Vision-Based Reconfigurable Intelligent Surface Beam Tracking for mmWave Communications

Juan Sanchez, Xuesong Cai, and Fredrik Tufvesson

Department of Electrical and Information Technology, *Lund University*, Lund, Sweden

{juan.sanchez, xuesong.cai, fredrik.tufvesson}@eit.lth.se

Abstract—Reconfigurable intelligent surfaces have emerged as a technology with the potential to enhance wireless communication performance for 5G and beyond. However, the technology comes with challenges in areas such as complexity, power consumption, and cost. This paper demonstrates a computer vision-based reconfigurable intelligent surface beamforming algorithm that addresses complexity and cost issues and analyzes the multipath components that arise from the insertion of such a device into the wireless channel. The results show that a reconfigurable intelligent surface can provide an additional multipath component. The power of this additional path can be critical in blockage scenarios, and a capacity increase can be perceived in both line-of-sight and non line-of-sight scenarios.

Index Terms—Reconfigurable intelligent surfaces, computer vision, wireless channel, beamforming, multipath component.

I. INTRODUCTION

The ever-increasing need for higher data rates and the scarcity of the radio spectrum have been influential factors for the exploration of communication at higher frequencies with each new technology release. Propagation at radio bands such as millimeter wave (mmWave) or subterahertz starts to show characteristics typical of optical propagation, where line-of-sight (LOS) and specular reflections are the dominant propagation mechanisms [1], [2]. Propagation at these frequencies also suffers from higher propagation loss and higher absorption effects. Massive multiple-input multiple-output (MIMO) systems have exploited a large number of antenna elements to enable coherent processing of electromagnetic waves (EM) and withstand the harsh wireless channel at higher frequencies at the transmitting and receiving ends.

Recently, the paradigm has further evolved into a joint manipulation of the EM waves from the transmitting and receiving ends and from the environment itself, by using so-called reconfigurable intelligent surfaces (RISs). These RISs are devices that are thought to scatter incident EM waves at will depending on the desired configuration. The technology comes with practical challenges such as manufacturing cost, power consumption, processing complexity – which in turn affects delay and power consumption, among all. Multiple studies have shown that RISs can leverage the geometry of
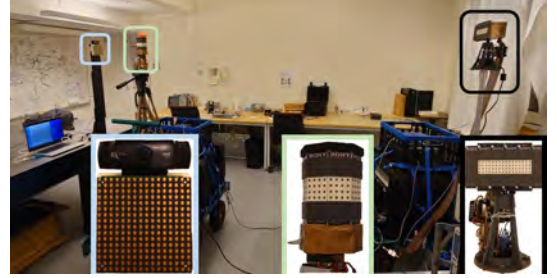
Fig. 1: mmWave setup and components. TX array (left), RX Array (center) and RIS with camera on top (right).

nodes and the environment to form a beam at low complexity. For this purpose, low-cost optical devices such as cameras can be used as sensors whose output can be processed using computer vision algorithms, revealing the geometry of the environment. The authors in [3], [4] demonstrated that the geometric information extracted from cameras can improve the SNR and consequently the achievable bitrate by using different processing methods.

The present paper demonstrates an arrangement of a computer vision-aided RIS that tracks the RX and TX location and updates its scattering pattern according to the beamforming algorithm presented in [5]. Furthermore, the effect of introducing such a structure into the wireless channel is analyzed. The rest of the paper is organized as follows: Section II presents the RIS arrangement and measurement setup. Section III establishes the post-processing and measurement analysis framework. Section IV presents the measurement results for the double-directional channel impulse response with and without RIS activation and parameterizes the wireless channels experienced in the measurements for each scenario. Section V gives conclusive remarks and outlook for future studies.

II. MEASUREMENT CAMPAIGN

For the measurement campaign, the mmWave channel sounder available at Lund University [6] was used together with a commercial RIS [7], in an indoor scenario. The channel sounder measures the channel transfer functions (CTFs) of all antenna combinations available in a so-called snapshot. The channel sounder RF ends consisted of a rectangular TX array and an octagonal RX array, both dual polarized. There were 128 TX and 256 RX antenna elements, for a total of 32768

TABLE I: Parameter setup.

