An Overview of Signal Processing Techniques for Terahertz Communications

Hadi Sarieddeen, Member, IEEE, Mohamed-Slim Alouini, Fellow, IEEE, and Tareq Y. Al-Naffouri, Senior Member, IEEE

Abstract—Terahertz (THz)-band communications are a key enabler for future-generation wireless communications systems that promise to integrate a wide range of data-demanding applications. Recent advances in photonic, electronic, and plasmonic technologies are closing the gap in THz transceiver design. Consequently, prospect THz signal generation, modulation, and radiation methods are converging, and corresponding channel model, noise, and hardware-impairment notions are emerging. Such progress establishes a foundation for well-grounded research into THz-specific signal processing techniques for wireless communications. This tutorial overviews these techniques, emphasizing ultra-massive multiple-input multiple-output (UM-MIMO) systems and reconfigurable intelligent surfaces, vital for overcoming the distance problem at very high frequencies. We focus on the classical problems of waveform design and modulation, beamforming and precoding, index modulation, channel estimation, channel coding, and data detection. We also motivate signal processing techniques for THz sensing and localization.

Index Terms—THz communications, signal processing, ultra-massive MIMO, intelligent reflecting surfaces, terabit-per-second.

I. INTRODUCTION

A. THz Communications are Emerging

The wireless communication’s frequency spectrum has been continuously expanding to satisfy ever-increasing bandwidth demands. Although millimeter-wave (mmWave)-band communications [1], [2] are already shaping the fifth-generation (5G) of wireless mobile communications, terahertz (THz)-band communications [3]–[8] will be essential to the future sixth-generation (6G) [9]–[21] and beyond. The THz band is sandwiched between the microwave and optical bands as the last unexplored area of the radio-frequency (RF) spectrum. Hence, technologies from both sides are being explored to support THz communications. RF engineers label THz all operations beyond the 100 gigahertz (GHz) threshold, below which most known mmWave use cases exist. In contrast, optical engineers label THz any frequency below 10 THz (the far-infrared). However, the THz range is 300 GHz to 10 THz according to IEEE Transactions on Terahertz Science and Technology and closely mapped to the tremendously high frequency (THF) band (300 GHz to 3 THz) according to ITU-R.

THz communications are called-for despite the maturity of neighboring technologies. In contrast to mmWave communications, THz communications can exploit the available spectrum to achieve a terabits per second (Tbps) data rate without additional spectral efficiency enhancement techniques. Furthermore, due to the shorter wavelengths, THz signals are less susceptible to free-space diffraction and inter-antenna interference and exhibit higher resilience to eavesdropping. THz systems can also support higher link directionality and can be achieved in much smaller footprints. In contrast, compared to visible light communications (VLC) [22], [23], THz signals are not as severely affected by alignment issues, ambient light, atmospheric turbulence, scintillation, fog, and temporary spatial variation in light intensity. THz communications can thus complement both mmWave and VLC by providing alternative quasi-optical paths. However, due to significant water vapor absorption above 1 THz, a gap might always exist for wireless communications at the high end of the THz range.

Consequently, THz communications will enable ultra-high bandwidth THz communications are thus expected to enable ultra-high bandwidth, and ultra-low latency communication paradigms [24]. For example, they can be used to achieve optical-fiber-like performance in network backhauling [17], backbone (rack-to-rack) connectivity in data centers [25]–[27], and high data rate kiosk-to-mobile communications [28]. Furthermore, THz wireless bridges enable the transparent integration of fiber networks without requiring detection, decoding, and re-modulation [29]. THz links supporting 100 gigabits per second (Gbps) have already been demonstrated over distances corresponding to such applications [30]. However, the holy grail of THz communications is enabling mobile communications at the device level and, in the context of medium-range indoor, vehicular, drone-to-drone [31], [32], or device-to-device communications, at the access level. When combined with other THz-band applications such as accurate localization, sensing, and imaging, THz communications can enable wireless remoting of human cognition, leading to ubiquitous wireless intelligence [24], [33].

B. Advances in THz Devices

The main contributions to THz technology are still at the device level rather than the system level. High-frequency electromagnetic radiation is perceived either as waves processed via electronic devices (the mmWave realm) or as particles processed via photonic devices (the optical realm). In between, the THz band is dubbed as a “THz gap” due to the lack...
of compact THz signal sources and detectors that have high power and sensitivity, respectively. However, recent electronic and photonic THz transceiver design advances have enabled efficient signal generation, modulation, and radiation [30], [34]–[37].

In electronic solutions [38], [39], silicon-based devices [40], [41] successfully used for mmWave systems are being advocated for the THz range. Electronic solutions are based primarily on silicon complementary metal-oxide-semiconductor (CMOS) and silicon-germanium (SiGe) BiCMOS technologies [42]–[45], and have demonstrated incredible compactness and compatibility with existing fabrication processes. CMOS solutions exhibit lower speed and power handling capabilities, but are the most compact. However, the corresponding highest unity current gain frequency ($f_T$) and unity maximum available power gain frequency ($f_{max}$) remain at 200 GHz and 320 GHz, respectively. Nevertheless, higher operating frequencies have been noted by III-V-based (Indium phosphide (InP)-based [46], for example) semiconductors [47] in high electron mobility transistors (HEMTs) [48]–[50], heterojunction bipolar transistors (HBTs) [46], [51], and Schottky diodes [52]. The $f_{max}$ of III-V based devices have crossed 1.5 THz [37], and in [45], SiGe-HBT devices could operate at 720 GHz. Resonant tunneling diodes (RTDs) have also exhibited oscillation frequencies up to 2 THz [53]–[56]. Other demonstrations at different frequencies and powers are highlighted in [57]–[60].

In photonic solutions [30], where the main design driver is the data rate, higher carrier frequencies are supported, but the degrees of integration and output power remain low (photonic devices have a larger form factor and are more expensive). Frequencies beyond 300 GHz have been supported using optical downconversion systems [30], quantum cascade lasers [61], photoconductive antennas [62], and uni-traveling carrier photodiodes [63]. Among other advances in THz photonics, knowledge on the topological phase of light is being exploited to demonstrate robust THz topological valley transport through several sharp bends in on-chip THz waveguides [64].

Satisfying emerging system-level properties requires designing efficient and programmable devices. This deviation from designing perfect THz devices resulted in the emergence of integrated hybrid electronic-photonic systems [35], such as combining photonic transmitters and III-V electronic receivers. However, more precise synchronization between the transmitter and receiver is required in hybrid solutions. Integrating miniature micro-electro-mechanical systems (mMEMS) within THz components also promises high degrees of reconfigurability and enhanced performance. By enabling electronic control of frequency, polarization, and beam steering, mMEMS can achieve features beyond what can be achieved in state-of-the-art optical and semiconductor technologies.

Also gaining popularity are plasmonic solutions [65], [66], where novel plasmonic materials such as graphene exhibit high electron mobility and reconfigurability [67]–[70]. The resultant surface plasmon polariton (SPP) waves in plasmonic antennas have much smaller resonant wavelengths than free space waves, which results in compact and flexible antenna array designs [71]. By leveraging the properties of plasmonic nanomaterials and nanostructures, transceivers and antennas that intrinsically operate at THz frequencies can be created, avoiding the upconversion and downconversion losses of electronic and photonic systems, respectively. Graphene can be used to develop direct THz signal sources, modulators (that manipulate amplitude, frequency, and phase), and on-chip THz antenna arrays [72]. Graphene solutions promise to have low electronic noise temperature and generate energy-efficient short pulses [73]. Alongside miniaturization, graphene antennas also promise high directivity and radiation efficiency [74]. All these advances demonstrate that the gap between designing THz technology is rapidly closing and that the THz-band will soon open for everyday applications.

Power consumption remains a significant hurdle toward the practical deployment of THz systems. Although studies on the power consumption of mmWave and sub-THz receivers are emerging, as reported in [75], [76], the power consumption of true-THz devices is still not well-established. For fully digital architectures, [75] proposes design options for mixers, noise amplifiers, local oscillators, and analog-to-digital converters (ADCs) by assuming a 90 nm SiGe BiCMOS technology. The corresponding optimizations can reduce power by 80% (fully-digital 140 GHz receiver with a 2 GHz sampling rate and power consumption less than 2 W). Given that cellular systems do not usually require high output signal-to-noise-ratio (SNRs), the researchers propose relaxing linearity requirements and reducing the power consumption of components. Low-resolution ADCs and mixer power consumption are significantly reduced, and fully digital architectures introduce significant benefits compared to analog beamforming with comparable power consumption under bit-width optimization. For a hybrid array-of-subarrays (AoSA) multiple-input multiple-output (MIMO) structure at 28 GHz and 140 GHz, and assuming a single-carrier (SC) system, the power dissipated by two major power-consuming components, namely the low noise amplifier (LNA) and the ADC, is estimated in [76]. The power consumption of a 140 GHz receiver is much higher than that of a 28 GHz receiver, indicating that considerable measures need to be taken to improve power consumption in high-frequency devices.

C. Standardization Efforts

Various leading 6G initiatives investigate THz communications, including the 6Genesis Flagship Program (6GFP) in Finland and the European Commission’s Horizon 2020 ICT-09 THz Project Cluster. The U.S. Defense Advanced Research Projects Agency ( DARPA) identifies THz technology as one of four major research areas that could impact society more than the Internet. Accordingly, THz-related research has attracted significant funding, and standardization efforts have been launched [77]–[80], [80]. The Federal Communications Commission (FCC) has recently allocated 21.2 GHz of spectrum between the 116 GHz and 246 GHz bands for unlicensed usage. Furthermore, the first IEEE 802.15.3d standardization efforts for sub-THz communications towards 6G were reported in [81], focusing on point-to-point links that can support 100 Gbps over few centimeters to several hundreds of meters.

The first IEEE standard, 802.15.3d, was approved in 2017, with channels in the sub-THz range of 253-322 GHz and
bandwidths up to 69 GHz (starting from 2.16 GHz), featuring up to 69 overlapping channels between 252.72 GHz and 321.84 GHz. All these bands were cleared for THz use at the World Radiocommunication Conference (WRC) 2019 (160 GHz of spectrum), without specifying the necessary conditions to protect passive services such as the radio astronomy earth exploration-satellite service [82]. However, the potential interference with existing ground-to-orbit links should not be overlooked. The standard includes medium access (MAC) and physical (PHY) layer designs. MAC considerations include initial solutions of directional channel access, neighbor discovery, and synchronization. More relevant to our work, the PHY layer considerations are in two modes, THz SC mode (THz-SC PHY) and THz on-off keying mode (THz-OOK PHY). THz-SC PHY supports high rates in fronthaul and backhaul links and data centers, where complex signals are used, including quadrature amplitude modulation (QAM) up to 64-QAM. THz-OOK PHY is more tailored to short-distance (up to several meters) kiosk and intra-device communications, featuring lower-cost sub-THz devices and a single low-complexity modulation scheme, on-off keying (OOK), that can still achieve tens of Gbps. Both high-rate (14/15) and low-rate (11/15) low-density parity-check (LDPC) codes are enabled in both modes, but with THz-OOK PHY, an additional Reed Solomon (RS) code with simple decoding without soft decision information is also enabled. The target minimal receiver sensitivity levels in both modes are set to –67 dBm (for 11/15 LDPC and 2.16 GHz bandwidth).

IEEE 802.15.3d-compliant solutions are already being proposed, as in [83], with THz waveform designs that exploit 99.3% of the total in-band energy and enable out-of-band interference management. In this paper, we expand on the tradeoffs between SC and multi-carrier designs to enhance future standards. We address signal processing solutions suitable for both modes of the standard, and we formulate signal processing problems that require further consideration in future THz standardization efforts. Because IEEE 802.15.3d is tailored to fixed point-to-point links, we define a generic system model that is quasi-deterministic and maintain the discussion at the link level, by assuming the transmit and receive antenna directions are known. We assume interference to be mitigated by appropriate link planning with simple procedures for initial access and device discovery. However, we also address the system-level considerations of interference, link mobility, and multiple-channel access. The stationarity in point-to-point links introduces significant complexity reductions, much needed for Tbps processing, as we highlight in the subsequent sections.

D. Significance of Signal Processing for THz Communications

Many signal processing and communications system challenges still need to be addressed. The factors to be considered in signal processing in the THz realm differ significantly from those in the systems at lower frequencies; they are closely linked to the transceiver or device architectures. Efficient THz-band signal processing is crucial for two reasons. First, signal processing must account for the use of ultra-massive MIMO (UM-MIMO) antenna systems [76], [84], [85] to overcome the very short communication distances due to severe power limitations and propagation losses. Second, signal processing must overcome the mismatch between the bandwidths of the THz channel and the digital baseband system [86], [87].

Because channel coding is the most computationally demanding component of the baseband chain, several projects are studying efficient coding schemes for Tbps operations [88]. However, the complete chain should be efficient and parallelizable. Therefore, joint algorithm and architecture co-optimization of channel estimation, channel coding, and data detection is required. Furthermore, the inherent sparsity in the angle and delay domains at THz frequencies can be exploited in solutions based on compressive sensing techniques. Low-resolution digital-to-analog conversion systems can also reduce the baseband complexity; even all-analog THz solutions have been considered.

Although THz communications exhibit quasi-optical traits, they retain several microwave characteristics. They can still use UM-MIMO antenna array processing techniques to support efficient beamforming and reflective surfaces to support non-line-of-sight (NLoS) propagation. Thus, efficient beamforming and beamsteering techniques and low-complexity precoding and combining algorithms are required. However, what seems predicted to be the norm in future THz systems are hybrid and adaptive AoSA antenna architectures, in which each subarray (SA) undergoes independent beamforming. Furthermore, given the large degrees of freedom in THz UM-MIMO systems at the transmitter side, several probabilistic shaping and index modulation schemes can be explored. This is particularly true in plasmonic solutions where each antenna element (AE) can be turned on and off or tuned to a specific frequency by simple material doping or electrostatic bias.

Molecular absorptions also result in band splitting and spectrum shrinking at larger communication distances. Distance-adaptive solutions in which antenna array designs and resource allocation criteria are optimized can tackle spectrum shrinking [89], [90]. Accordingly, the classical problems of waveform design and modulation need to be revisited. For instance, SC modulations can be favored over orthogonal frequency-division multiplexing (OFDM), which is challenging to implement in the THz band. Nevertheless, in some indoor THz scenarios, several multipath components might persist, resulting in frequency-selective channels. Frequency-selectivity might also arise at the receiver side also due to the behavior of THz components. Therefore, multi-carrier modulations might still be required, perhaps in the form of multiple orthogonal and independent SCs that can be achieved at low complexity and combined with a form of carrier aggregation.

E. Paper Contributions and Outline

In this paper, we provide an overview of recent advances in signal processing techniques for THz communications. It is unclear whether recent advances in THz devices have bridged the THz gap, primarily because there remains a gap between the promised data rates and the limited baseband-processing capabilities. Therefore, we seek to establish a clear link
TABLE I: Relevant overview articles on THz communications.

| Ref.                          | Area of Focus                                                                 |
|-------------------------------|-------------------------------------------------------------------------------|
| Kurner et al. [7] - 2014      | Focuses on THz channel modeling and standardization efforts.                  |
| Akyildiz et al. [4] - 2014    | Focuses on THz applications at the nano and macro levels and presents possible THz band transceivers and antennas; also summarizes challenges in the physical layer. |
| Nagatsuma et al. [30] - 2016  | Overviews recent THz advances at the device level.                             |
| Boulogeorgos et al. [91] - 2018 | Highlights how THz can deliver optical network quality; also touches on the channel, baseband challenges, and on beamforming. |
| Akyildiz et al. [92] - 2018   | Highlights the importance of MIMO configurations and distance-adaptive solutions to extend the THz communication distance. |
| Rappaport et al. [24] - 2019  | Details the opportunities and challenges in communications above 100 GHz towards 6G, mainly shaped by recent channel measurement campaigns led by the group. |
| Ghafoor et al. [93] - 2019    | Presents THz-band applications and various MAC-layer protocols.               |
| Lemic et al. [94] - 2019      | Focuses on THz communications for nanonetworks and the corresponding applications and networking considerations, especially in in-body environments. |
| Elayan et al. [79] - 2020     | Presents a timeline of THz research and funding standardization activities, comparing THz to rival technologies. |
| Sarieddeen et al. [21] - 2020 | Highlights the importance of merging communications, sensing, imaging, and localization in the THz band. |
| Tripathi et al. [80] - 2021   | Captures the physical layer differences between mmWave and THz communications. |
| Chaccour et al. [95] - 2021   | Describes joint THz communications and sensing, focusing on the possible applications and touching on both networking and physical layer considerations. |

between novel THz devices, THz-channel and -noise models, and THz signal-processing techniques. Accordingly, we strive to formulate THz-specific signal processing methods that can alleviate the quasi-optical behavior of THz signals to provide seamless connectivity and accurate localization and sensing capacities.

Samples of popular recent surveys and tutorials on THz communications in the literature are illustrated in Table I including notable papers written by pioneers in the field. The literature lacks a holistic approach to THz signal processing for communications and sensing, a nascent research field not well discovered but with a promising outlook. Papers by Rappaport et al. are more tailored to channel measurement and modeling campaigns and channel sounders; those by Kurner et al. are focused on channel modeling and standardization; those by Akyildiz et al. discuss THz signal processing but not in detail; and those by Nagatsuma et al. are focused on advances in THz devices.

We start our paper by examining how THz technology is widely believed to be the ultimate spectrum enabler for future-generation wireless communications systems. We then describe how the critical challenges in developing THz communications systems are, besides those inherent at the device level, at the infrastructure and algorithmic levels. We promote solutions on both levels, but most importantly, highlight the practical signal processing and hardware considerations critical to achieving the promised Tbps data rates, especially given the limitations in digital baseband solutions. The proposed signal-processing solutions should be low-cost and low-power-consuming. They should also be highly reliable, low-latency solutions.

