1-Bit MIMO for Terahertz Channels

Angel Lozano
Universitat Pompeu Fabra (UPF), Barcelona

***Abstract*—This paper tackles the problem of multiple-input multiple-output communication with 1-bit digital-to-analog and analog-to-digital converters. With the information-theoretic capacity as benchmark, the complementary strategies of beamforming and equiprobable signaling are contrasted in the regimes of operational interest. Line-of-sight settings under both spherical and planar wavefronts are considered, respectively representative of short and long transmission ranges at mmWave and terahertz frequencies. A judicious combination of beamforming and equiprobable signaling is shown to operate within a modest gap from capacity.**

I. Introduction

The next frontier in the quest for fresh spectrum over which to communicate wirelessly is the terahertz band, broadly defined as 100 GHz–10 THz. Although there are reasons why this band remains largely unexplored, the obstacles in terms of signal generation and detection look increasingly surmountable [1]. Because of the lack of diffraction, multipath propagation is limited and the communication is predominantly short-range, but that is compatible with a host of emerging applications.

A major challenge to ultrabroadband communication at terahertz frequencies is the power consumption associated with high-resolution digital-to-analog (DAC) and analog-to-digital (ADC) conversion. Precisely:

- The high-resolution DACs that go hand in hand with rich constellations force the transmit power amplifiers into exceedingly linear regimes where their efficiency is poor.
- The ADC power consumption at the receiver grows linearly with the bandwidth and exponentially with the resolution in bits [2]. Multiple watts might be required for 10-12 bits of resolution over bandwidths beyond 1 GHz.

Since both these aspects stem directly from the resolution of the converters, the natural roundabout is to lower such resolution and resort to simpler constellations. Taken to the limit, this leads to 1-bit DACs and ADCs, which drastically curb the power consumption and further enable dispensing with automatic gain control at the receiver.

The downside of 1-bit transceivers is a highly nonlinear behavior that severely distorts the signals. There is extensive literature on transmission strategies and performance with 1-bit DACs or ADCs (see [3]–[8] and references therein), and a smaller but growing body of work that considers 1-bit converters at both ends [9]–[18].

Because of the extremely high omnidirectional pathloss at terahertz frequencies, antenna arrays are instrumental, yet their usage with 1-bit converters is markedly different than with full-resolution converters. With the 1-bit multiple-input multiple-output (MIMO) capacity as benchmark, this paper provide characterizations of the performance of beamforming and equiprobable signaling, two transmission strategies that are information-theoretically motivated and complementary. It is shown how a judicious combination of these strategies suffices to operate within a modest gap from the 1-bit capacity in two channels of high relevance, side-stepping the need for general precoding solutions. Additional results reinforcing this observation in other channel types, as well as the proofs of the technical results, can be found in [19].

II. Signal and Channel Models

A. Signal Model

Consider a transmitter with N_{t} antennas and 1-bit DACs per complex dimension. The receiver, with N_{r} antennas and a 1-bit ADC per complex dimension, observes

$$\boldsymbol{y} = \mathrm{sgn}\left(\sqrt{\frac{\mathsf{SNR}}{2N_{\mathrm{t}}}}\,\boldsymbol{H}\boldsymbol{x} + \boldsymbol{z}\right) \tag{1}$$

where the sign function is applied separately to the real and imaginary parts of each entry, such that $y_n \in \{\pm 1 \pm \mathrm{j}\}$, while $\boldsymbol{H}$ is the $N_{\mathrm{r}} \times N_{\mathrm{t}}$ channel matrix, $\boldsymbol{z} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I})$ is the noise, and SNR is the signal-to-noise ratio per receive antenna. Each entry of $\boldsymbol{x}$ also takes the values $\pm 1 \pm \mathrm{j}$.