Carrier frequency	28 GHz
Signal bandwidth	768 MHz
Delay resolution	1.3021 ns (39.04 cm)
Switching time	18.8 μs
Snapshot time	600 ms
Trajectory length	1 m
Total number of snapshots	128

channels measured per snapshot. The RIS consisted of a 1-bit surface arranged in 16x16 antenna patches, which can be reconfigured via serial communication with another device. A form of geometry-based maximum ratio transmission (MRT) beamforming [5] was used to control the device's phase shifts upon visual input from a commercial, low-cost optical camera. Fig. 1 shows the measurement setup and close-ups of the mmWave components used for this measurement campaign, while Table I summarizes all relevant system parameters.

The RX was moved along a 1 m straight trajectory normal to the RIS boresight at a constant speed of 1 cm/s, in a line-of-sight (LOS) scenario, as seen in Fig. 2. The trajectory included a number of static measurements at the beginning and end of the trajectory, meaning that the experienced propagation channel should not vary significantly for these positions. This observation implies that the number of available positions is larger than the expected 100 positions for such a trajectory and speed, and that these static measurements are used to validate the consistency in the measurement campaign. The same trajectory was measured for subscenarios where the RIS was active and inactive. The shadowed region in the top left corner of Fig. 2 corresponds to a virtual blocker added after in the middle of the processing to generate a synthetic non line-of-sight (NLOS) scenario, as discussed in Section III. Calibration was performed in the position axis to align the trajectories for an active and an inactive RIS after measurements were taken.

III. MULTIPATH COMPONENT EXTRACTION AND CAPACITY ANALYSIS

The propagation channel can be characterized by a transfer function. The channel transfer function, in turn, is considered as a superposition of multipath components (MPCs) modeled as follows,

$$
\begin{aligned}
&\mathbf{H}(f, s, m_{\mathrm{T}}, m_{\mathrm{R}}; \boldsymbol{\Theta}) \\
&= \sum_{l=1}^{L} \mathbf{b}_{m_{\mathrm{R}}}^{T}(\phi_{\mathrm{R},l}, \varphi_{\mathrm{R},l}, f) \begin{bmatrix} \gamma_{\mathrm{HH},l} & \gamma_{\mathrm{HV},l} \\ \gamma_{\mathrm{VH},l} & \gamma_{\mathrm{VV},l} \end{bmatrix} \mathbf{b}_{m_{\mathrm{T}}}^{T}(\phi_{\mathrm{T},l}, \varphi_{\mathrm{T},l}, f) \\
&\quad \cdot \mathbf{b}(f) e^{-j2\pi f \tau_l} e^{-j2\pi f \nu_l t_{s,m_{\mathrm{T}},m_{\mathrm{R}}}} + \mathbf{N}(f, s, m_{\mathrm{T}}, m_{\mathrm{R}}),
\end{aligned}
\tag{1}
$$

where $\mathbf{H}$ is the CTF dependent on frequency f, snapshot index s, transmit/receive antenna indices m_{T}, m_{R}, and structural parameters of the propagation channel $\boldsymbol{\Theta}$ such as directions ϕ and φ, delay τ, Doppler ν and polarimetric coefficients $\boldsymbol{\gamma}$. Furthermore, $\mathbf{b}_{m_{\mathrm{T}}}$, $\mathbf{b}_{m_{\mathrm{R}}}$ are mappings from departure / arrival directions to polarimetric antenna responses, for transmit and receive antenna elements m_{T}, m_{R}, respectively. In turn, $\mathbf{b}(f)$ is the mapping from frequency to polarimetric responses of
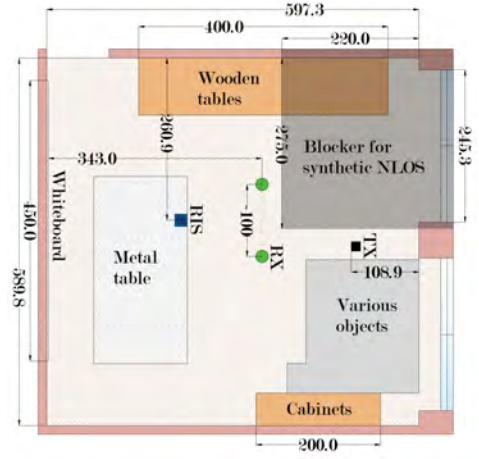