Approaching THz research from a signal-processing perspective is timely and should not wait until THz devices are mature. While we continue to monitor the latest THz devices and channel models to update our signal-processing solutions, this paper aims to study the signal-processing challenges and prospect THz use cases to guide research on THz transceivers. This paper’s novel contributions are thus expected to catalyze THz research by directly linking its components to broadly and significantly enable future 6G wireless systems to achieve an order-of-magnitude increase in capacity with improved robustness.

The paper is organized as follows. The system model is first presented in Sec. II followed by a discussion on THz channel and noise modeling in Sec. III. Then, recent performance analysis frameworks and experimental testbeds are summarized in Sec. IV. The latest advances in THz modulation and waveform designs are summarized in Sec. V the concept of THz “spatial tuning” and reconfigurable arrays are studied in Sec. VI and THz beamforming and precoding techniques are discussed in Sec. VII. The baseband signal processing problems of THz channel estimation, channel coding, and data detection are then illustrated in Sec. VIII. We examine the signal processing aspects of intelligent reflecting surface (IRS)-assisted THz communications in Sec. IX. Finally, in Sec. X-B, we highlight the importance of signal processing for THz sensing, imaging, and localization, briefly discussing THz networking and security before concluding in Sec. XI.
TABLE II: Frequently-Used Abbreviations

| Abbreviation | Definition |
|--------------|------------|
| 3D           | three-dimensional |
| 5G           | fifth-generation |
| 6G           | sixth-generation |
| ADC          | analog-to-digital converter |
| AE           | antenna element |
| AoA          | angle of arrival |
| AoD          | angle of departure |
| AoSA         | array-of-subarrays |
| AWGN         | additive white Gaussian noise |
| CFO          | carrier frequency offset |
| CMOS         | complementary metal-oxide-semiconductor |
| CPM          | continuous phase modulation |
| CSI          | channel state information |
| DAC          | digital-to-analog converter |
| DAoSA        | dynamic array-of-subarrays |
| Gbps         | gigabit/second |
| GHz          | gigahertz |
| GSM          | generalized spatial modulation |
| GVD          | group velocity dispersion |
| HITRAN       | high-resolution transmission molecular absorption database |
| IM           | index modulation |
| IRS          | intelligent reflecting surface |
| ISI          | inter-symbol interference |
| LDPC         | low-density parity-check |
| LIS          | large intelligent surface |
| LLR          | log-likelihood-ratio |
| LoS          | line-of-sight |
| MIMO         | multiple-input multiple-output |
| mmWave       | millimeter-wave |
| NLoS         | non-line-of-sight |
| NOMA         | non-orthogonal multiple access |
| OFDM         | orthogonal frequency-division multiplexing |
| OTFS         | orthogonal time-frequency modulation |
| PAPR         | peak-to-average power ratio |
| PN           | phase noise |
| QAM          | quadrature amplitude modulation |
| RF           | radio-frequency |
| RMS          | root mean square |
| SA           | subarray |
| SC           | single-carrier |
| SM           | spatial modulation |
| SNR          | signal-to-noise ratio |
| SPP          | surface plasmon polariton |
| Tbps         | terabit-per-second |
| THz          | terahertz |
| THz-TDS      | terahertz time-domain spectroscopy |
| UE           | user equipment |
| UM-MIMO      | ultra-massive MIMO |
| VLC          | visible light communications |
| WDM          | wavelength division multiplexing |

upper case letters correspond to scalars, vectors, and matrices, respectively. We denote by $(\cdot)^T$, $(\cdot)^H$, and $E[\cdot]$, the transpose, conjugate transpose, and expected value, respectively. Table II summarizes the abbreviations.

II. SYSTEM MODEL

It is challenging to define a generic system model for THz communications at this early stage. Nevertheless, the use of AoSAs of AEs is most likely to be the norm in future THz systems, as dynamic array gains are crucial for combating the distance problem [92]. A typical THz communications system model is illustrated in Fig. [I] where adaptive AoSAs are configured at the transmitting and receiving sides. After the digital-to-analog converter (DAC) and before the ADC, each SA is fed with a dedicated RF chain. Due to high directivity, each SA is detached from its neighboring SAs in a multi-user setting. The role of baseband precoding reduces to defining the utilization of SAs or simply turning SAs on and off. In a point-to-point setup, however, SA paths can be highly correlated due to low spatial resolution.

We adopt the three-dimensional (3D) UM-MIMO model of [97]–[99]. The AoSAs consist of $M_r \times N_r$ and $M_t \times N_t$ SAs, at the transmitter and the receiver, respectively. Each SA is composed of $Q \times Q$ AEs. Therefore, the overall configuration can be represented as a “large” $M_r N_r Q^2 \times M_t N_t Q^2$ MIMO system [100]. Such large and near-symmetric doubly-massive MIMO systems [101] differ from conventional massive MIMO systems. In the latter, large antenna arrays are typically configured at a transmitting base station to serve multiple single-antenna users at the receiver. The distances separating two SAs or two AEs are critical design parameters in reconfigurable settings, as discussed in subsequent sections. We denote these distances by $\Delta$ and $\delta$, respectively.

THz signal propagation is highly directional (quasi-optical) for three main reasons. First, limited THz reflected components and negligible scattered and diffracted components result in channels dominated by the line-of-sight (LoS) path and assisted by possibly very few NLoS reflected multipath components. Second, high-gain directional antennas are typically used to combat the distance problem instead of omnidirectional antennas with 0 dBi gains, which further reduces the surviving paths to a single path. Third, the high array gains of beamforming guarantee directional “pencil beams,” where typically each SA generates a single beam. Consequently, we assume for the generic system model an LoS transmission over an SC frequency-flat fading channel. The corresponding baseband system model is

$$y = W^H \cdot HW_t^H \cdot x + W^H \cdot n,$$  \hspace{1cm} (1)

where $x = [x_1, x_2, \cdots, x_N]_t^T \in \mathbb{C}^{N_x \times 1}$ is the information-bearing symbol vector of components belonging to a QAM constellation $X$; for example, $y \in \mathbb{C}^{N_t \times 1}$ is the received symbol vector, $H = [h_1, h_2, \cdots, h_M, N_t] \in \mathbb{C}^{M_r \times N_t}$ is the channel matrix, $W_t \in \mathbb{C}^{N_r \times M_t}$ and $W_r \in \mathbb{C}^{N_t \times N_t}$ are the baseband precoder and combiner matrices, and $n \in \mathbb{C}^{M_r \times N_t}$ is the additive white Gaussian noise (AWGN) vector of power $\sigma^2$.

An element of $H$, $h_{m_n m_n n_1}$, the frequency response between the $(m_r, n_r)$ and $(m_t, n_t)$ SAs, is thus defined as

$$h_{m, n, m, n} = a^H(\phi, \theta) G \gamma_{a_{m, n, m, n}} G_r a(\phi, \theta),$$  \hspace{1cm} (2)

for $m_r = 1, \cdots, M_r$, $n_r = 1, \cdots, N_r$, $m_t = 1, \cdots, M_t$, and $n_t = 1, \cdots, N_t$, where $\alpha$ is the path gain, $a_r$ and $a_t$ are the transmit and receive SA steering vectors, $G_r$ and $G_t$ are the transmit and receive antenna gains, and $\phi, \theta_1$ and $\phi, \theta_2$ are the transmit and receive angles of departure (AoD) and arrival (AoA), respectively (\phi$’$s are the azimuth angles and \theta$’$s the elevation angles).

The steering vectors can be expressed as a function of
the transmit and receive mutual coupling matrices, $C_x, C_r \in \mathbb{R}^{Q_x \times Q_r}$ as $a_r(\phi_r, \theta_r) = C_r a_0(\phi_r, \theta_r)$ and $a_t(\phi_t, \theta_t) = C_t a_0(\phi_t, \theta_t)$. By setting $C_t = C_r = I_Q^2$ ($I_N$ is an identity matrix of size $N$), the effect of mutual coupling is neglected. Such an assumption is especially valid in the plasmonic case when $\delta \geq \lambda_{spp}$ [102], where the SPP wavelength, $\lambda_{spp}$, is much smaller than the free-space wavelength, $\lambda$. The ideal SA steering vector at the transmitter side can thus be expressed as

$$a_0(\phi_t, \theta_t) = \frac{1}{Q} \left[ e^{j\phi_{1,1}}, \ldots, e^{j\phi_{1,0}}, \ldots, e^{j\phi_{Q_2,1}}, \ldots, e^{j\phi_{Q_2,0}}, \ldots, e^{j\phi_{Q_1,0}} \right]^T,$$

where $\Phi_{p,q}$ is the phase shift that corresponds to AE $(p, q)$, and is defined as

$$\Phi_{p,q}(\phi_t, \theta_t) = \psi_x^{(p,q)} \frac{2\pi}{\lambda} \cos \phi_t \sin \theta_t + \psi_y^{(p,q)} \frac{2\pi}{\lambda} \sin \phi_t \sin \theta_t + \psi_z^{(p,q)} \frac{2\pi}{\lambda} \cos \theta_t,$$

with $\psi_x^{(p,q)}$, $\psi_y^{(p,q)}$, and $\psi_z^{(p,q)}$ being the coordinate positions of AEs in the 3D space. Furthermore, for beamforming per SA, given the target AoDs, $\hat{\phi}_t, \hat{\theta}_t$, the beamforming vector $\hat{a}_t(\hat{\phi}_t, \hat{\theta}_t)$ can be similarly defined using

$$\Phi_{p,q}(\hat{\phi}_t, \hat{\theta}_t) = \psi_x^{(p,q)} \frac{2\pi}{\lambda} \cos \hat{\phi}_t \sin \hat{\theta}_t + \psi_y^{(p,q)} \frac{2\pi}{\lambda} \sin \hat{\phi}_t \sin \hat{\theta}_t + \psi_z^{(p,q)} \frac{2\pi}{\lambda} \cos \hat{\theta}_t.$$

The equivalent array response could then be represented as

$$a_{eq}^{(t)} = \frac{1}{\sqrt{M_t N_t}} \sum_{m_t=1}^{M_t} \sum_{n_t=1}^{N_t} e^{j \frac{2\pi}{\lambda} (\Phi_{p,q}(\phi_t, \theta_t) - \Phi_{p,q}(\hat{\phi}_t, \hat{\theta}_t))}.$$

At the receiver side, $a_r(\phi_r, \theta_r)$, beamforming vector $\hat{a}(\hat{\phi}_r, \hat{\theta}_r)$ ($\hat{\phi}_r, \hat{\theta}_r$ are the target AoA), and $a_{eq}^{(r)}$ can be similarly defined.

### III. THz-Band Channel Modeling

Channel modeling is essential for efficient signal processing in the THz band. Accurate THz channel models should consider the effect of both the spreading loss and the molecular absorption loss and should account for the LoS, NLoS, reflected, scattered, and diffracted signals. Channel modeling approaches are primarily deterministic or statistical [103]. Although deterministic channel modeling uses computationally extensive ray-tracing techniques to capture site geometry, matrix-based statistical modeling represents each independent sub-channel using a random variable of a specific distribution. Hybrid channel modeling schemes combine the advantages of both approaches, where dominant paths are captured deterministically and other paths are generated statistically [104].

#### A. Ray-Based THz Channel Modeling

Several extensive ray-tracing-based THz propagation measurements have been recently reported. For instance, a unified multi-ray THz-band channel model is proposed in [105], which covers the LoS, scattered, reflected, and diffracted paths and is experimentally validated over 0.06 – 1 THz. Ray-tracing techniques are similarly used to post-process sub-THz channel measurements in [106]. In [107], a deterministic channel model over 0.1 – 1 THz is proposed for LoS and NLoS scenarios, using the Kirchhoff scattering theory and ray tracing. Similarly, ray-based sub-THz channel characterization at 90 – 200 GHz is detailed in [108] using deterministic simulations in indoor office and outdoor in-street scenarios.

Other THz ray-tracing channel modeling attempts are tailored for the peculiarities of specific use cases. For instance, a ray-tracing channel model at 300 GHz is presented in [28], for close-proximity THz communications, such as for Kiosk downloading. Also, at 300 GHz, a ray-tracing simulator with calibrated electromagnetic parameters is used in [109] for vehicle-to-infrastructure THz communications. For reducing the complexity of ray tracing in UM-MIMO systems, select few virtual paths can be captured between virtual transmitting and receiving points, the response of which gets mapped to actual pairs of transmitting and receiving AEs.

#### B. Statistical THz Channel Modeling

Several statistical THz channel modeling alternatives to time-consuming and complex ray-tracing models in fixed geometries have been attempted. For example, by developing a wideband channel sounder system at 140 GHz, indoor wideband propagation and penetration measurements for common building materials are reported in [110]. In [111], indoor measurements and models for reflection, scattering, transmission, and large-scale path loss are provided by the same group for mmWave and sub-THz frequencies; a 3D statistical indoor channel model is further reported in [112]. A lower reflection loss is noted at higher frequencies in indoor drywall scenarios (stronger reflections). In contrast, the partition loss increases due to more prominent depolarizing effects. Furthermore, recent measurement-based models are reported by the same group for sub-THz urban microcell [113], indoor office [114], and [115] space scenarios. Natural isolation is noted between terrestrial networks and surrogate satellite systems and between terrestrial mobile users and co-channel fixed backhaul links.

The statistical characterization of three channel bands between 300 THz and 400 THz is presented in [116], based on a broad set of measurements in LoS and NLoS environments and including spatial and temporal variations. The large-scale losses are modeled using the single slope path loss model with shadowing, where variations due to shadowing are normally distributed. Metal, wood, and acoustic ceiling panels are confirmed as good reflectors. Strong multipath components reduce the coherence bandwidth in indoor environments significantly, and high channel correlation is retained. With a virtual antenna array technique, the same testbed is exploited in [117] to demonstrate 2x2 THz LoS MIMO channels. In [118], channel measurements agree closely with the new radios (NR) channel model of the Third Generation Partnership Project (3GPP) for a specific indoor scenario.

Another stochastic indoor 300 GHz spatio-temporal channel model is introduced in [119] that considers parameters such as polarization, ray amplitudes, times of arrivals, angles of arrivals and departures, and path-specific frequency dispersion. Furthermore, THz channel modeling via a mixture of gamma
distributions is proposed in [120]. In other notable studies, LoS broadband THz channel measurements are reported in [121] when using convex lenses at the transmitter, coupled with collimating lenses at the receiver. A beam domain channel model is also introduced in [122]. With large numbers of base stations and users, and given that several wavelengths typically separate users at high frequencies, the beam-domain channel elements are statistically uncorrelated [123], and their envelopes are independent of frequency and time. Moreover, a geometric-based stochastic time-varying model at 110 GHz is proposed in [124] for THz vehicle-to-infrastructure communications.

A comprehensive statistical simulator of 3D end-to-end wideband UM-MIMO THz channels has recently been introduced in [99]. The simulator, called TeraMIMO, models THz channel statistics such as coherence time, coherence bandwidth, Doppler spread, and root mean square (RMS) delay spread. TeraMIMO generates frequency-selective and time-varying THz channels for different communication distances, ranging from nanocommunications to short-range indoor and outdoor communications and LoS links of hundreds of meters. TeraMIMO also accounts for LoS-dominant and NLoS-assisted scenarios. By adopting the same AoSA architecture of our system model, TeraMIMO studies the resultant spatial effects of misalignment, spherical wave propagation, and beam split in wideband THz channels. Such tools can catalyze THz research, especially if they continue to follow the latest results from measurement-based THz channel modeling campaigns.

C. Effect of Scattering

The electromagnetic roughness of surfaces increases at higher frequencies, causing diffuse scattering and increased backscattering (at lower incident angles) [125]. The effect of scattering on the reflection coefficient in THz communications is studied in [126]. The scattered power increases with frequency and surface roughness relative to the reflected power, where smooth surfaces (like drywall) can be modeled as reflective surfaces. In [127], diffuse scattering in THz massive MIMO channels is studied by developing a hybrid modeling approach for 3D ray-tracing simulations by assuming realistic indoor environments over the 0.3–0.35 THz band. The channel capacity of indoor massive MIMO channels is calculated by assuming different surface roughnesses for LoS and NLoS scenarios. Scattering can be leveraged to achieve a trade-off between rich multipath and high received power in THz massive MIMO. Scattering can thus enhance the spatial multiplexing gains. Diffuse scattering can also be leveraged to identify surface types.

In [128], [129], it is argued that THz beams are more susceptible to snow than rain, suffering higher losses under an identical fall rate Mie theory approach for electromagnetic radiation. THz-band rain-induced co-channel interference is studied in [130], using the bistatic radar and the Mie scattering theory, and assuming first-order multiple scattering. The overall interference levels from rain are significantly lower in the THz band (20 dB difference between 300 GHz and 60 GHz), except when the receiver is very close to the LoS (forward-oriented scattering at high frequencies). Furthermore, in [115], rain attenuation measurements at 140 GHz reveal that communications above 70 GHz are not further impacted by rain.

D. Effect of Blockage

Statistical modeling can also be used to study the effect of blockage, which is significant at higher frequencies. For instance, NLoS THz channel modeling is conducted in a generic stochastic approach in [131] by assuming rectangular geometry and accounting for variable densities of reflecting objects (single reflection components) and blocking obstacles. THz signals are more sensitive to blockages than mmWave signals. Human blockage in indoor THz communications is studied in [132], where adding extra antennas to account for blocked streams is considered. The average network throughput and coverage probability in the presence of indoor blockage effects resulting from human bodies and walls are further studied in [133].

The dynamic blockages caused by moving humans are incorporated into the 3D THz channel model in [134], where deploying antennas at high altitudes would help prevent blockages. Furthermore, the effect and mitigation of blockage in THz relay systems are addressed in [135], [136]. Moreover, blockages can arise at the transmitter due to the condensation of particles, and such blockages can be mitigated using compressive sensing techniques [137], [138]. A similar remote array diagnosis technique that mitigates antenna blockage without the need for full channel state information (CSI) is reported in [139].