For a certain $\boldsymbol{H}$, (1) embodies a discrete memoryless channel with $4^{N_{\mathrm{t}}} \times 4^{N_{\mathrm{r}}}$ transition probabilities given by

$$p_{\boldsymbol{y}|\boldsymbol{x}} = \prod_{n=0}^{N_{\mathrm{r}}-1} p_{\Re\{y_n\}|\boldsymbol{x}}\, p_{\Im\{y_n\}|\boldsymbol{x}}, \tag{2}$$

where the factorization follows from the noise independence per receive antenna and complex dimension. Each such noise component has variance $1/2$, hence

$$p_{\Re\{y_n\}|\boldsymbol{x}}(1|\mathsf{x}) = \Pr\left[\sqrt{\frac{\mathsf{SNR}}{2N_{\mathrm{t}}}}\Re\{\boldsymbol{h}_n\mathsf{x} + z_n\} > 0\right] \tag{3}$$

$$= Q\left(-\sqrt{\frac{\mathsf{SNR}}{N_{\mathrm{t}}}}\Re\{\boldsymbol{h}_n\mathsf{x}\}\right) \tag{4}$$

where $\boldsymbol{h}_n$ is the nth row of $\boldsymbol{H}$ and $Q(\cdot)$ is the Gaussian Q-function. Similarly,

$$p_{\Re\{y_n\}|\boldsymbol{x}}(-1|\mathsf{x}) = Q\left(\sqrt{\frac{\mathsf{SNR}}{N_{\mathrm{t}}}}\Re\{\boldsymbol{h}_n\mathsf{x}\}\right). \tag{5}$$

From (4) and (5),

$$p_{\Re\{y_n\}|\boldsymbol{x}}(\Re\{\mathsf{y}_n\}|\mathsf{x}) = Q\left(-\Re\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_{\mathrm{t}}}}\Re\{\boldsymbol{h}_n\mathsf{x}\}\right) \tag{6}$$

and, mirroring it, finally

$$p_{\boldsymbol{y}|\boldsymbol{x}}(\mathbf{y}|\mathbf{x}) = \prod_{n=0}^{N_\mathrm{r}-1} Q\left(-\Re\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Re\{\boldsymbol{h}_n\mathbf{x}\}\right) \cdot Q\left(-\Im\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Im\{\boldsymbol{h}_n\mathbf{x}\}\right). \quad (7)$$

The transition probabilities correspond to (7) evaluated for the 4^{N_r} possible values of $\boldsymbol{y}$ and the 4^{N_t} values of $\boldsymbol{x}$.

The 4^{N_t} transmit vectors $\boldsymbol{x}$ can be partitioned into $4^{N_\mathrm{t}-1}$ quartets, each containing four vectors and being invariant under a 90° phase rotation of all the entries: from any vector in the quartet, the other three are obtained by repeatedly multiplying by j. Since a 90° phase rotation of $\boldsymbol{x}$ propagates as a 90° phase rotation of $\boldsymbol{H}\boldsymbol{x}$, and the added noise is circularly symmetric, the four vectors making up each transmit quartet are statistically equivalent and they should thus have the same transmission probability so as to convey the maximum amount of information.

Likewise, the set of 4^{N_r} possible vectors $\boldsymbol{y}$ can be partitioned into $4^{N_\mathrm{r}-1}$ quartets, and the four vectors $\boldsymbol{y}$ within each received quartet are equiprobable.

B. Channel Model

If the channel is stable over each codeword, then every realization of $\boldsymbol{H}$ has operational significance, and SNR is well-defined under the normalization $\mathrm{tr}(\boldsymbol{H}\boldsymbol{H}^*) = N_\mathrm{t}N_\mathrm{r}$. Two such channel types are specifically considered.