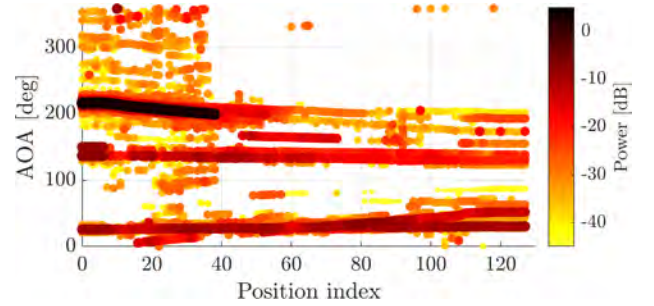Fig. 2: Measurement layout (measures in cm).

Fig. 3: Estimated AOA for the MPCs with active RIS, LOS region (before position 40) and NLOS region (after position 40).

the antenna elements. $\mathbf{N}$ denotes white Gaussian noise. The CTF can also be seen as the Fourier transform of the channel impulse response.

An implementation of the space-alternating generalized expectation maximization algorithm (SAGE) [8], [9] was used to detect the MPCs present at each measured position and estimate their parameters. Using the double-directional polarimetric capabilities of the channel sounder [6] and the knowledge of the effective aperture distribution function (EADF) of the sounder antenna arrays [10], it is possible to extract a very accurate representation of the characteristics of the propagation channel that is independent of the antenna architectures.

An additional synthetic NLOS region was generated after the estimation of the MPCs, for the assessment of the potential of the RIS technology under different conditions. The synthetic scenario considered an additional blocker shadowing the regions of the measurement layout shown in Fig. 2. The insertion of blocking was virtually performed by removing the MPCs with associated angular departure or arrival values that would impinge on the hypothetical surface. This means that the virtual blocker was assumed to be a perfect absorber. The synthetic scenario relied on equation (1) – with a reduced number of MPCs – for the reconstruction of the CTF. Fig. 3 shows the stark contrast in the transition from a LOS to a NLOS scenario, which is an approximation of the smoother

transition expected in a real scenario. The transition from LOS to NLOS is completed at position 40, implying that the NLOS scenario was bounded from position 40 and onward. Qualitative analysis of the estimated MPCs, as well as quantitative analysis of power and capacity values, were performed to assess the RIS capabilities in the LOS and NLOS scenarios. Since the RIS is, among other things, composed of a metal ground plane, a comparison between an active and an inactive RIS is equivalent to a comparison between a RIS and a metal plate located in the same position.

The effect of introducing a RIS in the environment can be characterized in power by the ratio

$$\Delta_P = \frac{||\mathbf{H}_{\mathrm{on}}||^2}{||\mathbf{H}_{\mathrm{off}}||^2} = \sum_{l=0}^{L} \frac{||\boldsymbol{\gamma}_{l,\mathrm{on}}||_F^2}{||\boldsymbol{\gamma}_{l,\mathrm{off}}||_F^2}, \quad (2)$$

where Δ_P is the ratio between the experienced channel gain for a RIS-enabled setup with respect to a scenario with a metal plate in place, $\mathbf{H}$ is the CTF from (1), and $||\boldsymbol{\gamma}_l||_F$ is the Frobenius norm of the polarimetric matrix for the lth channel path. Since the RIS used in this study was found to have polarization-dependent performance variations, we only considered the vertical to vertical polarization pair as a propagation mechanism and left the rest of the polarization pairs for further study. Therefore, (2) can be simplified to

$$\Delta_P = \sum_{l=0}^{L} \frac{|\gamma_{\mathrm{VV},l,\mathrm{on}}|^2}{|\gamma_{\mathrm{VV},l,\mathrm{off}}|^2}, \quad (3)$$

where $|\gamma_{\mathrm{VV},l,\mathrm{on}}|^2$ is the squared magnitude of the vertical-vertical polarization coefficient of the channel for the lth path. Δ_P was calculated for each position across the RX trajectory, and a moving average filter of 5 position samples was implemented on Δ_P to filter out noise-like fast variations and better visualize the tendency of the channel gain ratio.