E. The Molecular Absorption Effect

The path loss of a THz signal in the presence of water vapor is dominated by spikes that represent molecular absorption losses originating at specific resonant frequencies due to excited molecule vibrations. Higher densities of absorbing molecules strengthen and widen the peaks (broadening of absorption lines). The spectrum is divided into smaller windows (sub-bands) of tens or hundreds of GHz because of these lines. The windows are distance-dependent because some spikes only become significant at specific distances (by increasing the distance from 1 to 10 meters, the transmission windows are reduced by order of magnitude [140]). Hence, variations in the communication distance affect both the available bandwidths and the path loss, where the available bandwidth shrinks at higher frequencies.

The LoS path gain as a function of absorption is expressed as

$$d_{\text{LoS}}^{m,n_r,m_n} = \frac{c_0}{4\pi f d_{m_n,n_r,m_n}} e^{-\frac{1}{2}K(f)d_{m_n,n_r,m_n} e^{-\frac{2\pi f d_{m_n,n_r,m_n}}{c_0}}}$$

where $d_{m_n,n_r,m_n}$ is the distance between the transmitting and receiving SAs, $K(f)$ is the absorption coefficient, $f$ is the frequency of operation, and $c_0$ is the speed of light in vacuum. $K(f)$ is derived in [65] as a summation over contributions from isotopes $i (i = 1, \ldots, f)$ of gases $(g = 1, \ldots, G)$ that
\[ K(f) = \sum_{i,g} \frac{p}{p_0} T \frac{p}{RT} q^{i,g} N_A S^{i,g} \frac{f}{f_{c0} + \delta^{i,g} \frac{p}{p_0}} \tanh \left( \frac{h c f}{2 K_B T} \right) \left( 1 - q^{i,g} \right)^{\alpha^{i,g}_0 + q^{i,g} \alpha^{i,g}_0} \left( \frac{p}{p_0} \right) (T_0/T)^{\gamma} \frac{f}{f_{c0} + \delta^{i,g} \frac{p}{p_0}} \right) \]

\[ K(f) = K_1(f, v) + K_2(f, v) + K_3(f), \]

where

\[ K_1(f, v) = \frac{A(v)}{B(v) + \left( \frac{f}{1000} - c_1 \right)^2}, \]

\[ K_2(f, v) = \frac{C(v)}{D(v) + \left( \frac{f}{1000} - c_2 \right)^2}, \]

\[ K_3(f) = \rho_1 f^3 + \rho_2 f^2 + \rho_3 f + \rho_4, \]

with \( v \) being the volume mixing ratio of water vapor. The rest of the coefficients and functions are detailed in [142]. The updated model in [143] is valid over an extended sub-THz range (100–450 GHz), accounting for more absorption spikes. Both approximations are sufficiently accurate for links up to 1 kilometer (Km) under standard atmospheric conditions. Although such approximations reduce the computational complexity and are easy to track analytically, the exact HITRAN-based absorption model is still favored, especially in very high SNR settings. In the context of joint signal processing for communications and sensing, where exact knowledge of medium components is sought (Sec. X-B), precise HITRAN-based molecular absorption modeling is crucial.

The distance-dependent path loss is illustrated in Fig. 2.
response within a time margin $T_s$ is expressed as 
\[ h_{\mathrm{NL},m} = \sum_{n=0}^{N_{\mathrm{cl},m}-1} \sum_{v=0}^{N_{\mathrm{cl},m}^\mathrm{sv}} a_{\mathrm{NL}} \left( \phi_{m,v}^{(u,v)}, \theta_t^{(u,v)} \right) G_t \left( \phi_t^{(u,v)}, \theta_r^{(u,v)} \right), \]
where $N_{\mathrm{cl},m}$ is the number of path clusters and $N_{\mathrm{cl},m}^\mathrm{sv}$ is the number of paths within the $v$th cluster. Each path can have random angles of departure and arrival within a beam region. The expectation of path gain magnitude is 
\[ E[|h_{\mathrm{NL},m,n}(f)|^2] = \left( \frac{c_0}{4\pi f d_{\mathrm{m},n}} \right)^2 e^{-\kappa(f)d_{\mathrm{m},n}} e^{-\tau_v \frac{f}{c}} e^{-\tau_r \frac{f}{c}} \]
where $\tau_v$ and $\tau_r$ are the times of arrival that follow an exponential or paraboloid distribution and $\Gamma$ and $\gamma$ are the clusters and ray decay factors, respectively. The angles of departure and arrival are calculated as 
\[ \phi_t^{(u,v)} = \phi_t^{(u)}, \phi_r^{(u,v)} = \phi_r^{(u)} + \phi_t^{(u)}, \]
\[ \theta_t^{(u,v)} = \theta_t^{(u)}, \theta_r^{(u,v)} = \theta_r^{(u)} + \theta_t^{(u)}, \]
where $\phi_t^{(u)}$ and $\theta_t^{(u)}$ are the uniformly distributed cluster azimuth and elevation angles of departure/arrival, and $\phi_r^{(u)}/\phi_t^{(u)}$ and $\theta_r^{(u)}/\theta_t^{(u)}$ are the ray azimuth and elevation angles of departure/arrival. The latter can follow a zero-mean second-order Gaussian mixture model.

G. Wideband Channels and Beam Split

Frequency selectivity due to multipath or molecular absorption is more significant in wideband THz channels supported in several scenarios. Ultra-wideband pulse-based modulations are serious candidates for THz communications in which short pulses in time span the entire THz range in frequency [151, 152]. In such wideband scenarios, a deterministic frequency-selective fading, not captured by Rayleigh and Ricean models, arises due to molecular absorption. This fading results in delayed signal components in a wideband multipath scenario. Hence, group velocity dispersion (GVD) is another issue that becomes limiting at specific link distances, atmospheric water vapor densities, and channel bandwidths [153]. This phenomenon also results in inter-symbol interference (ISI) as data bits spread out of their assigned slots and interfere with neighboring slots. Therefore, dilating bit slots can solve this problem at the expense of the data rate. In [154], the atmospheric GVD of THz pulses (0.2 – 0.3 THz) is compensated for using stratified media reflectors.

Beam split is another critical phenomenon that arises in ultra-wideband THz signals with large fractional bandwidths (ratio between bandwidth and central frequency). The beam splits when different THz path components at different subcarriers squint into different spatial directions, resulting in array gain loss [155]. Beam squint can occur due to frequency-independent delays in analog beamforming phase shifters when the same phase shift is applied to different frequencies. The difference between $f_k$, the frequency of a subcarrier in a multi-carrier system, and $f_c$, the center frequency, is significant at THz frequencies. Therefore, the paths split into different spatial directions within an ultra-broadband bandwidth. The angle domain beam split is expressed as 
\[ \Phi_p,q(f_k, \theta_\tau, f_c) = \frac{f_k}{f_c} \Phi_p,q(f_c, \theta_\tau, f_c). \]

The large UM-MIMO AoSAs also add a beam split effect when the signal propagation time between SAs introduces a frequency-dependent phase shift that biases the estimated angle of arrival in steering vectors [155]. The highly narrow beamwidths under UM-MIMO beamforming worsen the beam split effect. The receiving antenna array might often be larger than the received beamwidth, which indicates that a path is not visible to all the antennas in the array. This phenomenon causes non-stationarity [156] and variations in time of arrival, angle of arrival (and angle of departure), and received amplitudes across the antenna array. The variation in the time of arrival causes ISI (not only a phase shift) in wideband scenarios [157].

Beam split can be mitigated in the digital domain of hybrid architectures [158, 159]. In [160], a beam squint mitigation scheme based on introducing true-time-delays over uniform planar arrays is investigated for channel estimation and hybrid combining. In [161], the use of radix-2 self-recursive sparse factorizations of delay Vandermonde matrices as alternatives to fast Fourier transform is robust to beam split effects. Furthermore, wideband beam zooming schemes based on beam tracking are proposed to combat beam split, such as the use of delay-phase precoding [162, 162, 163].

H. Spherical Wave Model

The near- and far-field considerations are also critical at THz frequencies. The plane wave propagation model applies to far-field scenarios where the distance between the transmitter and the receiver is greater than or equal to the Rayleigh distance of the antenna array [164], or when the distance is comparable to the array dimensions [165]. At lower microwave and mmWave frequencies, this distance is less than 0.5 m and for an array size of 0.1 m and an operating frequency of 6 GHz and 60 GHz, respectively. However, this distance grows to approximately 40 m at 0.6 THz, which is greater than most achievable THz communication distances, confirming the importance of the spherical wave propagation models. The spherical wave model should be considered when the distance is within the Fresnel region [166]:
\[ 0.62 \sqrt{\frac{D^3}{\lambda}} \leq d < 2D^2 \frac{D^2}{\lambda}, \]
where $D$ is the overall maximum antenna dimension (a slightly different MIMO near-field definition is adopted in [165]). Spherical waves can be modeled at AE or SA levels. In compact device technologies, such as plasmonics, an SA-level spherical wave model can be sufficient for most communication distances. Retaining the plane wave assumption at the AE level has the advantage of generating a simple and convenient equivalent array response. The spherical signal wave considerations are at the distance and angle levels—the distance between two SAs in the far-field should be calculated by accounting for the curvature, and the AoD/AoA differ for different SAs.
I. Antenna Gains

Directional antennas are essential for overcoming the high THz propagation losses. A simplified ideal sector model for antenna gains, \( G_t \) at the transmitter and \( G_r \) at the receiver, can be expressed as [167]:

\[
G = \begin{cases} \sqrt{\frac{\theta_{\text{max}}}{\theta_{\text{min}}}}, & \forall \phi \in \left[\phi_{\text{min}}, \phi_{\text{max}}\right], \forall \theta \in \left[\theta_{\text{min}}, \theta_{\text{max}}\right], \\
0, & \text{otherwise.}
\end{cases}
\]

This model can be applied for both LoS and NLoS components with the corresponding azimuth and elevation angles. For highly directional antennas, \( G_0 \) can be approximated as [166], [168]

\[
G_0 = \frac{4\pi}{\psi_A} \approx \frac{4\pi}{\psi_\phi \psi_\theta},
\]

where \( \psi_A \) is the beam solid angle, and \( \psi_\phi \) and \( \psi_\theta \) are the half-power beamwidths (HPBWs) in the azimuth and elevation planes, respectively. The antenna sectors are small under antenna directivity. For example, HPBW azimuth/elevation-plane angles \( \psi_\phi = \psi_\theta = 27.7^\circ \) only result in a 17.3 dBi gain [168]. However, much higher gains are required in medium-distance THz communications. Array gains can complement antenna gains. For instance, a massive number of THz-operating antennas can be fit into a few square millimeters.

J. Effect of Misalignment and Impairments

The performance of THz communications systems severely deteriorates under the effect of misalignment and hardware impairments. Misalignment occurs when the transmitter and receiver do not precisely point to each other, a highly probable scenario with narrow THz beams [169]. The joint effect of misalignment and hardware impairments on THz communications is studied in [170], including in-phase and quadrature imbalance (IQI) and non-linearities. The study models all impairments as Gaussian noise components and accounts for operation and design parameters alongside environmental parameters; it also introduces a misalignment fading model. However, the adopted model is based on the optical receiver’s intensity fluctuation derived in [171].

Alternative THz- and UM-MIMO-specific misalignment models need to be derived, by complementing the optical model with approximations of the effective radius of the receiving AoSA area, such as in [172]. The research in [170] is extended in [172] to capture the error analysis of mixed THz-RF wireless systems. The use of high-directivity antennas in THz systems results in small transceiver antenna beamwidths, which provide higher antenna gains but cause pointing errors and loss of connection. Due to the symmetry of the beam, the misalignment fading component depends only on the radial distance [173]. Earlier attempts to capture this misalignment effect are reported in [174]–[176]. The effect of small-scale mobility on THz systems is studied in [177], [178]. Simple shakes or rotations due to user equipment (UE) mobility can result in beam misalignment and SNR degradation, which result in a loss in communication time due to extra beam-search mechanisms. Therefore, there exists a trade-off between antenna directivity and capacity.

Misalignment can be modeled by expressing the effective channel coefficient between two SAs in terms of three components [170] as \( h_{\text{eff}} = hh_{\text{max}}h_d \), where \( h \) is expressed in \( \{2\} \), \( h_{\text{max}} \) is the misalignment fading, and \( h_d \) is the stochastic path gain (can be neglected or modeled as an \( \alpha - \mu \) process depending on the scenario). Due to beam symmetry, misalignment fading depends primarily on the pointing error, which can be expressed in the form of a radial distance \( r \) between the transmission and reception beams at a communication distance \( d \):

\[
h_{\text{max}}(r,d) \approx A_0 \exp\left(-\frac{2r^2}{w_{eq}^2}\right),
\]

where \( A_0 \) is the fraction of power collected at the receiver, and \( w_{eq} \) is the equivalent beamwidth. Hardware imperfections in both the transmitter and the receiver can be modeled as two additional distortion noises. The modified system model under impairments is approximated as

\[
y = H(x + n_t) + n_f + n,
\]

where \( n_t \in \mathbb{C}^{M_t \times 1} \) and \( n_r \in \mathbb{C}^{M_r \times 1} \) are two complex Gaussian distortion noise vectors at the transmitter and the receiver, with noise variances \( \eta_t \hat{\rho} \) and \( \eta_f \hat{\rho} \hat{h}^2 \), respectively, with \( \hat{\rho} \) being the average transmitted power and \( \eta_t \) and \( \eta_f \) being the impairment coefficients [170].

K. THz Noise Modeling

Accurate noise models are essential for understanding the behavior of THz systems. Although stochastic models for the electronic noise at THz receivers are still lacking, in [179], two primary sources of noise are noted: (1) thermal noise, which arises at the receiver multiplier and mixer chains, and (2) absorption noise that is channel-induced due to water vapor molecules. The corresponding histogram of the measured noise follows a Gaussian distribution, in accordance with the noise behavior at lower frequencies, as opposed to shot noise in optical receivers. In [180], the transmission-induced noise due to molecular absorption is discussed in more detail. The researchers differentiate between multiple models, most of which are based on the antenna temperature generated by the absorbed energy. However, this molecular absorption noise model has not yet been validated by measurements; it uses sky noise as a basis, which might overestimate the level of the self-induced noise. In [181], [182], molecular absorption is assumed to result in a rich scattering environment because absorptions are followed by re-radiations with minimal frequency shifts. Nevertheless, such coherent re-radiations can be more realistically lumped in a generic absorption noise factor [65], [180].

The channel-induced component dominates the overall noise in pulse-based systems (especially in low-noise graphene-based electronic devices [183]), and is colored over frequency. The total noise power at a distance \( d \) can be expressed as

\[
\sigma^2 = K_B \int_{B} T_{\text{noise}}(f,d)df,
\]

where \( T_{\text{noise}} = T_{\text{sys}} + T_{\text{mol}} + T_{\text{other}} \). \( T_{\text{sys}} \) and \( T_{\text{mol}} \) are the system electronic noise temperature and molecular absorption noise, respectively, and

\[
T_{\text{mol}}(f,d) = T_0(1 - e^{-\kappa f d}).
\]
In a carrier-based system with perfect frequency planning over absorption-free spectra, the effect of the channel-induced noise can be minimized. This is particularly true at shorter communication distances, where propagation losses dominate. At larger distances and higher frequencies, however, molecular absorption losses can overcome propagation losses. An additional low-frequency noise component exists at the transmission chain and power supply.

In addition to these noise sources, the effect of phase noise (PN) at THz frequencies should be considered. PN is caused by time-domain instability (jitter), which produces random rapid, short-term fluctuations in the phase. Precise THz-specific PN measurements are still lacking, despite some modeling attempts. PN can be modeled by the superposition of Wiener and Gaussian noises. Furthermore, PN is typically accompanied by strong phase impairment and carrier frequency offset (CFO)—both result from poorly performing high-frequency oscillators (more of a problem in multi-carrier systems). Novel signal processing techniques and optimized modulation/demodulation schemes can overcome these impairments and achieve PN robustness. THz CFO can be estimated via an under-sampling approach with narrow-band filtering and coprime sampling. Other notable studies that account for PN and impairments at the sub-THz band include on MIMO techniques, on SC transceivers, on OFDM systems, and on mesh backhaul capabilities.

IV. Analysis Frameworks and Testbeds

The more accurate the THz channel models get, the more insight we have on the achievable gains of THz systems. These gains can be captured via theoretical performance analysis frameworks and can be verified via experimental testbeds. In this section, we summarize recent results on both.

A. Fundamental Performance Analysis Frameworks

Early THz channel capacity studies are reported in in the context of nanonetworks, where numerical results are generated for different molecular compositions and power allocation schemes, with emphasis on pulse-based modulation. In a thorough analysis of THz channel characteristics is presented, where the variability of spectral window widths over communication distances is first observed. Water-filling power allocation can achieve more than 75 Gbps with 10 dBm transmit power over 0.06 – 1 THz. The study further illustrates that in multipath scenarios, the RMS delay spread is distance- and frequency-dependent, and the coherence bandwidth is less than 5 GHz (decreases with longer distance and lower carrier frequencies). Furthermore, the spacings between transmissions are affected by the increased temporal broadening effects at higher frequencies, wider pulse bandwidths, and longer distances. Therefore, distance-adaptive techniques are suggested alongside multi-carrier schemes.

Theoretical performance analysis of conventional THz communications systems is also progressing. For instance, in the two-path channel characteristics over 275 – 400 GHz are analyzed in terms of SNR and ergodic capacity by considering channel characteristics such as frequency selectivity, path-loss, and atmospheric conditions. The analysis also accounts for transceiver parameters such as antenna gains and transmit power. By assuming the signal and noise to be jointly Gaussian, classic Shannon results for coherent reception are used for capacity estimation. This approach is employed in several other studies: in to compute the capacity of THz nanosensor networks, in to capture the relation between transmission distance and absorption-free (transparency) windows, and in to study the data rates of fixed THz-links. Such studies are also extended in the context of reconfigurable MIMO systems. Furthermore, the effect of both deterministic (molecular absorption) and random (atmospheric turbulence and pointing errors) factors on the bit-error-rate performance and capacity of LoS THz links is studied in, where the log-normal, gamma-gamma, and exponentiated Weibull channel models are used.