a) Line-of-Sight (LOS) with Spherical Wavefronts: LOS is the chief propagation mechanism at mmWave and terahertz frequencies, and the spherical nature of the wavefronts is relevant for large arrays and short transmission ranges. For uniform linear arrays (ULAs) [20],

$$\boldsymbol{H} = \boldsymbol{D}_\mathrm{rx}\tilde{\boldsymbol{H}}\boldsymbol{D}_\mathrm{tx} \quad (8)$$

where $\boldsymbol{D}_\mathrm{rx}$ and $\boldsymbol{D}_\mathrm{rx}$ are diagonal matrices with entries

$$[\boldsymbol{D}_\mathrm{rx}]_{n,n} = e^{-j\pi\left[\frac{2n}{\lambda}d_\mathrm{r}\sin\theta_\mathrm{r}\cos\phi + \frac{n^2}{\lambda D}d_\mathrm{r}^2\left(1-\sin^2\theta_\mathrm{r}\cos^2\phi\right)\right]}$$
$$[\boldsymbol{D}_\mathrm{tx}]_{m,m} = e^{-j\pi\left[\frac{2m}{\lambda}d_\mathrm{t}\sin\theta_\mathrm{t} + \frac{m^2}{\lambda D}d_\mathrm{t}^2\right]} \quad (9)$$

and with D the range, λ the wavelength, d_t and d_r the antenna spacings at transmitter and receiver, θ_t and θ_r the transmitter and receiver elevations, and ϕ their relative azimuth angle. In turn, $\tilde{\boldsymbol{H}}$ is the Vandermonde matrix

$$\tilde{\boldsymbol{H}} = \begin{bmatrix} e^{j2\pi\eta\frac{0\times 0}{N_\mathsf{max}}} & \cdots & e^{j2\pi\eta\frac{(N_\mathrm{t}-1)\times 0}{N_\mathsf{max}}} \\ \vdots & \ddots & \vdots \\ e^{j2\pi\eta\frac{0\times(N_\mathrm{r}-1)}{N_\mathsf{max}}} & \cdots & e^{j2\pi\eta\frac{(N_\mathrm{t}-1)\times(N_\mathrm{r}-1)}{N_\mathsf{max}}} \end{bmatrix} \quad (10)$$

where $N_\mathsf{max} = \max(N_\mathrm{t}, N_\mathrm{r})$ while

$$\eta = \frac{(d_\mathrm{r}\cos\theta_\mathrm{r})(d_\mathrm{t}\cos\theta_\mathrm{t})N_\mathsf{max}}{\lambda D} \quad (11)$$

concisely describes any LOS setting with ULAs.

Uniform rectangular arrays can be expressed as the Kronecker product of ULAs, and expressions deriving from (8) emerge [21].

b) LOS with Planar Wavefronts: The planar wavefront counterpart to (8) is obtained by letting $D \to \infty$, whereby the channel becomes rank-1.

III. 1-Bit Capacity

Denote by $p_1, \ldots, p_{4^{N_\mathrm{t}}-1}$ the activation probabilities of the transmit quartets, such that $\sum_k p_k = 1$ and $p_{\boldsymbol{x}}(\mathbf{x}_k) = p_k/4$ with $\mathbf{x}_k$ any of the vectors in the kth quartet. Letting $\mathcal{H}(\cdot)$ indicate entropy, and with all the probabilities conditioned on $\boldsymbol{H}$, the mutual information is

$$\mathcal{I}(\mathsf{SNR}, \boldsymbol{H}) = \mathcal{H}(\boldsymbol{y}) - \mathcal{H}(\boldsymbol{y}|\boldsymbol{x}) \quad (12)$$
$$= 4\sum_{\ell=1}^{4^{N_\mathrm{r}-1}} p_{\boldsymbol{y}}(\mathbf{y}_\ell)\log_2\frac{1}{p_{\boldsymbol{y}}(\mathbf{y}_\ell)} - \mathcal{H}(\boldsymbol{y}|\boldsymbol{x}) \quad (13)$$

where $\mathbf{y}_\ell$ is any vector in the ℓth receive quartet and

$$p_{\boldsymbol{y}}(\mathbf{y}) = \sum_{k=1}^{4^{N_\mathrm{t}-1}}\frac{p_k}{4}\sum_{i=0}^{3} p_{\boldsymbol{y}|\boldsymbol{x}}(\mathbf{y}|\mathrm{j}^i\mathbf{x}_k) \quad (14)$$