For the calculation of capacities, it was assumed that channel state information was known at the transmitter (CSIT). Studies in time [11]–[13] and literature [14] have shown that optimal power allocation and consequent capacity maximization can be achieved by using the water filling algorithm. The achieved capacity per bandwidth unit can be computed as

$$C = \sum_{k=1}^{R_{\mathrm{H}}} \log_2 \left(1 + \frac{P_k}{\sigma_n^2} \sigma_k^2 \right), \quad (4)$$

where R_{H} is the rank of the CTF $\mathbf{H}$, σ_n^2 is the noise power density, σ_k^2 is the kth singular value of $\mathbf{H}$, and P_k is the power allocated to the kth eigenmode of the channel satisfying the constraint $\sum P_k = P$, with P being the total power allocated to a single subcarrier across all channel eigenmodes. The noise spectral density σ_n^2 was normalized to 0dB/Hz. The total power P was defined to be the same across all frequencies, and dictated by signal-to-noise ratio (SNR) values ranging from -10 dB to 30 dB. Capacities were calculated for every frequency bin within the sounder's bandwidth, every SNR value, and every trajectory position, and mean and standard deviation statistics were extracted across frequency. In addition, the ratio between the mean capacities for active and inactive RIS was

(a) Active RIS.
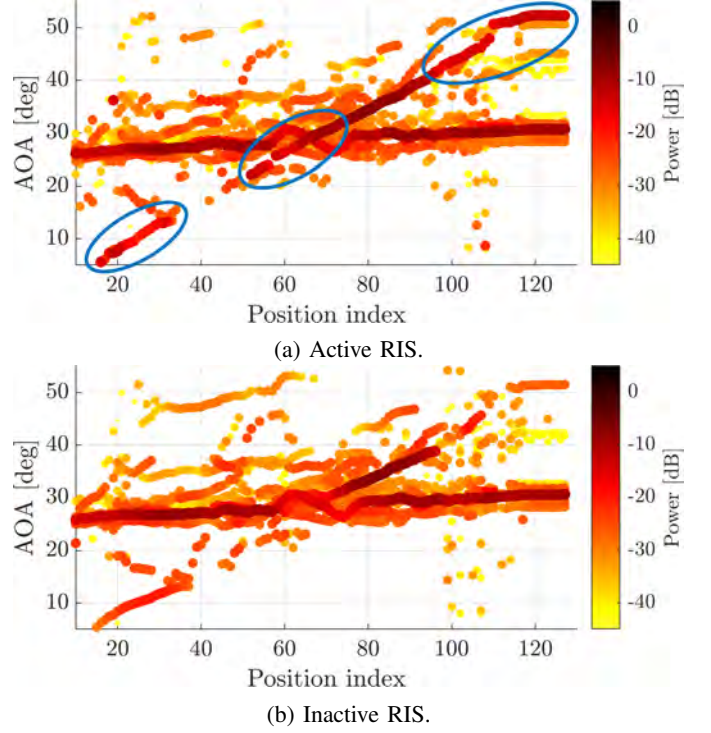

(b) Inactive RIS.

Fig. 4: Close-up of the RIS effect in the AOA range, with enhancements of MPCs being marked in (a).

calculated for each position and the SNR value. Finally, for both active and inactive RIS, the mean capacities are further averaged across the positions belonging to the LOS or NLOS scenarios and plotted against the SNR evaluation range.