THz systems are also susceptible to transceiver imperfections. In, it is argued that in the presence of PN and misalignment, the outage performance is not significantly enhanced by higher transmit power, and lower-order modulations might be required. The effect of local oscillator hardware impairments can also be more severe than misalignment issues. The performance of THz systems is studied under the joint impact of PN and misalignment in, and under the joint impact of PN and amplifier non-linearities in. When lumped into PN at the local oscillator, the impact of these errors examined under different transceiver architectures in.

The achievable data rates of indoor THz systems are studied in, where a single frequency network is advocated, and the corresponding ISI due to channel dispersion considered, and the effect of the density of access points on performance is studied. Similarly, in, the indoor interference and coverage under beamforming are studied. An analytical model for the distribution of indoor access points at different blocks of the THz spectrum is proposed in. In THz indoor NLoS scenarios, the total received signal is accumulated predominantly from diffuse reflections. Small-scale mobility in indoor THz scenarios is also studied as a function of several variables, such as frequency windows, beamwidths, distance, humidity, mobility type, and antenna placement—there exist optimal beamwidths for specific mobility types and AP placement strategies. In, the interference in mmWave and THz systems under blockage and directional antennas is studied.

B. Performance Analysis for Novel Use-Cases

It is crucial to study the system-level resource allocation considerations of THz communications. The purpose of such studies is to decide on the optimal use of THz signals in a future communications system, whether in the backhaul of networks to alleviate data rate bottlenecks or in the access domain to deliver novel communication and sensing applications. The performance under different medium conditions (snow, rain) should be studied for backhaul links, comparing THz links to other backhaul solutions. In the access domain, a stochastic
geometric approach [207] can be used, as an alternative to time-consuming and error-prone heavy simulations, to capture the outage probability and optimize resource allocation in dense deployments under interference and blockages. Despite THz signal directionality, inter-cell interference can still arise in dense deployments of THz base stations that overcome the high path and molecular absorption losses. Stochastic geometry can introduce a mathematically compliant formulation for inter-cell interference, which is extremely difficult under other approaches. Stochastic geometry at the THz band should account for a practical blockage model, highlighting the impact of the penetration loss on the communication quality of the THz channel in indoor and outdoor environments. The impact of the high molecular absorption loss, spreading loss, and scattering, the use of directional antennas and beamforming techniques, and the impact of misalignment can also be modeled by stochastic geometry.

A stochastic geometry approach for mean interference power and outage probability analysis is considered in [207], in the context of a dense THz network operating over 0.1 – 10THz. The researchers model the interference as a shot noise process, and they assume directional antennas. Furthermore, although high antenna gains result in a lower probability of interference in the THz band, the interference level is much higher when it occurs. Stochastic geometry is also used in [208] to derive the exact and approximate distributions of the received signal power and interference, respectively; semi-closed-form expressions are derived for the coverage probability and the average achievable rate. Eventually, heterogeneous networks (HetNets) comprising mmWave, THz, and optical wireless communications should be analyzed [209]. Integrating small cells operating at mmWave and THz frequencies is required to meet the ever-increasing demand for ultra-high data rates; retaining sub-6 GHz cells overcomes the limited coverage of THz communications. Densifying the network with THz base stations should increase the average rate significantly but not the coverage probability that decreases when THz small cells dominate the network.

The performance of THz systems is also studied in use-case-specific scenarios. For example, the reliability and latency of THz communications are studied in the context of wireless virtual reality in [210], [211]. High reliability can be achieved with proper densification, which is illustrated by deriving a tractable expression for system reliability as a function of THz system parameters. In another example [212], the channel capacity and reliability are studied for the special case of THz wireless networks-on-chip communications [213], [214], where performance can be enhanced by proper choice of silicon layers and their thickness. Furthermore, user- and network-centric metrics for THz information shower systems are evaluated in [215], where 95% of traffic from long-range networks can be offloaded, initiating heavy-traffic THz information shower sessions.

THz signals are also being considered for communications at atmospheric altitudes (where the concentration of water vapor decreases) among drones [216], [216], jets, unmanned aerial vehicles [217], high-altitude platform systems [218], and satellites [219], [220]. Calculating the absorption loss through the atmosphere at higher altitudes permits such applications. In [221], communication at 0.75 – 10THz is demonstrated as more feasible at higher altitudes than sea-level, with reported usable bandwidths of 8.218 THz, 9.142 THz, and 9.25 THz over a distance of 2 km. In [222], the capacity of optoelectronic THz Earth-satellite links is analyzed, where the claim is made that 10 Gbps per GHz can be supported.

Similarly, the use of the THz band for simultaneously providing high data rates and wide coverage data streaming services to ground users from a set of hotspots mounted on flying drones is studied in [223]. By solving an optimization problem for resource utilization, much higher throughput can be achieved in mobile environments than in static environments. Furthermore, a holistic investigation on THz-assisted vertical HetNets is conducted in [218]. The study comprises, in addition to terrestrial communication links, geostationary and low-earth orbit satellites, networked flying platforms, and in-vivo nanonetworks; accurate channel modeling is critical for harmony across all these applications. Analyses on THz-band nanocommunications are also reported in [224] and in particular for body-centric applications in [225] – [227].

C. Experimental Demonstrations

Several experimental demonstrations have been conducted to verify the corresponding channel models and predicted performance metrics. Experimental results for the first true-THz absorption-defined window above 1 THz (1.02 THz) are reported in [179], where tens of Gbps are demonstrated in a multi-carrier (OFDM) system over sub-meter distances. A typical software-defined PHY layer transceiver system consists of frame generation, modulation, pulse shaping, pre-equalization, noise filtering, frame synchronization, post-equalization, and demodulation. Although pre-equalization accounts for the frequency-selective response of components, post-equalization mitigates ISI and the frequency-selective channel response. A correlator filter is used for frame synchronization [179].

In [228], the researchers demonstrated a THz LoS link with digital beamforming, at an operating frequency of 300 GHz, a channel bandwidth of 1.5 GHz, and a communication distance reaching 0.6 m; the results confirm an 11 Gbit/sec data rate. Tbps speeds are demonstrated over THz LoS links in [229], where NLoS links through first-order reflections are also feasible in cases of signal obstruction. These observations inspire the development of MIMO mechanisms that exploit spatial diversity in transmission, reception, and reflection. Furthermore, in [230], a live streaming demonstration of an uncompressed 4K video using a photonics-based THz communication system (below 200 GHz) is reported, where error-free transmission is achieved at a distance of 1 m. Similarly, THz signals propagating through practical outdoor weather conditions and subject to indoor surface reflections are studied in [231]. According to [232], THz communications can support outdoor communications with proper planning.

Open-source large-scale distributed testbeds are being developed to facilitate experimental research in the mmWave and THz bands, such as MillimeTera [233]. Several experimental testbeds that demonstrate multi-Gbps THz links over several
distances have been reported for both electronic (at 240 GHz [234], 300 GHz [235], 625 GHz [236] and 667 GHz [237]) and photonic systems (near 300 GHz [238]). Other real-time testbeds are proposed in [239] (30 Gbit/sec at 325 GHz) and in [240] (100 Gbit/sec at 300 GHz). In [241], a versatile experimental testbed debunked the limited THz transmission distance myth by demonstrating a more than 1 Km distance at 100 GHz. Furthermore, THz-wireless fiber extenders are experimentally demonstrated at 300 GHz in [242], [243], and a spectral-efficient 64-QAM-OFDM THz link is demonstrated in [244]. However, we still lack experimental testbeds for THz UM-MIMO systems [245]. Nevertheless, thousands of graphene-based antennas promise to be embedded in small arrays, and electronic solutions promise high degrees of integration, which motivates research into UM-MIMO. The challenge is how to operate such arrays.

V. THz MODULATION SCHEMES AND WAVEFORM DESIGN

Designing efficient THz-specific waveforms and modulation schemes is crucial for unleashing the THz band’s true powers. Specific waveform designs can mitigate the limitations in THz sources and receivers. In contrast, optimized, adaptive modulation schemes can fully exploit the available spectrum. The choice of modulation schemes provides a compromise between low-complexity and high-rate PHY layer configurations.

Because carrier-based systems at higher frequencies tend to use larger spectrum bandwidths per channel, simple modulation schemes that require very low complexity digital demodulation are favored (binary phase-shift keying and amplitude-shift keying). Nevertheless, novel modulation schemes and optimized multi-carrier waveform designs should be tailored for specific THz communication use cases. Early results on IEEE 802.15.3d-compliant waveforms are documented in [83], which proposes an optimal THz envelope waveform that maximizes the spectral radiation efficiency under stringent emission constraints.

A. Single-Carrier versus OFDM

The use of multi-carrier waveform designs in future THz-band communications systems remains a controversial topic. The limited multipath components at higher frequencies result in frequency-flat channels that favor low-complexity wideband SC systems [246], [247]. Because THz beams are narrow under high antenna gains, the corresponding delay spread is reduced (survival of a single path), and the channel should be flat. This may require a deviation from OFDM. OFDM is challenging to implement in the context of ultra-broadband and ultra-fast THz systems (complex transceivers with Tbps digital processors still do not exist). The strict synchronization at the THz band severely hinders OFDM deployment, where frequency synchronization requires sampling rates on the order of multi-giga- or tera-samples per second [187]. Furthermore, the high peak-to-average power ratio (PAPR) requirements also render OFDM ineffective in the THz band. The limitations of DACs and ADCs also prevent the digital generation of multi-band orthogonal systems [140]. Special care should be given to mitigate the resultant Doppler effect with OFDM. Under perfect frequency and time synchronization, the cyclic prefix length in OFDM is chosen to account for the delay spread in the system, and the OFDM symbol length is proportional to the inverse of the Doppler spread. Hence, at THz frequencies, the cyclic prefix is relatively larger for the same delay spread conditions. This problem is highlighted in [122] and complicates OFDM design. Prospect solutions for this problem in the literature include beam-based Doppler frequency compensation schemes [248], [249]. Although variations of multi-carrier designs result in the use of distinct THz spectra, spectral efficiency need not be the primary concern in the presence of huge bandwidths, and baseband complexity constraints might be more critical.

SC modulation for above-90 GHz is proposed as a spectral- and energy-efficient solution for Tb/s wireless communications [246]. Nevertheless, because bandwidth is abundant in the THz band, non-overlapping and perhaps equally-spaced sub-windows are efficient, as confirmed in several photonic THz experiments [97], [149]. Non-overlapping windows can be understood as SC modulation with some form of carrier aggregation, which is much less complex than OFDM and would thus enable the use of high-frequency energy-efficient power amplifiers. Transmission can be conducted in parallel over these windows [250], where each carrier would occupy a small chunk of bandwidth that supports a lower data rate. This relaxes design requirements (simpler modulation and demodulation) and reduces energy consumption while retaining an overall high-rate THz system. These benefits are at the cost of operating multiple modulators in parallel, where a high-speed signal generator is required to switch between carriers.

SC transmission is further advocated in [251]. SC has been proposed in WiFi 802.11ad (WiGig) [252] for mmWave communications [253]. The SC waveform can be complemented with simple continuous phase modulation (CPM) schemes such as continuous phase modulated SC frequency division multiple access (CPM SC-FDMA) and constrained envelope CPM (ceCPM-SC) [254]. SCs can even provide higher power amplifier output power than cyclic-prefix OFDM. Furthermore, SC transceivers are resilient to PN [183], [190], especially when combined with a PN-robust modulation scheme. Low-order modulations have low PAPR and are robust to PN. By tracking the phase of local oscillators at the transmitter and receiver (e.g., using time-domain phase tracking reference signals), SCs can accurately estimate the PN at low complexity.

Frequency selectivity, however, can still arise due to frequency-dependent molecular absorption losses, frequency-dependent receiver characteristics, and the existence of several multipath components in indoor sub-THz systems. The bandwidth of each THz transmission window (approximately 0.2 THz [97]) can be much larger than the coherence bandwidth, which at 0.3 THz can be as low as 1 GHz in a multipath scenario and can reach 60 GHz with directional antennas [99]. Frequency selectivity is a function of the communication distance, pulse bandwidth, and center frequency [187]. A frequency-selective system can also exist because of the behavior of the THz receivers [179], [255].

Therefore, as a well-understood technology, OFDM can still be used for THz communications [244]. Candidate multi-
carrier schemes include cyclic-prefix OFDM (CP-OFDM), block-based discrete-Fourier-transform spread OFDM (DFT-s-OFDM) \[256\], continuous offset QAM-based filter-bank multicarrier (OQAM/FBMC) \[257\], and orthogonal time-frequency modulation (OTFS) \[258\]. A study on the performance and complexity trade-offs of such schemes in both the sub-THz and THz bands still lacks in the literature. Target parameters include PAPR performance, the complexity of different linear equalization schemes, and the impact of subcarrier spacing on PN.

OQAM/FBMC is considered a promising candidate for future wireless networks and cognitive radio applications as an alternative to CP-OFDM. OQAM/FBMC features very low out-of-band emissions, with the same bit error rate as CP-OFDM, but with enhanced spectral efficiency because no cyclic prefix is needed. Moreover, OQAM/FBMC relaxes synchronization requirements and offers less sensitivity to Doppler effects due to the use of well-localized prototype filters in the time-frequency domain such as PHYDYAS \[259\]. In contrast, all these advantages reflect negatively on the complexity of the equalizer because CP is not used to reduce ISI. The direct form representation of FBMC consists of OQAM pre-processing, synthesis filter bank, analysis filter bank, and OQAM post-processing.

OTFS \[258\] is promising for THz mobility scenarios with Doppler shifts that severely affect the automatic frequency control range. OTFS is robust to doubly selective channels because it modulates information symbols in the delay-Doppler domain. Alternative THz multi-carrier modulations such as wavelength division multiplexing (WDM) and Nyquist WDM \[260\] are also being studied.

### B. Optimized Modulation Schemes

Modulation schemes can be further optimized to use the fragmented THz bandwidth to mitigate the absorption effect and turn it into an advantage. Accordingly, and because the THz channel response is distance-dependent, distance-aware multi-carrier schemes that dynamically optimize transmission window allocations are proposed in \[250\]. Such schemes achieve Tbps data rates, an order of magnitude higher than fixed-bandwidth modulation schemes, over medium-range communications (10 m). Nevertheless, such schemes come at the expense of a slightly increased complexity because they typically require a control unit, a multi-carrier modulator, and a feedback path. In \[261\], a parallel sequence spread spectrum is proposed as an alternative to OFDM for THz communications; it achieves 100 Gbps with simple receiver architectures that can almost be implemented almost entirely in analog hardware. In \[262\], constrained PSK is introduced as an energy-efficient modulation scheme for sub-THz systems.

Many more resources, however, can be dynamically optimized. For example, frequency allocation per AE is optimized in \[89\] to maximize capacity as a function of the number of frequencies and AEs, antenna and array gains, and beamsteering angles. Moreover, a pulse-based multi-wideband waveform design is optimized in \[90\] to enable communication over long-distance networks by adapting the power allocation criteria over a variable number of frames. The design incorporates pseudo-random time-hopping sequences and polarity randomization and accounts for temporal broadening effects and delay spread; a communication range of 22.5 m and a data rate of 30 Gbps are reported. A single-user and multi-user distance-aware bandwidth-adaptive resource allocation solution is proposed in \[263\] that supports a data rate of 100 Gbps over a 21 m distance. The unique relationship between distance and bandwidth is also exploited to enable a multi-carrier transmission in \[105\].

A hierarchical modulation scheme is proposed in \[140\] for a single-transmitter multiple-receiver system that supports multiple data streams for various users at different distances, by adapting the modulation order and symbol time. These optimization problems are extended to cases of large densities and user mobility. In such scenarios, multi-user interference is unavoidable. In \[205\], the opportunistic use of resources under mobility is maximized to satisfy the constraints on humidity, distance, frequency bands, beamwidths, and antenna placement. Furthermore, a stochastic model of multi-user interference is proposed in \[264\], alongside modulation schemes that minimize the probability of collisions. THz OFDM adaptive distance- and bandwidth-dependent modulations are also considered in \[144\].

### C. Pulse-Based Modulation

Although continuous carrier-based transmission can be supported in the sub-THz range when the constraint on size is relaxed \[36\], \[265\], carrier-based transmission is still challenging at true THz frequencies. For example, it is not easy to generate more than short high-frequency pulses of few milli-watts with graphene at room temperature. Nevertheless, with large bandwidths, a reduction in spectral efficiency is acceptable, and pulse-based modulations can be used. Pulse-based SC on-off keying modulation spread in time (TS-OOK) that exchanges hundreds of femtosecond-long pulses between nanodevices is proposed in \[73\]. By assuming time-slotted operations with a time slot \(T\), the Gaussian pulse is expressed as \(p(t) = a \exp \left(-\left(t - b\right)^2/2T_p^2\right)\), where \(a\), \(b\), and \(T_p\) are the amplitude, center, and spread of the pulse, respectively (Gaussian pulse has a duration \(\tilde{T}_p < \tilde{T}\)). Conversely, the raised cosine pulse of the carrier-based system is expressed as \(q(t) = \text{sinc} \left(t/\tilde{T}\right) \left(\cos \left(\pi \tilde{\alpha}t/\tilde{T}\right)\right) / \left(1 - (2\tilde{\alpha}t/\tilde{T})^2\right), \) where \(\tilde{\alpha}\) is the roll-off factor (0 ≤ \(\tilde{\alpha}\) < 1).

