a higher level of abstraction and allows for the analysis of signal flow, power distribution, and control mechanisms within the aperture system.

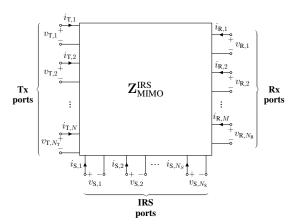


Fig. 10: Equivalent circuit-theoretic model for IRS-assisted MIMO communication channels.

Most research studies assess the performance of IRS-aided communication systems by assuming a linear relationship between the received signal and the scattering parameters of the IRS. When mutual coupling is taken into account, the scattering parameters of the IRS cannot be represented as a diagonal matrix, and a full matrix description is required to capture the interactions between the element ports. As a result, the relationship between the received signal and the scattering parameters of the IRS becomes highly non-linear, requiring a sophisticated

tuning of the IRS elements even with known CSI. In this context, circuit theory provides more accurate scattering models for imperfect phase shifts, while information theory can account for the non-linearity of the problem by characterizing new approximations and bounds.

Another line of work [53] suggests using IRS to directly embed data by manipulating the tunable reactance on a symbol-by-symbol basis. This framework is known as space-time IRS and allows the IRS to perform information embedding through backscatter communication. Analyzing and optimizing the information transfer of such a system requires a careful circuit modeling of the interactions within the IRS elements and the tunable component, which generally leads to a highly nonlinear nonstandard channel.

The dynamic nature of highly reconfigurable antennas, with their ability to adjust radiation properties in real-time, introduces new challenges and opportunities in information transmission. The traditional assumptions of uncontrolled channels and interference environments may no longer hold, requiring non-standard information theory approaches to account for the dynamic and controllable nature of the reconfigurable antenna systems.

- 2) Communication, sensing, and localization: The integration of traditional communication and sensing systems requires an understanding of the trade-off between the power and bandwidth of a full-duplex system. For better illustration, consider a full-duplex base station both localizing and communicating with an unmanned aerial vehicle (UAV) equipped with an IRS. The knowledge of the scattering behavior of the IRS by the base station is key to localizing the UAV. Moreover, the IRS can also be used to embed/modulate information for back-scattering communication. The characterization of the scattering properties of the IRS elements can be performed using circuit-based models so as to characterize the imperfection of the IRS elements and the mutual coupling effects between the IRS elements. For this reason, finding the optimal trade-off between the achievable rate of communication and the accuracy of sensing is a problem that lies at the interface of communication, estimation, and antenna theories.
- 3) Near-field MIMO: Circuit models offer accurate models to describe the antenna parameters whose control is critical for reactive and radiative near-field communication scenarios. The near-field of antenna arrays can extend far beyond physically close distances from the array depending on the operating frequency and the array aperture. The Fraunhofer distance, which predicts the radiated near-field boundary of an antenna, gives distances at around 100m away from the array at mmWave frequencies for typical array apertures [54]. By means of

appropriate antenna terminations, more realistic models of mutual coupling can be leveraged to study near-field MIMO systems [10], [55]. Circuit models could be leveraged to provide a better understanding of near-field MIMO communication systems because the extrinsic noise of transmit and receive antenna ports are coupled [11]. In other words, the modeling of the mutual coupling effects in the near-field region must capture the coupling within and *across* antenna arrays. This calls for the need to adopt the circuit-based modeling approach to design circuit components for both communication and radiation purposes.

RF chains consist of a number of digital and analog components that process incoming and outgoing signals. These include RF mixers, digital-to-analog converters (DACs) and analog-to-digital converters (ADCS), matching networks, and amplifiers. Matching networks, for example, are critical in decoupling antenna ports and ensuring power is properly radiated. DACs and ADCs are another set of components that can significantly affect communication depending on the quantization resolution. The circuit model can also be applied to model coupling between the arrays. In a way, current assumptions about the system assume ideal processing before transmission and after reception. Doing so, however, ignores the implications of hardware limitations on the near-field communication performance.

4) The THz challenge: Next-generation wireless networks demand high data rates and require reliable connectivity and large bandwidths in the order of tens of GHz. Accordingly, the terahertz (THz) frequency band (0.1-10 THz) is becoming an attractive band that can provide over 20 GHz of bandwidth for realizing such ultra-high data rates [56], [57]. At such frequencies, however, the efficiency of circuit components is heavily reduced while significant beamforming gain in the direction of interest is required. Incorporating the RF circuit components, particularly the power amplifier in the capacity analysis is even more critical.