IV. RESULTS AND DISCUSSION

A. Multipath Component Analysis

Fig. 3 shows the estimated received power and azimuth of arrival (AOA) values of each detected MPC at all positions. The behavior of the entire surface plot remains fairly constant for some few first and last measurement instances, validating the consistency of the measurement results for stationary positions at the beginning and end of the trajectory. It is clear that the strongest MPC that starts around 220° corresponds to the LOS path between the TX and the RX. This path decreases in angle along with the movement of the RX across the measurement trajectory, and disappears when transitioning from the LOS to the NLOS region. The MPC cluster correlated in azimuth with the LOS path also gradually disappears. As expected for the LOS region, rich scattering is observed across the whole angular domain, with numerous MPCs coming from various elevation angles when further evaluated in the rest of the directional domains. The MPC cluster with AOA values just below 150° corresponds to reflections from the cabinets in Fig. 2. The evolution of these MPCs shows power-level variations correlated with the scattering characteristics of the different components in the cabinets, mainly wood and glass. Since these reflections are coming from a direction close in angle with the direction of movement of the RX, their

AOA remains relatively constant with respect to other MPCs observed in the measurement. Finally, there is a second most dominant MPC cluster that approximately varies from 0° to 60°. This cluster corresponds to reflections coming from both the laboratory whiteboard and the RIS. As this cluster is of special importance for the purpose of this paper, let us take a closer look at the evolution of these MPCs, and foremost, at the comparison between the active and inactive RIS scenarios.

Fig. 4 shows a close-up of the MPC estimates for the AOA range of interest, for both active and inactive RIS scenarios. In both cases, we observe an MPC cluster with AOA values approximately ranging from 25° to 30°, corresponding to reflections from the whiteboard. As the whiteboard is made up of metal, the reflections coming from this cluster have a significant contribution at the power level. Moreover, the reflections from the whiteboard are not affected by the blocker introduced into the layout and remain independent of the visibility region in which the RX is located. The same can be said of the contribution from the RIS MPC, seen in Fig. 4a as the relatively straight line ranging from 5° to 52° in AOA and "crossing" the MPC cluster coming from the whiteboard, and as the rather shorter/interrupted line in Fig. 4b mainly with power contributions from 32° to 40° in AOA. As the RIS in its inactive state behaves as a metal surface, reflection occurs as a natural process for the latter AOA interval. Even though the area of the RIS is much smaller than the area where reflections from the whiteboard occur, the RIS is closer to the RX and the wavefront travels a shorter distance until it reaches the RX. This results in less spreading and lower pathloss, which in turn implies that the received power is higher than for longer path distances. The AOA values for the whiteboard's MPC cluster increase along with the trajectory except for some noticeable variations seen approximately between positions 60 and 80. These variations can be explained by the geometry of the measurement trajectory: At a certain position in the trajectory, the MPCs coming from the whiteboard and the RIS align, but the RIS is closer to the RX, hence it blocks the radiation coming from the whiteboard. However, diffraction happens around the edges of the RIS, which changes the AOA with which the RX receives some of the whiteboard's MPCs with respect to the expected AOA for no obstruction from the RIS. Hence, the oscillation around the straight linear evolution of the AOA and the reduction in received power for these MPCs is, at a level where the RIS MPC becomes the most dominant among all MPCs in the environment.

Looking further at the RIS MPC, it is very clear that the MPC's angular range for the active RIS is greatly extended, to around 400% of the angular range of a metal surface of the same dimensions. The straight evolution line of the MPC is interrupted between 15° and 25° in AOA. This artifact is a manifestation of the blockage that the RX performs when crossing the path between the TX and the RIS on its trajectory.

Let us now look at how the power relations between MPCs in the different scenarios across the visibility regions. Fig. 5 shows the received power ratio between all MPCs for the active RIS with respect to the inactive RIS. There is a
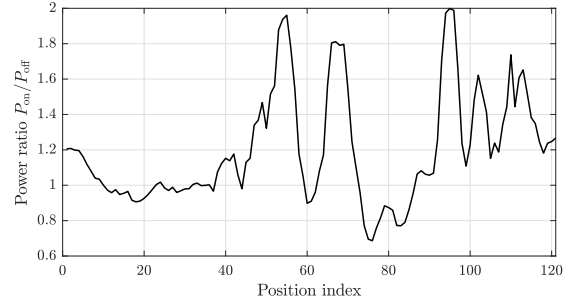

Fig. 5: Received power ratio between active and inactive RIS scenarios.