Pulse-based THz communications can achieve a Tbps data rate in nanonetwork scenarios \[73\]. They have also been used in ultra-wide-band impulse-radio systems \[266\] and free-space optics \[267\]. However, pulse-based systems are power-limited, especially for extremely wideband signals. Wideband pulses are thus typically used to achieve low-power, compact, and low-complexity sub-band transmissions. Nevertheless, by jointly optimizing the modulation and power allocation iteratively, Tbps rates are demonstrated in indoor communication paradigms in \[268\] by assuming realistic transmit, receive, and equalization filters under practical error rate constraints. An SC pulse-based approach is first addressed by optimizing the choice of modulation scheme by assuming a very
long dispersive channel impulse response that accounts for ISI. Then, frequency division over multiple orthogonal sub-bands is considered, alongside efficient power allocation, to minimize the loss in the rate that is caused by finite-alphabet modulations.

VI. RECONFIGURABLE UM-MIMO ARRAYS

Antenna gains and array gains are pivotal for overcoming path loss, where specific combinations of these gains are recommended for a specific communication distance and frequency of operation. Although a mmWave system typically requires a footprint of few square centimeters for several tens of antennas (not even sufficient to overcome the path loss over few tens of meters), a vast number of AEs can be embedded in a few square millimeters at THz frequencies. Furthermore, given the quasi-optical behavior under LoS dominance, a THz MIMO channel is sparse with multi-user beamforming and low-rank with spatial multiplexing.

As described in Sec. II, due to high directivity and because beamforming is typically configured at the level of AEs within a SA, each SA is effectively detached from its neighboring SAs in a multi-user setting. Moreover, the role of baseband precoding reduces to defining the utilization of SAs, or to simply turning SAs on and off. However, in a point-to-point setup, the SA paths are highly correlated, and the channel is ill-conditioned. Nevertheless, good multiplexing gains are still achievable using sparse antenna arrays [269] that reduce spatial correlations in point-to-point LoS scenarios. The capacity of LoS MIMO uniform linear array channels is studied in [270] over all possible antenna arrangements.

A. THz Spatial Tuning

To enhance THz LoS channel conditions, spatial tuning techniques that optimize the separations between AEs can be applied [98]. Depending on the communication range (d) and the SA separation (Δ), three modes of operation can be distinguished: (1) a mode where Δ is large enough so that the channel paths are independent and the channel is always well-conditioned, (2) a mode where d is large compared to Δ, and therefore the channel is ill-conditioned, and (3) a mode where d is much larger than Δ, larger than the Rayleigh distance, where the channel is highly correlated. If the communication range is less than the Rayleigh distance in the second mode, Δ can be adapted to enhance the channel conditions and achieve near-orthogonality [269], [270]. Hence, the Rayleigh distance is an important metric for capturing THz system performance. For the number of transmit antennas N and receive antennas M, this distance is expressed as [271]

\[ d_{\text{Ray}} = \max(M, N) \Delta_{r} \Delta_{t} / \lambda, \]

where λ is the operating wavelength, and Δr and Δt are the uniform separation between SAs (also applicable to AEs) at the receiver and transmitter, respectively.

The LoS Rayleigh distances are illustrated in Fig. 3 as a function of array dimensions, operating frequencies, and antenna separations. Two AoSA configurations are simulated, \( M = N = 128 \times 128 \) and \( M = N = 2 \times 2 \) by assuming Δr = Δt. For small antenna separations, a large number of antennas are required to achieve multiplexing gains beyond several meters. For the same Δ, higher frequencies and larger arrays result in larger Rayleigh distances. However, for the same footprint, increasing the number of antennas incurs a reduction in Rayleigh distance that is quadratic in Δ. By finely tuning Δ, multiple data streams over eigenchannels can be transmitted when [98]

\[ \Delta_{\text{opt}} = \sqrt{\frac{\Delta_{r} \Delta_{t}}{M}}, \]

for odd values of z. Such optimizations cannot be achieved in the third mode due to the limitation in physical array sizes. Alternative optimization schemes for antenna separations in LoS communications beyond the Rayleigh distance have been studied in [271], [272].

When combined with numerical optimization, we denote the corresponding adaptability in design by “spatial tuning.” Spatial tuning is typically illustrated in the context of plasmonic antennas with a sufficiently large uniform sheet of AEs and where optimal configurations can be tuned in real-time [98], [273]. For example, a graphene-based sheet can consist of hundreds of uniformly-spaced active graphene elements mounted on a dielectric layer, which itself is mounted on a common metallic ground [84]. Hence, AEs can be contiguously placed over a 3D structure, and SAs can be virtually formed and adapted. For the desired communication range, a required number of AEs per SA is allocated. Then, the number of possible SA allocations, bounded by array dimensions and the number of RF chains, dictates the diversity/multiplexing gain.

Spatial tuning can be extended to include multi-carrier design constraints. For instance, nanoantenna spacings in plasmonic antenna arrays can be reduced to λ_{SPP} while still avoiding the effects of mutual coupling. Mutual coupling in graphene-based THz antenna arrays is studied in [102], [274]. With the couple mode theory, the impact of mutual coupling on the response of nanoantennas is modeled via a coupling coefficient. Even at separations much less than λ, near field mutual coupling is negligible. This promising realization enables the
practical implementation of compact THz systems in small footprints. In [274], the use of a frequency-selective surface structure that can be mounted between the array elements of an UM-MIMO array (behaving as a spatial filter) is proposed to reduce the mutual coupling effects further to negligible values.

Placing AEs very close to each other, however, is not always beneficial because it reduces the spatial resolution and the achievable multiplexing gains. Furthermore, the inter-AE separation distance should not exceed $\lambda/2$, beyond which grating-lobe effects arise. By setting the separation between two active AEs to be $\lambda/2$, all the AEs in between would be idle. These antennas can be used for several purposes. Besides using them to increase the array gain, they can be configured to operate at different frequencies in a multi-carrier scheme that supports more users with the same array footprint, as illustrated in Fig. 4. Alternatively, neighboring AEs can be configured to operate at the same frequency in a spatial oversampling setup [24], lowering the spatio-temporal frequency-domain region of support of plane waves [275]. The latter approach can be exploited for noise shaping, which results in a reduced noise figure and increased linearity.

**B. Index Modulation and Blind Parameter Estimation**

Spatial modulation (SM) schemes for THz communications are promoted in [98], [276] as power- and spectrum-efficient solutions. The efficiency of SM at higher frequencies is highly dependent on the array design and channel conditions, as illustrated in [277]–[279] for mmWave systems. By mapping information bits to antenna locations adaptively, hierarchical SM solutions can be designed at the level of SAs or AEs. With SM, the number of bits that can be accommodated in a single channel use is

$$N_b = \log_2 (M_f N_f) + \log_2 (Q^2) + \log_2 (|X|).$$

The transmitted binary vector over one symbol duration can be expressed as $b = [b_m b_q b_s]$ $\in$ $\{0, 1\}^{N_b}$, where $b_m \in \{0, 1\}^{\log_2(M_f N_f)}$ represent SA selection, $b_q \in \{0, 1\}^{\log_2(Q^2)}$ represent AE selection, and $b_s \in \{0, 1\}^{\log_2(|X|)}$ correspond to the actual QAM symbol. The number of AEs and SAs and the constellation size can be tuned for the desired bit rate in such a design. By enabling the selection of various combinations of antennas simultaneously, a generalized index modulation scheme can be defined [280], [281]. Typical massive MIMO SM and generalized SM solutions [282], [283] should be revisited in the ultra-massive THz context. Furthermore, when enabling adaptive antenna-frequency maps, generic index modulation (IM) solutions that take full advantage of the available resources can be configured, in which information bits are also mapped to frequency allocations. The number of bits per channel use with IM can increase to

$$N_b^{(IM)} = \log_2 \left( \frac{F}{\bar{F}} \right) + \log_2 \left( \frac{S}{M_f Q^2} \right) + \log_2 (|X|),$$

where $F$ is the total number of available narrow frequency bands, and $\bar{F}$ and $S$ are the numbers of frequencies that can be supported and antennas that can be activated at a specific time, respectively.

Even more generalized IM schemes can be achieved by jointly designing the spatial and frequency bit maps. Such designs could confirm to be particularly efficient in the THz band because a considerable number of AEs can fit in small footprints. The fragmented nature of the THz spectrum enables allocating multiple absorption-free spectral windows concurrently. However, the efficiency of these adaptive schemes is limited by the speed at which frequency hops can be executed. Nevertheless, changing the frequency of operation in the THz band can be achieved quickly without changing the physical dimensions of the transmitting antennas. Simple material doping or electrostatic bias change the Fermi energy of graphene, which dictates the frequency of operation. Software-defined plasmonic metamaterials are also a candidate solution, primarily for frequencies below 1 THz. For SC systems with IM, constant-envelope modulations such as CPM are power-efficient [277].

Such design compactness and flexibility can be further enhanced when complemented by THz-specific signal processing...
techniques at the receiver side. Instead of communicating transmission parameters with the receiver, blind parameter estimation can be conducted. For instance, in \cite{284}, a tertiary hypothesis test based on power comparison for THz antenna index and modulation mode detection is proposed and analyzed, alongside low-complexity frequency index detectors and modulation type estimators. Information bits can be assigned for the choice of modulation type as well. Although modulation classification \cite{285,286} is a classical signal processing problem, its applicability to the THz-band \cite{287} can be particularly beneficial. Given the enormous possibilities of map-bit combinations, compressed sensing and machine learning techniques can be applied for detection and estimation.

VII. BEAMFORMING AND PRECODING

As previously discussed, beamforming and precoding are critical to overcome the high path losses at high frequencies and exploit the THz channel’s distance- and frequency-dependent characteristics. Beamforming enhances power versus distance such that the transmit power of CMOS THz devices can suffice for communication purposes. Because maintaining alignment is much harder at THz frequencies due to collimation, THz beamforming schemes should be fast. By the time 6G matures, fully digital arrays operating at mmWave frequencies and below will likely be readily available and capable of achieving near-optimal beamforming performance. However, it is unlikely that this would be the case at sub-THz and THz frequencies.

The motivation for hybrid beamforming in the THz realm is similar to that in the mmWave realm \cite{288}—there is a bottleneck in achieving prohibitively-complex and high-power-consuming fully-digital arrays (as highlighted in Sec. I-B) power consumption remains a significant hurdle for the practical deployment of THz systems). THz hybrid beamforming is described in \cite{97}, where a distinction is made between a tactical deployment of THz systems). THz hybrid beamforming is a modification of (1), where the precoding and combining power consumption, near-optimal and low-complexity THz hybrid precoding algorithms are proposed in \cite{147}. The design problem is divided into two sub-problems: hybrid DAoSA precoding problem and switch selection. The system model corresponding to hybrid precoding in DAoSA architectures is a modification of (1), where the precoding and combining matrices, \( W_f^H \) and \( W_r^T \), are decomposed into their digital and analog components as

\[
y = \sqrt{\bar{p}}C_A^H\tilde{H}P_DP_Dx + C_A^Hn,
\]

where \( \bar{p} \) is the transmitting power and \( y \) and \( x \) here have the dimension \( N_r \times 1 \), with \( N_r \) being the actual number of data streams, \( H \in \mathbb{C}^{M_r,N_r,Q^2 \times M_r,N_r} \) and \( n \in \mathbb{C}^{M_r,N_r,Q^2 \times 1} \). Furthermore, \( C_A \in \mathbb{C}^{M_r,N_r,Q^2 \times M_r,N_r} \) and \( C_D \in \mathbb{C}^{M_r,N_r \times N_r} \) are the analog and digital combining matrices, and \( P_A \in \mathbb{C}^{M_r,N_r,Q^2 \times M_r,N_r} \) and \( P_D \in \mathbb{C}^{M_r,N_r \times N_r} \) are the analog and digital precoding matrices, respectively. The achievable rate of this system model (subject to optimization) can be expressed as

\[
R = \log_2 \left( |M_r,N_r,Q^2| + \frac{\bar{p}}{N_r \sigma^2} \tilde{H}P_DP_D(P_AP_D)^{H}H^{H} \right),
\]

where \( | \cdot | \) is the matrix determinant operator. Another dynamic SA architecture is proposed in \cite{299}, which analyzes both quantized- and fixed-phase shifters.

Hybrid beamforming often entails user grouping and SA selection. In \cite{149}, a THz multi-carrier distance-dependent hybrid beamforming scheme is proposed, in which user grouping is achieved through analog beamforming, alongside digital beamforming, power allocation, and SA selection mechanisms. The proposed solution enables users of different user groups
Fig. 5: High-frequency hybrid beamforming architectures: Fully-connected, fixed AoSAs, and dynamic switching.

to share frequencies while avoiding interference in the analog domain. Users in the same group, however, are assigned orthogonal frequencies based on a distance-aware multi-carrier scheme. SAs are then assigned to the data streams of a user group in the digital domain. The same authors address the hybrid beamforming problem for THz indoor scenarios in [167]. A bound on the ergodic capacity is derived, and the impact of random phase shifter errors is analyzed. The relation between the required size and number of SAs and the communication distance is also established. The spectral efficiency gap between hybrid and digital precoding is smaller when the channel is sparser [300], making hybrid schemes suitable for multipath-limited THz channels. In [301], joint inter- and intra-multiplexing and hybrid beamforming are proposed for widely-spaced multi-SA THz systems.

Developing wideband hybrid beamforming schemes for THz communications is also essential. For instance, in [302], the researchers propose an OFDM-based normalized codebook search algorithm for beamsteering and beamforming in the analog domain and a regularized channel inversion method for precoding in the digital domain. Two digital beamformers are used in a three-stage scheme to account for the loss in performance due to hardware constraints and the difference between subcarriers. Similarly, in [122], a beam division multiple access scheme is proposed for wideband massive MIMO sub-THz systems, which schedules a mutually non-overlapping subset of beams for each user. The algorithm is based on per-beam synchronization in time and frequency, considering the delay and Doppler frequency spreads, the latter of which are orders-of-magnitude larger at THz frequencies. An inaccurate narrowband assumption, in which the precoding and combining matrices of a wideband system are designed for a specific carrier frequency, induces the effect of beam split in THz communications. This phenomenon is mitigated in [303] using a THz-specific delay-phase controlled precoding mechanism, in which time-delay components are introduced between the RF chains and the phase shifters. The addition of such components creates frequency-dependent beams uniformly aligned with the spatial directions over the entire bandwidth.

In addition to the considerations above for beamforming in the THz band, all of which are extensions to similar considerations in the mmWave band, recent studies consider novel THz-specific beamforming schemes based on novel THz circuitry. For instance, in [304], a graphene-based dense antenna array architecture is proposed, in which each element is integrated by a THz plasmonic source, direct signal modulator, and nanoantenna. For such an architecture, novel dynamic beamforming schemes at the level of single elements and the level of the integrated array are proposed, where full phase and amplitude weight control can be achieved by tuning the Fermi energy of the modulator and AE. The researchers propose a codebook design for Fermi energy tuning that results in reasonably accurate beamforming and beamsteering; the power-density of the array increases non-linearly with its size. Controllable THz frequency-dependent phase shifters can also be achieved via low-loss integrally gated transmission lines [291], the length of which determines the signal travel time, and hence the phase shift. Furthermore, graphene/liquid crystals have been proposed for magnet- or voltage-controlled THz phase shifters [305], [306]. Because they are digitally controlled, such phase shifters only generate quantized angles.

Several other beamforming considerations need to be considered for future THz networks. For example, in a cell-free massive MIMO scenario [307], distributed access points can each provide an excess of 100 GHz to a user, especially under low mobility. Dense deployment of access points can guarantee short THz communication distances, even under significant blockage. Mobility and blockage are addressed in [308] in the context of network-massive MIMO scenarios in the mmWave and THz bands, where per-beam synchronization is proposed to mitigate the channel Doppler and delay dispersion; precoding beam-domain power allocation is reduced to a network sum-rate maximization problem. In other studies with system-level considerations, distance-aware multi-carrier hybrid beamforming based on beam division multiple access is proposed in [309] to enable massive micro-scale THz networks; relay-assisted THz hybrid precoding designs are proposed in [310].

B. One-Bit Precoding

The circuit power consumption, hardware complexity, and system cost significantly increase at higher frequencies and ultra-massive dimensions. The dominant sources of power consumption are ADCs in the uplink and DACs in the downlink. The power dissipation in converters scales exponentially in the number of resolution bits, and state-of-the-art DACs and ADCs can only achieve 100 gigasamples-per-second rates
Furthermore, the capacity requirements on the fronthaul interconnect links are also severe in large MIMO systems. Jointly reducing system costs, power consumption, and interconnect bandwidth with minimal performance degradation remains a challenge.

As an alternative to reducing the number of converters using hybrid beamforming, the bit resolutions can be reduced through coarse quantization. This approach has the extra benefit of lowering the linearity and noise requirements, which is crucial in THz settings. In the extreme case of one-bit quantization, only simple comparators are required; automatic gain control circuits are no longer required. For high-amplitude resolutions, the power consumption of ADCs grows quadratically with the sampling rate. A one-bit quantization is thus more likely to be conducted over point-to-point doubly-massive MIMO links.

In such scenarios, the concept of multiple access reduces to superposition coding of multiple data streams over a single link. Nevertheless, calling the resultant configuration NOMA is not a misnomer because each THz beam can still be configured to serve multiple users. In this scenario, the role of NOMA can be essential in mitigating the hardware constraints in THz devices that limit the beamforming capabilities.