with $p_{\boldsymbol{y}|\boldsymbol{x}}$ depending on SNR and $\boldsymbol{H}$ as per (7). Moreover,

$$\begin{aligned} p_{\boldsymbol{y}}(\mathbf{y}) = \sum_{k=1}^{4^{N_\mathrm{t}-1}}\frac{p_k}{4}\Bigg[&\prod_{n=0}^{N_\mathrm{r}-1} Q\left(-\Re\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Re\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &\cdot Q\left(-\Im\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Im\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &+ \prod_{n=0}^{N_\mathrm{r}-1} Q\left(\Re\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Im\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &\cdot Q\left(-\Im\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Re\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &+ \prod_{n=0}^{N_\mathrm{r}-1} Q\left(\Re\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Re\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &\cdot Q\left(\Im\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Im\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &+ \prod_{n=0}^{N_\mathrm{r}-1} Q\left(-\Re\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Im\{\boldsymbol{h}_n\mathbf{x}_k\}\right) \\ &\cdot Q\left(\Im\{\mathsf{y}_n\}\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}}\Re\{\boldsymbol{h}_n\mathbf{x}_k\}\right)\Bigg] \end{aligned} \quad (15)$$

and, because of the factorization of $p_{\boldsymbol{y}|\boldsymbol{x}}$ in (7),

$$\mathcal{H}(\boldsymbol{y}|\boldsymbol{x}) = \sum_{n=0}^{N_\mathrm{r}-1}\left(\mathcal{H}(\Re\{y_n\}|\boldsymbol{x}) + \mathcal{H}(\Im\{y_n\}|\boldsymbol{x})\right) \quad (16)$$
$$= \sum_{k=1}^{4^{N_\mathrm{t}-1}}\frac{p_k}{4}\sum_{i=0}^{3}\sum_{n=0}^{N_\mathrm{r}-1}\left(\mathcal{H}(\Re\{y_n\}|\boldsymbol{x} = \mathrm{j}^i\mathbf{x}_k) + \mathcal{H}(\Im\{y_n\}|\boldsymbol{x} = \mathrm{j}^i\mathbf{x}_k)\right) \quad (17)$$
$$= \sum_{k=1}^{4^{N_\mathrm{t}-1}}\frac{p_k}{4}\sum_{i=0}^{3}\sum_{n=0}^{N_\mathrm{r}-1}\Big[\mathcal{H}_\mathrm{b}\Big(Q\Big(-\sqrt{\tfrac{\mathsf{SNR}}{N_\mathrm{t}}}\Re\{\boldsymbol{h}_n\mathrm{j}^i\mathbf{x}_k\}\Big)\Big) + \mathcal{H}_\mathrm{b}\Big(Q\Big(-\sqrt{\tfrac{\mathsf{SNR}}{N_\mathrm{t}}}\Im\{\boldsymbol{h}_n\mathrm{j}^i\mathbf{x}_k\}\Big)\Big)\Big] \quad (18)$$

where $\mathcal{H}_\mathrm{b}(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function. Since changing i merely flips the sign of some of the Q-funcion arguments, and $Q(-\xi) = 1-Q(\xi)$ such that $\mathcal{H}_\mathrm{b}(Q(-\xi)) = \mathcal{H}_\mathrm{b}(Q(\xi))$, it follows that

$$\mathcal{H}(\boldsymbol{y}|\boldsymbol{x}) = \sum_{k=1}^{4^{N_\mathrm{t}}-1} p_k \sum_{n=0}^{N_\mathrm{r}-1} \left[\mathcal{H}_\mathrm{b}\left(Q\left(-\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}} \Re\{\boldsymbol{h}_n \mathbf{x}_k\} \right) \right) + \mathcal{H}_\mathrm{b}\left(Q\left(-\sqrt{\frac{\mathsf{SNR}}{N_\mathrm{t}}} \Im\{\boldsymbol{h}_n \mathbf{x}_k\} \right) \right) \right]. \tag{19}$$

The union of (13), (15), and (19) gives $\mathcal{I}(\mathsf{SNR}, \boldsymbol{H})$, whose evaluation involves $\mathcal{O}(4^{N_\mathrm{t}+N_\mathrm{r}-2})$ terms; this is prohibitive even for modest N_t and N_r, hence the interest in analytical characterizations. The 1-bit capacity is

$$C(\mathsf{SNR}, \boldsymbol{H}) = \max_{\{p_k\}:\sum_k p_k = 1} \mathcal{I}(\mathsf{SNR}, \boldsymbol{H}) \tag{20}$$

solvable with convex optimization tools or, alternatively, with the Blahut-Arimoto algorithm [**?**, ch. 13].