slight improvement of the power ratio at the beginning of the trajectory followed by a rather unitary relation for the remaining positions belonging to the LOS region. This relation shows that the power enhancement that the RIS brings in a small indoor LOS environment is small, but noticeable in this setup. In contrast, the improvement induced in the NLOS region of the measurements is much higher and stays for the major part of the trajectory over 1, with most of the peaks above 1.4 and going up to 2 in the best case. The overall improvement corresponds to logarithmic gains of up to 0.8 dB for the LOS region and up to 3 dB for the NLOS region in our setup. Notice that for the trajectory around position 80 we see a drop in the ratio, going as low as 0.7. Two factors are involved in this result. The first factor is the natural reflection of the RIS that occurs in this position range for the scenario with inactive RIS. The only gain from an active RIS would be a focusing effect, which is not as pronounced because the beamforming algorithm only took an approximate value for the distance between the RX and the RIS for the whole trajectory. Clearly, for such a trajectory, the distance between the RX and the RIS changes, and beamforming loses focus. The use of multiple optical camera lenses is expected to increase the beamforming gain by accurately estimating the distance to the RX. The second factor is the absence of the strongest MPC that comes from the reflection of the whiteboard at position 82 in Fig. 4a, which reduces the entire power contribution for the active RIS scenario. After a closer inspection, some disturbances to the smooth evolution of the mentioned MPC can be seen around the area, which happens to be very close to the area where the whiteboard reflection impinges on the RIS. Further studies would be needed to assess whether the activation of the RIS is playing a role in the disturbances observed for the whiteboard's MPC cluster.

B. Capacity Analysis

For the capacity analysis, a normalization of the CTFs across frequency bins and antenna indices was carried out, thus removing the additional gain from the active RIS seen in Fig. 5 and allowing an analysis of the small-scale effects of the RIS. Fig. 6 shows the mean capacity across frequency points, for different positions and SNR values. For low SNR values the ratio remains relatively unitary, but as the SNR increases, there are more pronounced variations in the regions where the
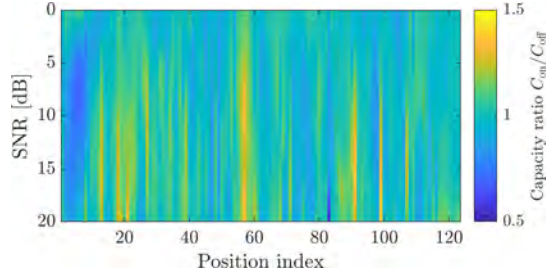
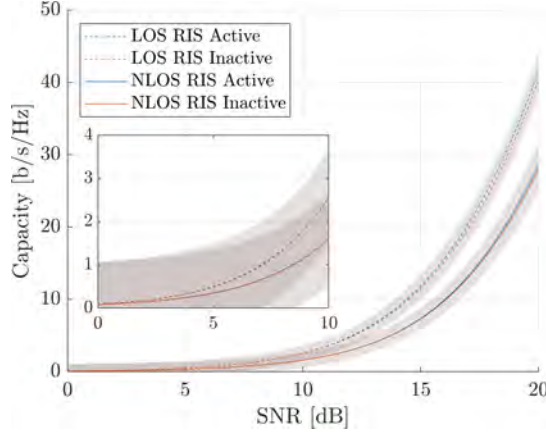Fig. 6: Ratio between capacities with active and inactive RIS.

Fig. 7: Mean capacity across all positions for all scenarios. The shadowed regions represent the mean standard deviation of the capacity curves.

ratio is greater or less than 1. There is correspondence between the increases in mean capacity around positions 20, 60, 90 and 120, and the extension of the RIS MPC when going from an inactive to an active state in Fig. 4.

Furthermore, averaging the capacities on all the positions belonging to a certain visibility region, Fig. 7 shows the mean capacity between frequencies and positions for different SNR values, visibility regions, and active/inactive RIS scenarios. The overall gap between mean capacities for the LOS and NLOS regions is very clear, whereas the gap between mean capacities of active and inactive RIS scenarios is less evident, but still noticeable. An evaluation of the relative ratio between capacities shows that there are mean relative capacity increases of 0.29% and 0.00% and maximum relative capacity increases of 3.21% and 2.20% for the active RIS with respect to the inactive RIS across the entire SNR range, for the LOS and NLOS regions, respectively. This indicates that although the RIS influences a greater overall power increase in the NLOS scenario, its use to slightly enhance capacity in mmWave communications has a similar effect for both LOS and NLOS scenarios.