For power-domain NOMA, and based on our system model, multiple data streams can be sent concurrently via superposition coding over different combinations of transmitting and receiving SAs that form overlapping effective channel matrices. Assume, for ease of construction, that the superposition of data symbols occurs at the lower layers of the MIMO channel matrix. Denote by $S$ the set of power-domain multiplexed data streams of dimensions $S_i$, $i = 1, \ldots, |S|$, such that $S_1 \geq S_{i+1}$ and $S_1 = M_1 N_1$. The multiplexed transmitted symbol vector $\bar{x}_t = [x_1^T \cdots x_n^T \cdots x_n^T] \in \mathbb{C}^{S_1}$ is allocated the contiguous set of SAs $N - S_i + 1$ to $M_i N_i$. We thus use the effective channel matrices $H_i \in \mathbb{C}^{M_i N_i \times S_i}$, comprised of the columns $M_1 N_i - S_i + 1, M_1 N_i - S_i + 2, \ldots, M_1 N_i$ of $H$. The equivalent baseband input-output system relation can then be expressed as

$$\mathbf{y} = \sum_{i=1}^{|S|} H_i \bar{x}_i + \mathbf{n},$$

in which NOMA is achieved by assigning different power levels to the multiplexed transmitted symbol vectors.

For example, we can allocate a higher power level $p_i$ to the symbol vectors $i$ of smaller dimensions, i.e., $p_i < p_{i+1}$. Each symbol $x_i^T$ thus belongs to a scaled complex constellation $X_i^T$ ($\mathbb{E}[x_i^T x_i^H] = p_i$), and we use $x_i \in X_i$, the lattice that includes all possible symbol vectors generated by $S_i$ $X_i$ constellations. With higher degrees of reconfigurability in THz antenna arrays, generalized coding schemes encompassing antenna selection, frequency selection, and power allocation can be achieved. However, low-complexity detection and decoding schemes should complement such designs, particularly efficient successive interference cancellation at the receiver.

When sufficient multipath components exist, perhaps in scenarios where lower antenna gains are required (indoor sub-THz scenarios, for example), conventional single-cell multiuser MIMO-NOMA settings can still be achieved. Assume that the cellular users are divided into two groups, where the first group of users is uniformly distributed in an inner disk ($C_1$) centered at the base station and of radius $R_N$, and the second group of users is uniformly distributed in an outer disk ($C_2$) from $R_N$ to $R_C$. A base station with $N$ transmitting antennas simultaneously services two users, user 1 with $M_1$ antennas in $C_1$ and user 2 with $M_2$ antennas in $C_2$, in the same frequency and time slot via power-domain superposition coding. The received vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ at user 1 and user 2 can be expressed as

$$\mathbf{y}_1 = H_1 \mathbf{x}_1 + H_1 \mathbf{x}_2 + \mathbf{n}_1,$$

$$\mathbf{y}_2 = H_2 \mathbf{x}_1 + H_2 \mathbf{x}_2 + \mathbf{n}_2,$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the power-multiplexed transmitted symbol vectors and $H_1$ and $H_2$ are the corresponding channel matrices.
In this scenario, NOMA is achieved by clustering users from the inner disk $C_1$ with users from the outer disk $C_2$ and assigning different power levels to the multiplexed transmitted symbol vectors.

Adaptive superposition coding and subspace detection are considered for THz NOMA in [273]. Energy efficiency in a THz MIMO-NOMA system is also addressed in [322] by optimizing the user clustering, hybrid precoding, and power allocation mechanisms. Furthermore, energy-efficient resource allocation in downlink THz-NOMA systems is studied in [323]. Although the attainable gains of power-domain MIMO-NOMA remain unclear, THz scenarios suggest a compelling use case.

In their recent study, [324], the researchers argue that MIMO-NOMA can misuse the spatial dimension by incurring a multiplexing gain loss because of fully decoding streams in SIC. Such loss is particularly evident when comparing MIMO-NOMA to other MIMO schemes, such as conventional multi-user linear precoding (MU-LP) or rate splitting (RS). Nevertheless, compared to OMA, the benefits of NOMA are clear. Furthermore, the proposed efficient SIC using subspace detectors in [273] combats this multiplexing gain reduction. Spatial precoding in MU-LP fails to overcome inter-stream interference under ill-conditioned near-singular THz channels matrices. Accordingly, power-domain NOMA is a vital enabler for THz data multiplexing.

VIII. THz BASEBAND SIGNAL PROCESSING

Efficient baseband signal processing is critical for mitigating the impairments of novel THz-band devices and enabling operations beyond 100 Gbps [325]. The true bottleneck at the baseband is the lack of energy-efficient transceivers that can approach a Tbps data rate [87], where the sampling frequency is still on the order of tens of gigasamples-per-second in state-of-the-art ADCs and DACs. Efficient signal processing across all baseband blocks is required to fill this gap. Due to Moore’s Law’s diminishing effect, limited advances in chip power density and baseband computations are expected from silicon scaling. Therefore, incorporating application-specific integrated circuit (ASIC) architectures in a holistic framework is vital for achieving Tbps data rates and decreasing the time-to-market of THz-operating equipment. Joint algorithm and architecture optimization across all baseband blocks optimizes latency, area efficiency, power consumption, and overall throughput.

Although channel code decoding dominates baseband computations [88], the entire baseband chain should be optimized. Accordingly, algorithm and architecture co-optimization for synchronization, data decoding and detection, and channel estimation is required. Such optimization can exploit the specific THz propagation characteristics, such as angular and temporal sparsity. Implementing THz-band high-resolution precoding and beamforming algorithms in the baseband’s digital domain is challenging due to massive dimensionalities and the lack of processing capabilities. Such holistic solutions are especially crucial for massive antenna dimensions and high mobility wideband scenarios. We highlight recent advances in low-complexity channel estimation, channel coding, and data detection schemes in the following sub-sections.

A. Channel Estimation

Channel estimation in the THz band is very challenging in mobile scenarios, where accurate CSI is required in beamforming mechanisms and for accurately directing beams to avoid misalignment issues. Accurate CSI is essential in the absence of an LoS path. Furthermore, frequent channel estimation might also be required for fixed LoS point-to-point THz links because, at the micrometer wavelength scale, slight variations in the environment can introduce significant channel estimation errors. Classical channel estimation techniques should also be revisited by considering low-resolution quantization and hybrid analog and digital designs. Several techniques can be considered to reduce the complexity of THz-band channel estimation, such as fast channel tracking algorithms, lower-frequency channel approximations (exploiting outband signals), compressive-sensing-based techniques, and learning-based techniques, to name several.

Compressive-sensing techniques for sparse channel recovery in THz channel estimation [326] are inspired by their successful use in mmWave communications [327], where channels are even less sparse. These approaches can exploit sparsity in the dictionary, delay, or beamspace domains. Researchers investigate variations of greedy compressive sampling matching pursuit [326], orthogonal matching pursuit (OMP) [328], and the least absolute shrinkage and selection operator (LASSO) [329]. Furthermore, in [330], approximate message passing (based on belief propagation in graphical models) and iterative hard-thresholding are argued to be an efficient compressed-sensing-based technique for THz channel estimation. Nevertheless, learning-based THz channel estimation schemes are most efficient at higher dimensionalities [331]. Deep kernel learning based on the Gaussian process regression is explored in [332] for multi-user channel estimation in UM-MIMO systems over 0.06 – 10 THz.

Despite channel sparsity, the real-time THz channel estimation complexity overhead can be significant in a dense multi-user wideband scenario with many paths; a large number of measurements might be required for compressive-sensing-based estimation. Accordingly, traditional minimum mean square error (MMSE) and least square channel estimation methods can be used to estimate the second-order statistics of THz channels [97]. Furthermore, joint activity detection and channel estimation is an efficient technique to reduce the use of pilots and the complexity of computations in wideband random massive-access THz systems [333]. Fast channel tracking is an alternative approach to reduce the channel estimation overhead in high-mobility scenarios, as illustrated in [334] for THz beamspace massive MIMO. Eventually, THz channel estimation algorithms should be complemented by efficient hardware implementations to verify their practicality, as demonstrated in [335] for sparsity-exploiting mmWave beamspace channel estimation algorithms [336].
B. Channel Coding

Data detection and decoding algorithms and architectures require low latency and high energy efficiency and throughput to bridge the Tbps gap in baseband signal processing. They should be highly parallelizable (spatial and functional parallelism) and should exhibit significant data locality and structural regularity. Channel coding is hitting the implementation wall because it is the most computationally demanding baseband process [88]. The three central candidate coding schemes for 6G are Turbo, LDPC, and Polar codes. Although Turbo and LDPC decoders are both executed on data-flow graphs, Turbo decoding is inherently serial, and LDPC decoding is inherently parallel. In contrast, Polar decoding is typically performed on a tree structure and is inherently serial. Due to their parallel nature, LDPC decoders provide higher throughput [337]. However, Polar and Turbo codes provide much flexibility in code rates and block sizes [87], which is much required in 6G.

A modular framework for generating and evaluating high-throughput Polar code decoders is presented in [338], where soft cancellation algorithms are explored. Turbo codes have advanced significantly towards beyond-100-Gbps operations [339]–[341]. Achieving a Tbps throughput with Polar codes is also addressed in [342], where low-latency majority logic and low-complexity successive cancellation are combined for decoding, alongside an adaptive quantization scheme for log-likelihood-ratios (LLRs). This scheme achieves Tbps in a 7 nm technology implementation while occupying a 10 mm² chip area and consuming 0.37 W of power. Furthermore, Polar codes complemented with guessing random additive noise decoding (GRAND) are shown to be computationally efficient for short-length, high-rate codes [343], [344], which is promising for THz-based control channel communication, especially if further knowledge on THz system noise is developed.

Although such advances in coding schemes serve the ultimate goal of THz communications, which is achieving Tbps operations, they are blind to the inherent characteristics of THz channels. Nevertheless, the THz channel can be considered in MIMO detection schemes and joint modulation, coding, and detection algorithms and architectures.

C. Data Detection

Although channel code decoding is the most computationally demanding baseband processing block, data detection also adds a significant computational burden, especially in doubly-massive MIMO systems. Channel hardening occurs in conventional massive MIMO systems at lower frequencies, with many antennas at the base station and several antennas at the receiving equipment. With channel hardening, simple linear detection schemes such as zero-forcing and MMSE can achieve near-optimal performance [355]. This is not the case in THz systems, where symmetric doubly-massive MIMO systems are common [101], especially because compact large THz antenna arrays can be embedded in the UE. In the latter scenario, the channel tends to be highly correlated, especially under THz LoS-dominance. Inter-channel interference prohibits using simple linear detection schemes that fail to decouple spatial streams and result in noise amplification.

Consequently, more sophisticated non-linear detection schemes should be considered. However, the complexity of optimal non-linear detection schemes that achieve near-maximum likelihood performance is prohibitive at large dimensions. Therefore, novel THz-specific MIMO detectors that can achieve near-optimal performance with reasonable complexity are required. Conventional near-optimal detectors mainly replace the full-lattice search over all candidate transmit vectors in maximum likelihood detection with a reduced search over a reduced space of vectors closer to the truly transmitted vector (from a Hamming distance perspective). Such reduced-complexity detectors are mainly variations of sphere decoding schemes [346], [347]. Although sequential processing in sphere decoding results in variable complexity and limits parallelism, several algorithmic and architectural optimizations have been proposed [348] to fix its complexity [349]. In [350], the complexity of sphere decoding is reduced by casting memory-bound computations into compute-bound operations, and real-time processing is maintained by using graphics processing units.

However, even fixed-complexity sphere decoding is prohibitively complex if used for UM-MIMO detection. Recently, several detection algorithms suitable for large doubly-massive MIMO systems have been proposed. Such algorithms are based primarily on local search criteria [352], heuristic tabu search algorithms [353], message passing on graphical models [354], Monte Carlo sampling [355], and lattice reduction [356]. The Bell Laboratories Layered Space-Time (BLAST) detection algorithm is also modified to support ultra-high data rates in massive MIMO scenarios in [357]. Furthermore, perturbation-based regularizations can be used for equalization with ill-conditioned channels [358].

One family of detectors in particular that can achieve an acceptable trade-off between performance and complexity in large highly-correlated MIMO channels is the family of subspace detectors [100], [359]–[363]. Subspace detectors mainly exploit channel puncturing to reduce complexity and enhance parallelism. Because the computational cost of MIMO detectors is proportional to the number of nonzero elements in a channel matrix (most detectors involve back-substitution and slicing operations), by puncturing the channel into a specific structure, the detection process can be simplified and accelerated. More importantly, subspace detectors can break the interconnection between spatial streams, significantly enhancing parallelism at a marginal cost of multiple channel decompositions. Channel puncturing can be generalized to channel shortening, which can mitigate ISI in SC THz systems [364]. The performance gap between optimal channel shortening (from an information-theoretic perspective) and channel puncturing can be covered by adding MMSE prefilter and channel-gain compensation stages [365], [366]. Subspace-marginalized belief propagation can also be adopted for high-frequency data detection [367].

Other THz-specific data detectors include envelope- and energy-based detectors [368], which enable direct baseband operations without frequency down-conversion, bypassing
phase impairments and enabling non-coherent detection that is inherently robust to PN. Low-complexity energy detection is studied in [369] for pulse-based systems. Compressed detection with orthogonal matching pursuit is also considered for sparse pulse-based multipath THz communications [328].

MMSE precoding and detection are explored in [370] by assuming sparse channel matrices for a broadband SC THz system. Finite-alphabet equalization can also confirm helpful in THz scenarios. Coarsely quantizing the equalization matrix reduces the complexity, power consumption, and circuit area [371], increasing the speed of most critical baseband tasks, such as downlink precoding and data detection. Such coarse equalization enables all-digital massive multi-user MIMO equalization in mmWave systems [372] and should scale up efficiently with THz transceivers. Indoor THz communications Tbps rates under finite alphabets are demonstrated in [268], where a frequency-division scheme of multiple sub-bands is used to relax the requirements on ADCs and DACs.

D. Joint Coding, Modulation, and Detection

One method of incorporating the effect of the THz channel in decoding is to consider iterative detection and decoding schemes. If MIMO detectors generate soft-output LLRs, these LLRs can be fed as soft inputs to the decoder, whose output LLRs can then be fed again as soft inputs to the detector [373]. However, the extra complexity in iterations and computing soft-output values in the detector should be considered. In the particular case of MIMO detection with Polar code decoding, iterations can be configured per stream detection after each successive decoding cancellation; the output of every step in the decoder can thus be used to enhance channel equalization.

Parallelizable detectors are favored for such designs. However, with parallelizability, the transmission vectors per stream would typically consist of a smaller number of bits. Although Polar and Turbo decoders can cope with this, LDPC decoders might not perform well. Larger modulation types can be considered to increase the number of bits per stream for enhanced decoding, 1024 QAM and beyond [374], for example. However, THz systems do not perform well with higher-order modulations due to the increased complexity and the PN effect. Alternatively, deeply-pipelined MIMO architectures can be used to aggregate data for the decoders, but this comes at the expense of reduced throughput. Nevertheless, because different transmission vectors within a decoder block can be independent, multiple detectors can operate in parallel.

In addition to joint channel coding and data detection, joint modulation design and coding can use resources most efficiently, especially in multi-wideband THz communications. Efficient probabilistic signal shaping techniques [375]—[377] can be used accordingly, especially in highly reconfigurable UM-MIMO systems. In [378], the researchers demonstrate that due to the peculiarity of noise in the THz band, proper use of channel codes can increase single-user and network capacity beyond classical networks with AWGN.

Furthermore, THz-specific coding schemes are also being introduced at the network level. For instance, systematic random linear network coding (sRLNC) is proposed in [379], [380] for generic THz systems, in which coded low rate channels carry redundant information from parallel high rate channels. By tuning the transmission and code rates, the number of channels, and the modulation format, fault-tolerant high-throughput THz communications can be supported at different communication ranges. Furthermore, in the context of THz index modulation, joint data detection and parameter estimation can be executed at the receiver side [381]. In the particular case of subspace detection, the entire process can be parallelized [382], [383].

Finally, THz digital baseband operations of any complexity might be prohibitive in some use cases. Therefore, designing an all-analog THz baseband chain is a reasonable solution. Such design should include, in addition to analog modulation and detection schemes, all-analog decoders [384], which are based on soft-output transmission and detection (bit LLRs represented by currents and voltages). All-analog MIMO schemes can be achieved using a continuous mapping scheme, which is noise-limited but might perform well under interference.

IX. RECONFIGURABLE SURFACES

A. Emergence of IRS Technology

One of the recent hot research topics in wireless communications is creating an intelligent, programmable environment for communication [385]. A straightforward approach is to install large active arrays of AEs, also known as active large intelligent surfaces (LISs) [386]—[388] on indoor and outdoor walls and other structures. This approach is a special case of UM-MIMO and is suitable for THz scenarios. With minimal restrictions on how to spread antennas over a surface, the mutual coupling effects can be avoided, and channel correlation can be reduced in LoS environments (spatial tuning of Sec. VI-A can be easily configured). Furthermore, channel estimation and feedback mechanisms can be easily achieved in active LIS setups, essential for achieving low-latency THz communications. Another form of large surfaces is the concept of holographic MIMO surfaces [389]. Because holographic MIMO borrows techniques from the optical domain, its implementation in the THz band should be more convenient than at lower frequencies [390].

Passive IRSs [391], [392] are gaining most of the attention. IRSs are typically implemented using reflective arrays or software-defined metasurfaces, which introduce phase shifts at the level of reflecting elements (non-specular reflections) to focus and scale up the power of reflected signals and steer beams in a particular direction. These results can be achieved without requiring complex encoding and decoding schemes or additional RF operations [393].

Phase shifts can also be implicitly achieved by tuning the impedance or length of delay lines. Metasurface elements might be much smaller than those of reflectarrays (which typically follow the half-wavelength rule). Hence, they support more functionalities, such as polarization manipulation and absorption of incident waves. Reflections from tiny reflecting elements (of sub-wavelength size) form scattering in all directions, of which the combined effect is beamforming. The control complexity of metasurfaces, however, can be higher.
Compared to LISs, both types of IRSs are passive (LIS and IRS are used interchangeably in the literature). Nevertheless, IRSs should be electronically active at some level, perhaps to transmit pilots and for operational purposes.