IRS can be also used at THz frequencies to mitigate signal blockage vulnerability and reshape the wireless environment at a low cost. In [58], a passive IRS that operates at THz frequencies composed of three layers, a conductive top layer, a middle layer composed of a dielectric with tunable dielectric properties, and a bottom dielectric layer is presented. The reflective surface controls the phase of the reflected wave from each unit cell. By using a voltage source to control the dielectric properties of the tunable substrate, beamforming becomes attainable. The development of circuit models along with the information-theoretic analysis for this type of dielectric IRS at THz frequencies is a crucial

area of research.

V. CONCLUSION

This magazine paper reviews the advancements in the physically consistent modeling of MIMO antennas and their impact on information theory. It explores the integration of information theoretical concepts with equivalent antenna circuit models, aiming to enhance the overall performance of future wireless systems. Equivalent circuit models can accurately model the physical behavior of modern antenna structures based on principles as fundamental as the conservation of energy, while information theory provides the tools to analyze, optimize, and exploit their full potential. Additionally, it investigates the utilization of broadband equivalent circuits in modern highly reconfigurable large aperture designs, introducing new challenges and opportunities in information transmission.

VI. ACKNOWLEDGMENT

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery Grants Program, in part by Research Manitoba through New Investigator Operating Grant, in part by the National Science Foundation under Grant NSF-ECCS- 2153698, Grant NSF-CCF-2225555, and Grant NSF-CNS- 2147955, in part by Federal Agency and industry partners through the Resilient & Intelligent NextG Systems (RINGS) program, in part by Qualcomm, in part by Nokia Bell Labora- tories, and in part by Samsung.

REFERENCES

- J. L. Massey, Applied Digital Information Theory 1: Lecture Notes. Zurich, Switzerland: ETH Zurich, 1998, [online] www.isiweb.ee.ethz.ch/archive/massey_scr/adit1.pdf.
- [2] M. Franceschetti, Wave Theory of Information. Cambridge University Press, 2018.
- [3] L. Stankovic, D. Mandic, M. Dakovic, M. Brajovic, B. Scalzo, S. Li, and A. G. Constantinides, "Graph signal processing – part iii: Machine learning on graphs, from graph topology to applications," 2020. [Online]. Available: https://arxiv.org/abs/ 2001.00426
- [4] H.-A. Loeliger, F. Lustenberger, M. Helfenstein, and F. Tarkoy, "Probability propagation and decoding in analog vlsi," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 837–843, 2001.
- [5] L. J. Chu, "Physical limitations of omni-directional antennas," Journal of applied physics, vol. 19, no. 12, pp. 1163–1175, 1948.
- [6] R. M. Fano, "Theoretical limitations on the broadband matching of arbitrary impedances," *Journal of the Franklin Institute*, vol. 249, no. 1, pp. 57–83, 1950.
- [7] D. Gabor, "Communication theory and physics," *Transactions of the IRE Professional Group on Information Theory*, vol. 1, no. 1, pp. 48–59, 1953.