V. Conclusions

This paper showed a detailed analysis of the detected multipath components (MPCs) of a measured indoor scenario with a virtual blocker for the line-of-sight (LOS) and non-LOS (NLOS) regions, where a reconfigurable intelligent surface (RIS) performing beam tracking sought to enhance communi-

cation at mmWave frequencies. Beam tracking was performed under a low-complexity geometry-based beamforming algorithm that exploits the angular and range information obtained by visual sources – a low-cost optical camera in this case. The measurement campaign considered positions across a straight trajectory where two scenarios were depicted, namely, one with an active RIS and the other with an inactive RIS, equivalent to placing a metal surface of the same dimensions. The results show that the activation of the RIS induces a significant extension of the visibility of the associated MPC over the AOA range, to the point where the RIS component becomes the dominant MPC for propagation for some positions in the trajectory. Furthermore, measurements show that the inclusion of a RIS could increase the overall received power and thus the channel propagation gain by up to 3 dB. This indicates that the implementation of RIS technology can positively influence the propagation environment, and that this can be done at low complexity and low cost.

References

[1] K. Haneda, J. Järveläinen, A. Karttunen, M. Kyrö, and J. Putkonen, "A statistical spatio-temporal radio channel model for large indoor environments at 60 and 70 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 6, pp. 2694–2704, 2015.

[2] W. Fan, I. Carton, J. Ø. Nielsen, K. Olesen, and G. F. Pedersen, "Measured wideband characteristics of indoor channels at centimetric and millimetric bands," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, pp. 1–13, 2016.

[3] S. Jiang, A. Hindy, and A. Alkhateeb, "Sensing aided reconfigurable intelligent surfaces for 3GPP 5G transparent operation," *IEEE Transactions on Communications*, pp. 1–1, 2023.

[4] M. Ouyang, F. Gao, Y. Wang, S. Zhang, P. Li, and J. Ren, "Computer vision-aided reconfigurable intelligent surface-based beam tracking: Prototyping and experimental results," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.

[5] J. Sanchez, E. Bengtsson, F. Rusek, J. Flordelis, K. Zhao, and F. Tufvesson, "Optimal, low-complexity beamforming for discrete phase reconfigurable intelligent surfaces," in *IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 01–06.

[6] X. Cai, E. L. Bengtsson, O. Edfors, and F. Tufvesson, "A switched array sounder for dynamic millimeter-wave channel characterization: Design, implementation and measurements," *IEEE Transactions on Antennas and Propagation*, 2023.

[7] L. Yezhen, R. Yongli, Y. Fan, X. Shenheng, and Z. Jiannian, "A novel 28 GHz phased array antenna for 5G mobile communications," *ZTE Communications*, vol. 18, no. 3, pp. 20–25, 2020.

[8] X. Cai *et al.*, "Enabling complexity-efficient high-resolution parameter estimation for wideband switched array channel sounding," *Under preparation*, 2023.

[9] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on selected areas in communications*, vol. 17, no. 3, pp. 434–450, 1999.

[10] X. Cai, M. Zhu, A. Fedorov, and F. Tufvesson, "Enhanced effective aperture distribution function for characterizing large-scale antenna arrays," *IEEE Transactions on Antennas and Propagation*, Jun. 2023.

[11] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2007–2019, 1999.

[12] M. Vu and A. Paulraj, "MIMO wireless linear precoding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 86–105, 2007.

[13] L. You, J. Xiong, A. Zappone, W. Wang, and X. Gao, "Spectral efficiency and energy efficiency tradeoff in massive MIMO downlink transmission with statistical CSIT," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2645–2659, 2020.

[14] A. F. Molisch, *Wireless communications*. John Wiley & Sons, 2012, vol. 34.