IRS systems are particularly favorable in the THz band, where they can introduce controlled scattering to extend the very limited achievable communication distances and enable multicasting. Hence, IRSs can add synthetic multipath components to enhance the performance of multipath-limited THz systems. The necessity for IRSs operating at THz frequencies also arises from the fact that regular coherent large antenna arrays are not easily achieved with the very small size of AEs. Even the relaying technology at THz is not mature. IRS systems are thus a viable solution.

Another argument for IRS THz deployments stems from the limitations of surfaces themselves. For IRSs to achieve SNRs comparable to those of massive MIMO, or outperform a classical half-duplex relay system, a large number of reflecting elements is required. This could result in physically large arrays that are harder to deploy and subject to beam squinting. The near-field behavior of IRSs is studied in [165], [394], where a physically accurate near-field channel gain expression is derived for planar arrays, also considering the mismatch between the incident wave and the polarization of antennas. At THz frequencies, however, an electronically-large IRS (compared to the operating wavelength) can be achieved in very small footprints, suggesting that dense THz-band IRS deployments for short-range communications can be easily achieved. In addition to regular IRS functionalities, at THz frequencies, reflectarrays can be used at the transmitter to generate and direct a THz beam excited by a close THz source [174]. IRS signal processing enables both communications and sensing [395].

**B. THz IRS Material Properties**

In addition to the THz-specific communications system considerations, several device and material properties favor THz-band IRS deployments, where the corresponding technologies might be mature before THz MIMO technology. Although THz-IRS CMOS technology [396], [397] has low power consumption, it is limited by the maximum clock speed. Furthermore, CMOS integrates easily with different cell designs but is challenged by parasitic capacitance leakage. The CMOS unit cells of subwavelength dimensions can also result in large footprints. Micro-electro-mechanical systems (MEMS) for THz surface design [398], [399] are also limited by switch movement speed and control signaling, where each switch requires independent current and the resultant power consumption is relatively high. Because they are mechanical, MEMS are also more subject to faults and error and exhibit relatively big footprints.

Conversely, graphene-based technology is low-power-consuming, easy to implement and integrated into small footprints, and has a simple biasing circuit. Nevertheless, graphene is limited by the voltage-implementing controller. Graphene-based metasurfaces can control the chemical potential of reflecting elements via electrostatic biasing, which varies the complex conductivity to achieve phase control [400]. Graphene-based digital metasurfaces combining reconfigurable and digital approaches are studied in [401], where beamsteering is achieved by dynamically adjusting a phase gradient along the metasurface plane. A graphene-based metasurface is also proposed in [402], where a two-dimensional periodic array of graphene meta-atoms guarantees a wideband perfect-absorption polarization-insensitive reconfigurable behavior at THz frequencies.

THz metasurface-based beamsteering techniques can achieve wide-angle ranges at high compactness and low weight. In [403], the proposed curvilinear THz metasurface design is argued to be independent of the geometry and the frequency. Metamaterials and metasurfaces operating in the THz band can provide flexibility for generating orbital angular momentum and polarization conversion [404]. Furthermore, the concept of HyperSurfaces is proposed in [405] for THz communications. HyperSurfaces are composed of a stack of virtual and physical components that can enforce lens effects and custom reflections per tile.

In addition to graphene, thermally or electrically tunable vanadium dioxide and liquid crystals and micro-electromechanical systems have also been considered as candidates for efficient THz steering technologies [404]. Novel intelligent plasmonic antenna array designs for transmission, reception, reflection, and waveguiding of multipath THz signals are further studied in [406], where an end-to-end physical model is developed. Nevertheless, future metasurfaces should support higher reconfigurability and sensing accuracy to support spatially-sensitive THz communications. A novel distributed control process, perhaps aided by optical internet-
working, can guarantee fast adaptation [405].

C. Signal Processing for THz IRS

Performance analysis studies for IRS-assisted communications include [394, 407]–[413]. However, the corresponding performance limits at high frequencies [174], [414]–[417], are still lacking. Most analytical studies consider lower-frequency scenarios and assume a downlink multi-user system model, in which an IRS is in the LoS of a base station assisting in reaching multiple users, each having a small number of antennas. For such setups, in [407], the minimum achievable signal-to-interference-plus-noise ratio is studied, both when the channel between the IRS and the base station is full-rank or rank-one.

The optimality of passive beamforming in IRSs is studied in [409], where a novel modulation beamform that avoids interference with existing users is proposed, alongside a resource allocation algorithm. Opportunistic scheduling is further studied in [408] as a means to achieve good multi-user diversity gains in spatially correlated LoS scenarios. Furthermore, the effect of random blockages in large-scale surface deployments is studied in [418] (blockage mitigation is crucial at THz frequencies). However, the THz multi-user channel is very sparse, and most of the analytical frameworks must be revised in the THz context. THz RIS systems can also assist virtual reality applications, as illustrated in [419].

Several attempts for channel modeling in IRS-assisted scenarios at high frequencies are noted [420]. In [421], a 3D channel model for indoor hypersurface-assisted communications at 60 GHz is developed. Channel estimation for IRS-assisted THz communications is studied along with hybrid beamforming in [416], where cooperative channel estimation is achieved via beam training and exploiting the advantages of high dimensionalities and poor scattering at THz frequencies. End-to-end 3D channel modeling of radiation patterns from graphene-based reflectarrays at true THz frequencies is also presented in [174]. For an IRS consisting of $M_i \times N_i$ reflecting elements, the IRS-assisted NLoS communications system model by assuming AoSAs being placed at the transmitter and the receiver, is a simple extension of [1] and can be expressed as

$$y = W^H_r (H_{IR} \Phi H_{TI}) W^H_r x + \hat{n},$$

where $H_{IR} \in \mathbb{C}^{M_i \times N_i \times M_i \times N_i}$ is the channel between the IRS and the receiving array, $H_{TI} \in \mathbb{C}^{M_i \times N_i \times M_i \times N_i}$ is the channel between the transmitting array and the IRS, $\hat{n}$ is the equivalent noise vector at the receiver, and $\Phi = \text{diag}(\beta_{11} e^{j \Psi_{11}}, \ldots, \beta_{M_i N_i} e^{j \Psi_{M_i N_i}}) \in \mathbb{C}^{M_i \times N_i \times M_i \times N_i}$ is the diagonal matrix that comprises the gains ($\beta's$) and phase shifts ($\Omega's$) at each IRS element.

Unlike at lower frequencies, where channel hardening effects can arise with large IRSs [411], the channel is highly correlated and low-rank at high frequencies. Hence, in addition to increasing the signal strength, IRSs operating at high frequencies should enhance the system performance by increasing the overall channel rank and suppressing interference. The researchers in [413] demonstrate how an IRS can be used to increase the channel rank, leading to substantial capacity gains.

We illustrate this issue at THz frequencies by extending the concept of spatial tuning in Sec. VI-A to IRS-assisted THz NLoS environments. In a simplified proof-of-concept binary IRS operation, we assume that each reflecting element can either fully absorb an incident signal or reflect it toward a target direction. Hence, by controlling which element to reflect, a spatial degree of freedom is added at the IRS level, which could enhance the multiplexing gain of the NLoS system, but at the expense of a reduced total reflected power. Global solutions can be derived by jointly optimizing $\Delta_1, \Delta_2,$ and $\Delta_3,$ the inter-SA spacing at the transmitting array, inter-reflect-element spacing at the intermediate IRS, and the inter-SA spacing at the receiving array, respectively, as illustrated in Fig. 3. For this special case, the channel matrices (less the molecular absorption factor) can be approximated as

$$H_{IR} \approx \frac{c}{4\pi f d_2} e^{-j \pi f d_2} \begin{pmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1, M_i N_i} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2, M_i N_i} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{M_i N_i, 1} & \omega_{M_i N_i, 2} & \cdots & \omega_{M_i N_i, M_i N_i} \end{pmatrix},$$

$$H_{TI} \approx \frac{c}{4\pi f d_1} e^{-j \pi f d_1} \begin{pmatrix} \tilde{\omega}_{1,1} & \tilde{\omega}_{1,2} & \cdots & \tilde{\omega}_{1, M_i N_i} \\ \tilde{\omega}_{2,1} & \tilde{\omega}_{2,2} & \cdots & \tilde{\omega}_{2, M_i N_i} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\omega}_{M_i N_i, 1} & \tilde{\omega}_{M_i N_i, 2} & \cdots & \tilde{\omega}_{M_i N_i, M_i N_i} \end{pmatrix},$$

where $\omega_{m_r n_r m_i n_i} = e^{-j\Psi_1 (\Delta_2, \Delta_3) / d_2}, \omega_{m_r n_r m_i n_i} = e^{-j\Psi_2 (\Delta_3, \Delta_1) / d_1}, \Psi_1$ and $\Psi_2$ are functions of the specific geometry and coordinate system, and $d_1$ and $d_2$ are the distances between the centers of the IRS and the transmitting array and the IRS and the receiving array, respectively.

Compressive sensing and machine learning techniques for IRS-assisted THz communications are promising. For instance, in [422], the training overhead for channel estimation and the baseband hardware complexity are both reduced by assuming a sparse channel sensor configuration for surfaces. In this architecture, several elements in the IRS remain active (without RF resources; standard reflecting elements cannot send pilot symbols for channel estimation), and compressive sensing is used to acquire the channel responses on all other passive elements. This knowledge can then be exploited in a deep learning-based solution to design the reflection matrices with no training overhead. For IRS systems with imperfect CSI, distributed reinforcement learning techniques are also considered for channel estimation in [423]. Deep learning is also exploited for beam and handoff prediction in drones with IRSs in [424]. Furthermore, deep reinforcement learning for IRS-based hybrid beamforming is studied in [425], [426].

Several other recent signal processing solutions for THz-IRS systems have been proposed. Sum rate maximization for IRS-assisted assisted THz communications is studied in [427], [428]. In [413], indoor IRS-assisted THz communications are studied, where a near-optimal low-complexity phase shift search scheme is proposed as an alternative to a complex, exhaustive search. Beamforming in THz scenarios incorporat-
ing both graphene-based UM-MIMO arrays and metasurfaces is also studied in [429], demonstrating the potential of combining these two technologies. IRS-based index modulation schemes [430] can also be very efficient at THz frequencies. A Taylor-expansion-aided gradient descent scheme is proposed in [431] for optimizing the desired phase shift combination of IRS elements. Cooperative beam training using hierarchical codebooks for IRS channel estimation is also proposed in [432], and channel estimation for THz-IRS-MIMO systems is studied in [433].

X. Extensions

A. THz Localization

Instead of being a byproduct of the communications system, localization in 6G is indispensable for location-aware communications. High-resolution localization capabilities [24], [434]–[442] are critical to THz communications, especially because the beams are narrow and mobile users are hard to track. Higher directionality, array compactness, and larger bandwidths are all features that can be exploited to enhance THz-based localization. LoS THz propagation conditions can significantly improve the required distance estimation for localization.

In contrast, with tiny device footprints and dense networks, THz localization becomes a prerequisite for communications. For example, channel estimation, spatial multiplexing, beamforming, and resource allocation can benefit from THz-based localization information. Furthermore, maintaining the relative positions of UEs is advantageous for efficient tracking and link re-establishment. THz localization thrives on the availability of adequate infrastructure and access to wider bandwidths, where cooperative localization can further improve the system performance. Therefore, the interaction between THz localization and communication can synergistically contribute to a versatile system that can perform multiple functions beyond data communications.

The angular and triangular accuracy in 3D localization can be enhanced by THz-band massive array signal processing, ensuring accurately beamformed THz beams in the elevation and azimuth directions of interest. Furthermore, conventional ranging techniques based on the received signal strength, time of arrival, angle of arrival, and time difference of arrival can be adjusted for THz characteristics. Nevertheless, novel localization techniques are more suitable for high-frequency operations. For instance, simultaneous localization and mapping (SLAM) can use THz-generated high-resolution environment images to improve location precision [435]. Multidimensional scaling (MDS) is another technique that can be used for THz network localization [21].

THz localization algorithms can use both IRS phase shifting and base station precoding for the delay and angle estimation. THz radar also promises to achieve millimeter accuracy. In the context of IRS systems, by acquiring the accurate position of an UE, joint precoding at the base station and phase shifting at the IRS can guarantee accurate angle or delay estimation. In [443], [444], a leaky-wave antenna with a broadband transmitter is proposed for single-shot link discovery of neighboring nodes and mobility tracking. Learning-based solutions for high-frequency localization are also promising [445]–[447]. THz sensing, imaging, and localization applications can all be piggybacked onto THz wireless communication or supported via dedicated resource allocation schemes [21].

B. THz Sensing and Imaging

The unique THz spectral fingerprints of biological and chemical materials have been exploited in many sensing and imaging applications [24], [448]–[452], such as quality control, food safety, and security. Because THz frequencies are close to the optical realm, high-energy electromagnetic waves behave as photons, often interacting with other particles and matter. Such light-matter interactions with small particles (reflections, diffraction, and absorption) create unique electromagnetic signatures that can be exploited for THz sensing.

THz signals can penetrate several materials and are strongly reflected by metals. They can also be used to analyze water dynamics (due to molecular coupling with hydrogen-bonded networks) and gas compositions (rotational spectroscopy). However, based on recent advances in THz technology and the prospect of realizing THz capabilities in hand-held devices, THz sensing applications will extend beyond the traditional industrial and pharmaceutical domains to reach everyday applications. Sensing, imaging, and localization applications will likely be piggybacked onto THz wireless communications [21].

Novel THz-specific signal processing techniques are required to enable efficient joint THz sensing and communications. The first stage of THz sensing is signal acquisition, usually achieved via THz time-domain spectroscopy (THz-TDS). THz-TDS can be achieved in reflection mode or transmission mode, with the latter being more useful for sensing and imaging from a distance. Furthermore, reflection-based spectroscopy is more convenient in the context of joint communications and sensing.

Beyond signal acquisition, several signal processing and machine learning techniques can be used to pre-process the received signals, extract characteristic features, and classify target materials into appropriate classes [453]. Furthermore, the accuracy of sensing and imaging is greatly enhanced in the THz band due to the vastly wider available channel bandwidths and the high directionality that accompanies massive MIMO beamforming. Smart metasurfaces operating in the THz band are also capable of sensing environments [454]. In [417], artificial intelligence is used in the context of IRS-assisted intercell mmWave communications for sensing, programmable computing, and actuation facilities within each unit cell.

For massive THz MIMO systems, carrier-based sensing and imaging can be more efficient as multiple RF chains can be tuned to multiple frequencies, generating multiple responses over the THz spectrum. Compared to short pulses covering the entire THz frequency range, carrier-based THz systems (frequency-domain spectroscopy) provide greater flexibility for choosing the carriers of interest for specific sensing applications. Only several carefully selected carriers can provide an efficient test for the existence of a specific molecule.
In the particular case of carrier-based THz-band wireless gas sensing (also known as electronic smelling [34]), the estimated channel responses can be correlated with the HITRAN database [141] so that a decision is made on the gaseous constituents of the medium. We demonstrate that this sensing procedure can be seamlessly piggybacked over a communications system by considering an UM-MIMO AoSA scenario and assuming each SA to be tuned to a specific frequency symmetrically at the transmitter and receiver. Multiple SAs can still be tuned to the same frequency while assuming the channel to be orthogonal by design (following spatial tuning). Hence, the corresponding channel is diagonal, and the MIMO problem can be resolved into multiple single-input-single-output problems. Each diagonal entry of $\mathbf{H}$ thus represents the channel response between a particular SA at the transmitter (tuned to a particular frequency) and its corresponding SA at the receiver side. The received vector can then be expressed as in (4), where $h_{m,n,r,m,n_t}$ is defined as

$$h_{m,n,r,m,n_t} = a_{M,N_r,N_t} e^{-j \frac{\pi}{2} \gamma G N_r N_t} G_t \alpha_{m,n,r,m,n_t} G_r \alpha_r (\theta_t, \phi_t),$$

with $\gamma G N_r N_t$ the absorption coefficient of gas $g$ at frequency $f_n$. By solving for $\mathbf{H}$, we identify both the gasses (or even specific isotopes of gases) that exist in the medium and their concentrations (by inspecting equation (3)). The larger the AoSA size is, the more observations can be accumulated per channel use, and the faster the decision is made on the constituents of the medium.

Several methods can be used to solve for the absorption coefficients in $\mathbf{H}$, including optimal maximum likelihood detectors, variations of compressed sensing techniques, and machine learning algorithms. For instance, instead of comparing the exact values of channel measurements, we can set thresholds to check the presence or absence of specific spikes and build decision trees for classification [355]. $\mathbf{x}$ can be assumed to be a fixed or a random vector for sensing. However, in joint sensing and communications setups, the entries of $\mathbf{x}$ would belong to a specific constellation with a specific structure. This knowledge can be exploited to enhance the sensing performance further.

### C. Networking and Security

Having addressed several THz-specific signal processing techniques, both the signal processing and networking problems differ in the THz band; both are linked to the underlying THz device architectures. Multiple access and networking paradigms for highly varying THz mobile environments are required. Hence, THz-specific MAC protocols [33, 34, 168, 169, 456–460] need to be optimized jointly with PHY layer signal processing schemes under the constraints of state-of-the-art THz devices.

Examples of joint optimization schemes include the studies in [461], where energy harvesting THz nanonetworks are designed for controlling software-defined metamaterials, and in [462], where joint THz power allocation and scheduling are optimized in mesh networks. Similarly, an on-demand multi-beam power allocation MAC protocol for THz MIMO networks is proposed in [458]. A receiver-initiated handshake procedure is proposed in [168] to guarantee link-layer critical synchronization between high-speed THz networks. Synchronization of ultra-broadband THz signals, spectrum access and sharing, and neighbor discovery (given narrow beams) are open problems that require solutions on both the PHY and network layers.

THz security issues at the PHY and network layers are also important [463–466]. THz communications are more secure with higher propagation losses and increased directionality than communication paradigms at lower frequencies. Nevertheless, this enhanced security is not perfect. Security and eavesdropping in THz links are first studied in [467], where it is argued that security protocols should be designed on multiple levels, including hardware and the PHY layer. Signal processing techniques for waveform design can be proposed for the latter.