- [8] M. Franceschetti, M. D. Migliore, and P. Minero, "The capacity of wireless networks: information-theoretic and physical limits," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3413–3424, 2009. [Online]. Available: https://doi.org/10.1109/TIT.2009.2023705
- [9] A. Pizzo, L. Sanguinetti, and T. L. Marzetta, "Fourier planewave series expansion for holographic MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 6890– 6905, Sep. 2022.
- [10] M. Akrout, V. Shyianov, F. Bellili, A. Mezghani, and R. W. Heath, "Super-wideband Massive MIMO," *IEEE Journal on Selected Areas in Communications*, 2023.
- [11] M. T. Ivrlač and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Transactions on Circuits and Systems I:* Regular Papers, vol. 57, no. 7, pp. 1663–1683, 2010.
- [12] D. Nie and B. M. Hochwald, "Improved Broadband Matching Bound," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 11, pp. 5878–5885, 2017.
- [13] S. Phang, M. T. Ivrlač, G. Gradoni, S. C. Creagh, G. Tanner, and J. A. Nossek, "Near-Field MIMO Communication Links," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 9, pp. 3027–3036, 2018.
- [14] N. Deshpande, M. R. Castellanos, S. R. Khosravirad, J. Du, H. Viswanathan, and R. W. Heath, "A wideband generalization of the near-field region for extremely large phased-arrays," *IEEE Wireless Communications Letters*, pp. 1–1, 2022.
- [15] M. Akrout, V. Shyianov, F. Bellili, A. Mezghani, and R. W. Heath, "Achievable Rate of Near-Field Communications Based on Physically Consistent Models," *IEEE Transactions on Wireless Communications*, vol. 22, no. 2, pp. 1266–1280, 2023.
- [16] V. Shyianov, M. Akrout, F. Bellili, A. Mezghani, and R. W. Heath, "Achievable rate with antenna size constraint: Shannon meets chu and bode," *IEEE Transactions on Communications*, vol. 70, no. 3, pp. 2010–2024, Mar. 2022.
- [17] H. Wang, N. Shlezinger, Y. C. Eldar, S. Jin, M. F. Imani, I. Yoo, and D. R. Smith, "Dynamic metasurface antennas for MIMO-OFDM receivers with bit-limited ADCs," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2643–2659, Apr. 2021.
- [18] C. E. Shannon, "A mathematical theory of communication," The Bell System Technical Journal, vol. 27, no. 3, pp. 379–423, 1948.
- [19] R. M. Fano, Theoretical limitations on the broadband matching of arbitrary impedances. Technical Report NO. 41, Research Laboratory of Electronics, MIT, 1948.
- [20] H. W. Bode, Network analysis and feedback amplifier design. Van Nostrand, New York, 1945.
- [21] D. F. Sievenpiper, D. C. Dawson, M. M. Jacob, T. Kanar, S. Kim, J. Long, and R. G. Quarfoth, "Experimental Validation of Performance Limits and Design Guidelines for Small Antennas," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 1, pp. 8–19, 2012.
- [22] M. Gustafsson, C. Sohl, and G. Kristensson, "Physical Limitations on Antennas of Arbitrary Shape," *Proceedings: Mathematical, Physical and Engineering Sciences*, vol. 463, no. 2086, pp. 2589–2607, 2007.
- [23] S. A. Schelkunoff and H. T. Friis, Antennas: theory and practice. Wiley New York, 1952, vol. 639.
- [24] F. Baccelli and B. Blaszczyszyn, "Stochastic Geometry and Wireless Networks: Volume I Theory," *Foundations and Trends in Networking*, vol. 3, no. 3–4, pp. 249–449, 2010. [Online]. Available: http://dx.doi.org/10.1561/1300000006
- [25] R. W. Heath Jr, T. Wu, Y. H. Kwon, and A. C. Soong, "Multiuser MIMO in distributed antenna systems with outof-cell interference," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4885–4899, 2011.

- [26] H. L. Van Trees, Optimum array processing: Part IV of detection, estimation, and modulation theory. John Wiley & Sons, 2002.
- [27] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2526–2537, Nov. 2020.
- [28] E. Björnson and L. Sanguinetti, "Power scaling laws and near-field behaviors of massive MIMO and intelligent reflecting surfaces," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1306–1324, Sep. 2020.
- [29] M. Jensen and J. Wallace, "A review of antennas and propagation for MIMO wireless communications," *IEEE Transactions* on Antennas and Propagation, vol. 52, no. 11, pp. 2810–2824, 2004
- [30] M. T. Ivrlač and J. A. Nossek, "Toward a circuit theory of communication," *IEEE Transactions on Circuits and Systems I:* Regular Papers, vol. 57, no. 7, pp. 1663–1683, Jul. 2010.
- [31] J. A. Nossek, "On the relation of circuit theory and signals, systems and communications," AEU-International Journal of Electronics and Communications, vol. 65, no. 8, pp. 713–717, 2011.
- [32] D. M. Kerns and E. S. Dayhoff, "Theory of diffraction in microwave interferometry," *Journal of Research of the National Bureau of Standards Section B Mathematics and Mathematical Physics*, p. 1, 1960.
- [33] A. Mezghani, F. Bellili, and E. Hossain, "Reconfigurable Intelligent Surfaces for Quasi-Passive mmWave and THz Networks: Should They be Reflective or Redirective?" in 2022 56th Asilomar Conference on Signals, Systems, and Computers, 2022, pp. 1076–1080.
- [34] W. Wasylkiwskyj and W. K. Kahn, "Theory of mutual coupling among minimum-scattering antennas," *IEEE Transactions on Antennas and Propagation*, vol. 18, no. 2, pp. 204–216, 1970.
- [35] R. F. Harrington, "Effect of antenna size on gain, bandwidth, and efficiency," J. Res. Nat. Bur. Stand, vol. 64, no. 1, pp. 1–12, 1960.
- [36] M. T. Ivrlač and J. A. Nossek, "On multistreaming with compact antenna arrays," in 2011 International ITG Workshop on Smart Antennas, 2011, pp. 1–8.
- [37] D. Nie and B. M. Hochwald, "Bandwidth Analysis of Multiport Radio Frequency Systems-Part I," *IEEE Transactions on Anten*nas and Propagation, vol. 65, no. 3, pp. 1081–1092, 2017.
- [38] S. Saab, A. Mezghani, and R. W. Heath, "Optimizing the Mutual Information of Frequency-Selective Multi-Port Antenna Arrays in the Presence of Mutual Coupling," *IEEE Transactions* on Communications, vol. 70, no. 3, pp. 2072–2084, 2022.
- [39] FCC, "Evaluating compliance with FCC guidelines for human exposure to radiofrequency electromagnetic fields," Tech. Rep. OET Bull. 65, Suppl. C, ed. 01-01, Jun. 2001.
- [40] ICNIRP, "Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz)," *Health Phys.*, vol. 118, no. 5, pp. 483–524, May 2020.
- [41] D. Ying, D. J. Love, and B. M. Hochwald, "Sum-rate analysis for multi-user MIMO systems with user exposure constraints," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7376–7388, Sep. 2017.
- [42] B. M. Hochwald, D. J. Love, S. Yan, and J. Jin, "SAR codes," in *UCSD Information Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2013, pp. 1–9.
- [43] M. R. Castellanos, D. Ying, D. J. Love, B. Peleato, and B. M. Hochwald, "Dynamic electromagnetic exposure allocation for Rayleigh fading MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 728–740, Feb. 2021.
- [44] F. Héliot, M. A. Jamshed, and T. W. C. Brown, "Exposure modelling and minimization for multi-antenna communication