A scattering object can still be placed within the broadcast sector of a transmitting antenna, despite the increased spatial resolution, which would then scatter radiation towards a nearby eavesdropper. By perfectly characterizing the backscatter of the channel, such a security breach can be avoided. Furthermore, narrow beams can still cover a relatively large area around the receiver, a vulnerability that can be exploited for eavesdropping. In [468], this vulnerability is mitigated by THz multipath propagation at the expense of slightly reduced capacity (sending shares of data over multiple paths). In other notable studies, secure-transmission IRS-assisted THz systems are studied in [469], and adding receiver artificial noise to enable THz secure communications is proposed in [470].

Covert THz communications, in which an adversary residing inside the beam sector is prevented from knowing the occurrence of transmission, are also gaining attention. For instance, in [471], covert THz communication is studied at the network level in dense internet-of-things systems using reflections and diffuse scattering from rough surfaces. Furthermore, in [472], covertness is achieved by designing novel modulation schemes, such as distance-adaptive absorption peak hopping, in which frequency hopping is strategically selected at THz molecular absorption peaks. The covert distance is dictated by the transmit power and the SNR thresholds.
XI. CONCLUSIONS

In this paper, we present a first-of-its-kind tutorial on signal processing techniques for THz communications. We detail the THz channel characteristics and summarize recent literature on THz channel modeling attempts, performance analysis frameworks, and experimental testbeds. We also highlight problem formulations that extend classical signal processing for wireless communications techniques into the THz realm. We study THz-band modulation and waveform design, beamforming and precoding, channel estimation, channel coding, and data detection, extending the discussion to the role of reflecting surfaces in the THz band and THz sensing, imaging, and localization. We also highlight THz-band networking and security issues. The techniques discussed in this paper will continue to evolve in the near future, driven by advances in THz transceiver design and system modeling.

XII. ACKNOWLEDGMENTS

We thank Prof. Josep Miquel Jornet and Dr. Onur Sahin for the fruitful discussions on the topic and Mr. Simon Tarbouch for his input on THz channel modeling.

REFERENCES

[1] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, C. I., and A. Ghosh, “Millimeter wave communications for future mobile networks,” IEEE J. Sel. Areas Commun., vol. 35, no. 9, pp. 1909–1935, Sep. 2017.
[2] S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: Potentials and challenges,” Proceedings of the IEEE, vol. 102, no. 3, pp. 366–385, Mar. 2014.
[3] I. F. Akyildiz, J. M. Jornet, and C. Han, “TeraNets: ultra-broadband communication networks in the terahertz band,” IEEE Wireless Commun., vol. 21, no. 4, pp. 130–135, Aug. 2014.
[4] ——, “Terahertz band: Next frontier for wireless communications,” Physical Communication, vol. 12, pp. 16–32, Sep. 2014.
[5] T. Kleine-Ostmann and T. Nagatsuma, “A review on terahertz communications research,” Journal of Infrared, Millimeter, and Terahertz Waves, vol. 32, no. 2, pp. 143–171, 2011.
[6] J. Chen, X. Ma, B. Zhang, Y. Zhang, Z. Niu, N. Huang, W. Chen, L. Li, and S. Li, “A survey on terahertz communications,” China Communications, vol. 16, no. 2, pp. 1–35, Feb. 2019.
[7] K. Huang and Z. Wang, “Terahertz terabit wireless communication,” IEEE Commun. Mag., vol. 12, no. 4, pp. 108–116, Jun. 2011.
[8] H. Song and T. Nagatsuma, “Present and future of terahertz communications,” IEEE Trans. THz. Sci. Technol., vol. 1, no. 1, pp. 256–263, Sep. 2011.
[9] N. Rajatheva, I. Atzeni, E. Björnson, A. Bourdoux, S. Buzzi, J.-B. Dore, S. Erkucuk, M. Fuentes, K. Guan, Y. Hu et al., “White paper on broadband connectivity in 6G,” arXiv preprint [arXiv:2004.14247] 2020.
[10] N. Rajatheva, I. Atzeni, S. Bicais, E. Björnson, A. Bourdoux, S. Buzzi, C. D’Andrea, J.-B. Dore, S. Erkucuk, M. Fuentes et al., “Scoring the terabit/s goal: Broadband connectivity in 6G,” arXiv preprint [arXiv:2008.07220] 2020.
[11] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, “What should 6G be?” Nature Electronics, vol. 3, no. 1, pp. 20–29, 2020.
[12] K. M. S. Haq, S. A. Busari, J. Rodriguez, V. Frascolla, W. Buzzi, and D. C. Sicker, “Terahertz-enabled wireless system for beyond-5G ultrafast networks: A brief survey,” IEEE Netw., vol. 33, no. 4, pp. 89–95, Jul. 2019.
[13] J. Zhang, K. Kang, Y. Huang, M. Shafi, and A. F. Molisch, “Millimeter and THz wave for 5G and beyond,” China Commun., vol. 16, no. 2, pp. ii–iii, Feb. 2019.
[14] A.-A. A. Boulogeronos, A. Alexiou, D. Kricharidis, A. Katsiotis, G. Ntouri, J. Kokkoneni, J. Lethomaki, M. Juniti, D. Yankova, A. Mokhtar et al., “Wireless terahertz system architectures for networks beyond 5G,” arXiv preprint [arXiv:1810.12260] 2018.
structure,” in Proc. IEEE Int. Conf. Commun. (ICC), May 2019, pp. 1–6.
[298] D. Zhang, Y. Wang, X. Li, and W. Xiang, “Hybridly connected structure for hybrid beamforming in mmwave massive MIMO systems,” IEEE Trans. Commun., vol. 66, no. 2, pp. 662–674, 2018.
[299] L. Yan, C. Han, N. Yang, and J. Yuan, “Dynamic-subarray with quantized- and fixed-phase shifters for terahertz hybrid beamforming,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), 2020, pp. 1–6.
[300] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, “Spatially sparse precoding in millimeter wave mimo systems,” IEEE Trans. Wireless Commun., vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[301] L. Yan, C. Han, and J. Yuan, “Joint inter-and-intra-multiplexing and hybrid beamforming for terahertz widely-spaced multi-subarray systems,” arXiv preprint arXiv:2011.07929, 2020.
[302] H. Yuan, N. Yang, K. Yang, C. Han, and J. An, “Hybrid beamforming for MIMO-OFDM terahertz wireless systems over frequency selective channels,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec. 2018, pp. 1–6.
[303] J. Tan and L. Dai, “Delay-phase precoding for THz massive MIMO with beam split,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec. 2019, pp. 1–6.
[304] M. Andrello, A. Singh, N. Thawdar, and J. M. Jornet, “Dynamic beamforming algorithms for ultra-directional terahertz communication systems based on graphene-based plasmonic nano-antenna arrays,” in Proc. Asilomar Conf. Signals, Systems and Computers (Asilomar), Oct. 2018, pp. 1558–1563.
[305] C.-Y. Chen, C.-F. Hsieh, Y.-F. Lin, R.-P. Pan, and C.-L. Pan, “Magnetically tunable room-temperature 2×2 liquid crystal terahertz phase shifter,” Optics express, vol. 12, no. 12, pp. 2625–2630, 2004.
[306] Y. Wu, X. Ruan, C.-H. Chen, Y. J. Shin, Y. Lee, J. Niu, J. Liu, Y. Chen, K.-L. Yang, X. Zhang et al., “Graphene/liquid crystal based terahertz phase shifters,” Optics express, vol. 21, no. 18, pp. 21395–21402, 2013.
[307] J. Zhang, E. Björnson, M. Matthaiou, D. K. W. Ng, H. Yang, and D. J. Love, “Prospective multiple antenna technologies for beyond 5G,” IEEE J. Sel. Areas Commun., vol. 38, no. 4, pp. 629–643, 2014.
[308] H. Yuan, N. Yang, K. Yang, C. Han, and J. An, “Enabling massive connections using hybrid beamforming in terahertz micro-scale networks,” in Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC), 2020, pp. 1–7.
[309] T. Mir, M. Waqas, U. Mir, S. M. Hussain, A. M. Elbir, and S. Tu, “Joint precoding design for two-way massive MIMO systems,” IEEE Access, vol. 8, pp. 22660–22672, 2021.
[310] C. Laperle and M. O’Sullivan, “Advances in high-speed DACs, ADCs, and DSP for optical coherent transceivers,” J. of Lightwave Tech., vol. 32, no. 4, pp. 629–643, 2014.
[311] J. Mo and R. W. Heath, “High SNR capacity of millimeter wave MIMO systems with one-bit quantization,” in Proc. Inf. Theory and Applications Workshop (ITA Workshops), Feb. 2014, pp. 1–5.
[312] P. Neuhaus, M. Dörpinghaus, H. Halbauer, S. Wessmann, M. Schlüter, F. Gast, and G. Fettweis, “Sub-THz wideband system employing 1-bit quantization and temporal oversampling.”
[313] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, “Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink,” IEEE Trans. Signal Process., vol. 65, no. 17, pp. 4624–4634, Sep. 2017.
[314] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, “Quantized precoding for massive MU-MIMO,” IEEE Trans. Commun., vol. 65, no. 11, pp. 4670–4684, 2017.
[315] Ö. Tüfge Demir and E. Björnson, “The bussgang decomposition of nonlinear systems: Basic theory and MIMO extensions [lecture notes],” IEEE Signal Process. Mag., vol. 38, no. 1, pp. 131–136, 2021.
[316] D. Li, D. Qiao, L. Zhang, and G. Y. Li, “Performance analysis of indoor THz communications with one-bit precoding,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec. 2018, pp. 1–7.
[317] G. Fettweis, M. Dörpinghaus, S. Bender, L. Landau, P. Neuhaus, and M. Schlüter, “Zero crossing modulation for communication with temporally oversampled 1-bit quantization,” in Proc. Asilomar Conf. Signals, Systems and Computers (Asilomar). IEEE, 2019, pp. 207–211.
[318] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, “A survey of non-orthogonal multiple access for 5G,” IEEE Commun. Surveys Tuts., vol. 20, no. 3, pp. 2294–2323, thirdquarter 2018.
[433] X. Ma, Z. Chen, Y. Chi, W. Chen, L. Du, and Z. Li, “Channel estimation for intelligent reflecting surface enabled terahertz mimo systems,” in IEEE Int. Conf. on Commun. Work. (ICC Work.), 2020, pp. 1–6.

X. Panah, S. P. M., T. Panigrahi, and M. Hassan, “Direction of arrival estimation for intelligent reflecting surface enabled terahertz mimo systems,” in IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP), 2019, pp. 4539–4543.

[441] O. Kanhere and T. S. Rappaport, “Position location for full-scale 6G terahertz wireless systems,” in IEEE Global Commun. Conf., (GLOBECOM), 2018, pp. 206–212.

[445] M. El-Absi, A. Alhajj Abbas, A. Abuelhajja, F. Zheng, K. Soltbach, and T. Kaiser, “High-Accuracy indoor localization based on chipless RFID systems at THz band,” IEEE Access, vol. 6, pp. 54 355–54 368, Sep. 2018.

[458] X. Yao, F. Ni, and C. Wang, “Multi-beam on-demand power allocation MAC protocol for MIMO terahertz communication networks,” in IEEE Trans. Commun., vol. 70, no. 2, pp. 70–75, 2021.

[459] O. Kianoush, S. Savazzi, and V. Rampa, “Passive detection and discrimination of body movements in the sub-thz band: A case study,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP), May 2019, pp. 1587–1591.

[460] D. L. Woolard, R. Brown, M. Pepper, and M. Kemp, “Terahertz frequency domain spectroscopy identification system based on decision trees,” Acta Physica Polonica-Series A General Physics, vol. 122, no. 5, p. 891, 2012.

[461] Z. Hassien, Q. Xia, and J. M. Jornet, “TeraSim: An ns-3 extension to simulate terahertz-band communication networks,” Nano Communication Networks, 2017, pp. 36–44, 2018.

F. Lemic, R. U. Akbar, J. Marquez-Barja, and J. Famaey, “Assessing the reliability of energy harvesting terahertz nanonetworks for controlling software-defined materials,” in Proc. of the Sixth Annual ACM Int. Conf. on Nanoscale Comput. and Commun., 2019, pp. 1–6.

[463] R. Ryniec, P. Zagrajek, and N. Palka, “Terahertz frequency domain spectroscopy identification system based on decision trees,” Acta Physica Polonica-Series A General Physics, vol. 122, no. 5, p. 891, 2012.

[464] X. Yao, F. Ni, and C. Wang, “Multi-beam on-demand power allocation MAC protocol for MIMO terahertz communication networks,” in IEEE Trans. Commun., vol. 70, no. 2, pp. 70–75, 2021.

[465] O. Kianoush, S. Savazzi, and V. Rampa, “Passive detection and discrimination of body movements in the sub-thz band: A case study,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP), May 2019, pp. 1587–1591.

[466] D. L. Woolard, R. Brown, M. Pepper, and M. Kemp, “Terahertz frequency domain spectroscopy identification system based on decision trees,” Acta Physica Polonica-Series A General Physics, vol. 122, no. 5, p. 891, 2012.

[467] Z. Hassien, Q. Xia, and J. M. Jornet, “TeraSim: An ns-3 extension to simulate terahertz-band communication networks,” Nano Communication Networks, 2017, pp. 36–44, 2018.

[468] F. Lemic, R. U. Akbar, J. Marquez-Barja, and J. Famaey, “Assessing the reliability of energy harvesting terahertz nanonetworks for controlling software-defined materials,” in Proc. of the Sixth Annual ACM Int. Conf. on Nanoscale Comput. and Commun., 2019, pp. 1–6.

[469] X. Yao, F. Ni, and C. Wang, “Multi-beam on-demand power allocation MAC protocol for MIMO terahertz communication networks,” in IEEE Trans. Commun., vol. 70, no. 2, pp. 70–75, 2021.

[470] Q. Ma, G. D. Bai, H. B. Jing, C. Yang, L. Li, and T. J. Cui, “Smart metasurface with self-adaptively reprogrammable functions,” Light: Science & Applications, vol. 8, no. 1, pp. 1–12, 2019.

[471] M. Poiese, J. M. Jornet, T. Melodia, and M. Zorzi, “Toward end-to-end, full-stack 6G terahertz networks,” IEEE Commun. Mag., vol. 58, no. 11, pp. 48–54, 2020.

[472] D. Burghal, A. T. Ravi, V. Rao, A. A. Alghafis, and A. F. Molisch, “Assessing the reliability of energy harvesting terahertz nanonetworks for controlling software-defined materials,” in Proc. of the Sixth Annual ACM Int. Conf. on Nanoscale Comput. and Commun., 2019, pp. 1–6.

[473] M. Yu, A. Tang, X. Wang, and C. Han, “Joint scheduling and power allocation for 6G terahertz mesh networks,” in 2020 Int. Conf. on Computing, Networking and Commun. (ICNC), 2020, pp. 631–635.

[474] R. Singh and D. Sicker, “THz communications—a boon and/or bane for security, privacy, and national security,” Privacy, and National Security (December 16, 2020).

[475] R. Wang, Y. Mei, X. Meng, and J. Ma, “Secrecy performance of terahertz wireless links in rain and snow,” Nano Communication Networks, vol. 28, p. 100350, 2021.

[476] R. Wang, Y. Mei, X. Meng, and J. Ma, “Secrecy performance of terahertz wireless links in rain and snow,” Nano Communication Networks, vol. 28, p. 100350, 2021.

[477] J. Ma, R. Shrestha, J. Adelberg, C.-Y. Yeh, Z. Hassien, E. Knightly, J. M. Jornet, and D. M. Mittelman, “Secrecy and eavesdropping in terahertz wireless links,” Nature, vol. 563, no. 7729, p. 89, 2018.

[478] V. Petrov, D. Moltchanov, J. M. Jornet, and Y. Koucheryavy, “Exploiting multipath terahertz communications for physical layer security in beyond 5G networks,” in IEEE INFOCOM 2019 - IEEE Conf. on Comp. Commun. Workshops (INFOCOM WKSHPS), Apr. 2019, pp. 865–872.

[479] J. Qiao and M. S. Alouini, “Secure transmission for intelligent reflecting surface-assisted mmWave and terahertz systems,” IEEE Wireless Commun. Lett., vol. 9, no. 5, pp. 1743–1747, 2020.

[480] W. Gao, C. Han, and Z. Chen, “Receive artificial noise aided terahertz communications with eavesdropper in close proximity,” in GLOBECOM, 2020, pp. 1–6.

[481] Z. Liu, J. Liu, Y. Zeng, and J. Ma, “Covert wireless communication in IoT network: From AWGN channel to THz band,” IEEE Internet Things J., pp. 1–1, 2020.

[482] W. Gao, Y. Chen, C. Han, and Z. Chen, “Distance-adaptive absorption-peak hopping (DA-APH) modulation for terahertz covert communications,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec. 2019, pp. 1–6.

[483] R. Ryne, P. Zagrzejek, and N. Palka, “Terahertz frequency domain spectroscopy identification system based on decision trees,” Acta Physica Polonica-Series A General Physics, vol. 122, no. 5, p. 891, 2012.
in 2009. His current research interests include the modeling, design, and performance analysis of wireless communications systems.

**Tareq Y. Al-Naffouri** (M’10-SM’18) Tareq Al-Naffouri received his B.S. degrees in Mathematics and Electrical Engineering (with first honors) from KFUPM, Saudi Arabia, his M.S. degree in Electrical Engineering from Georgia Tech, Atlanta, in 1998, and his Ph.D. degree in Electrical Engineering from Stanford University, Stanford, CA, in 2004. He is currently a Professor at the Electrical Engineering Department, KAUST. His research interests lie in the areas of sparse, adaptive, and statistical signal processing and their applications, localization, machine learning, and network information theory.