- systems," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC Spring)*, May 2020, pp. 1–6.
- [45] M. Abdelghany, U. Madhow, and M. Rodwell, "An efficient digital backend for wideband single-carrier mmwave massive mimo," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
- [46] N. Deshpande, M. R. Castellanos, S. R. Khosravirad, J. Du, H. Viswanathan, and R. W. Heath, "A wideband generalization of the near-field region for extremely large phased-arrays," *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 515–519, Mar. 2023.
- [47] A. Mezghani and R. W. Heath, "The Information and Wave-Theoretic Limits of Analog Beamforming," in 2018 Information Theory and Applications Workshop (ITA), 2018, pp. 1–6.
- [48] E. Ghaderi, A. Sivadhasan Ramani, A. A. Rahimi, D. Heo, S. Shekhar, and S. Gupta, "An Integrated Discrete-Time Delay-Compensating Technique for Large-Array Beamformers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 9, pp. 3296–3306, Sep. 2019.
- [49] E. Zolkov, A. Madjar, R. Weiss, and E. Cohen, "Analysis and Design of *N* -Path True-Time-Delay Circuit," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 12, pp. 5381–5394, Dec. 2020.
- [50] D. I. Lialios, N. Ntetsikas, K. D. Paschaloudis, C. L. Zekios, S. V. Georgakopoulos, and G. A. Kyriacou, "Design of True Time Delay Millimeter Wave Beamformers for 5G Multibeam Phased Arrays," *Electronics*, vol. 9, no. 8, p. 1331, Aug. 2020.
- [51] D. M. Pozar, Microwave engineering. John wiley & sons, 2011.
- [52] X. Chen, S. Zhang, and Q. Li, "A review of mutual coupling in MIMO systems," *IEEE Access*, vol. 6, pp. 24706–24719, 2018.
- [53] H. U. Rehman, F. Bellili, A. Mezghani, and E. Hossain, "Modulating Intelligent Surfaces for Multiuser MIMO Systems: Beamforming and Modulation Design," *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 3234–3249, 2022.
- [54] D. Dardari and N. Decarli, "Holographic communication using intelligent surfaces," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 35–41, Jun. 2021.
- [55] T. S. Bird, Mutual Coupling Between Antennas. John Wiley & Sons. 2021.
- [56] H.-J. Song and N. Lee, "Terahertz Communications: Challenges in the Next Decade," *IEEE Transactions on Terahertz Science* and Technology, vol. 12, no. 2, pp. 105–117, 2022.
- [57] S. Saab, D. Kim, S. Abu-Surra, G. Xu, and J. Zhang, "Path Towards Tbps Communications: LoS MIMO Theory, Simulation and Measurement Analysis," in 2022 IEEE Globecom Workshops (GC Wkshps), 2022, pp. 1796–1801.
- [58] S. Saab, S. Abu-Surra, G. Xu, and R. W. H. Jr., "An inverted l-shaped multi-layer reconfigurable intelligent surface for thz communications," in 56th Asilomar Conference on Signals, Systems, and Computers, ACSSC 2022, Pacific Grove, CA, USA, October 31 - Nov. 2, 2022. IEEE, 2022, pp. 1060–1063.