The 1-bit capacity exhibits three distinct regimes:

- Low SNR. This is a key regime at terahertz frequencies, given the difficulty in producing strong signals, the high propagation losses, and the noise bandwidth.
- Intermediate SNR. Here, the spectral efficiency improves sustainedly with the SNR.
- High SNR. This is a regime of diminishing returns, once the capacity nears $2\min(N_\mathrm{t}, N_\mathrm{r})$.

In the sequel we concentrate on the two leading regimes. The high-SNR regime, in contrast, is undesirable.

A. Low SNR

The low-SNR behavior is most conveniently examined with the mutual information expressed as function of the normalized energy per bit at the receiver,

$$\frac{E_\mathrm{b}}{N_0} = \frac{\mathsf{SNR}}{\mathcal{I}(\mathsf{SNR})}. \tag{21}$$

The mutual information is positive beyond [22, sec. 4.2]

$$\frac{E_\mathrm{b}}{N_{0\,\min}} = \lim_{\mathsf{SNR}\to 0} \frac{\mathsf{SNR}}{\mathcal{I}(\mathsf{SNR})} = \frac{1}{\dot{\mathcal{I}}(0)} \tag{22}$$

$$= \frac{\pi N_\mathrm{t}}{\mathrm{tr}(\boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{x}}\boldsymbol{H}^*)\log_2 e} \tag{23}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{x}} = \mathbb{E}\big[\boldsymbol{x}\boldsymbol{x}^*\big] = \sum_k p_k \mathbf{x}_k \mathbf{x}_k^*$, with (23) descending from (13), (15), and (19).

IV. Transmit Beamforming

Transmit beamforming corresponds to $\boldsymbol{\Sigma}_{\boldsymbol{x}}$ being rank-1, i.e., to $\boldsymbol{x}$ being drawn from a single quartet, with such quartet generally dependent on $\boldsymbol{H}$.

A. Low SNR

Here, transmit beamforming is not only conceptually appealing, but optimum. Indeed, (23) can be rewritten as

$$\frac{E_\mathrm{b}}{N_{0\,\min}} = \frac{\pi N_\mathrm{t}}{\sum_k p_k \|\boldsymbol{H}\mathbf{x}_k\|^2 \log_2 e}, \tag{24}$$

which is maximized by assigning probability 1 to the quartet $k^\star = \arg\max \|\boldsymbol{H}\mathbf{x}_k\|^2$. Therefore, it is optimum to beamform, and the optimum beamforming quartet is the one maximizing the received power. The task is then to determine $k^\star$ from within the $4^{N_\mathrm{t}}-1$ possible quartets.

For $N_\mathrm{t} = 1$, there is no need to optimize over k and

$$\frac{E_\mathrm{b}}{N_{0\,\min}} = \frac{\pi}{2N_\mathrm{r}\log_2 e}, \tag{25}$$

which amounts to 0.37 dB for $N_\mathrm{r} = 1$.

For $N_\mathrm{t} > 1$, it is useful to recognize that the choices for $\boldsymbol{x}$ that are bound to yield high values for $\|\boldsymbol{H}\boldsymbol{x}\|^2$ are those that project maximally on the dimension of $\boldsymbol{H}$ that offers the largest gain, namely the maximum-eigenvalue eigenvector of $\boldsymbol{H}^*\boldsymbol{H}$. This, in turn, requires that $\boldsymbol{x}$ mimics, as best as possible, the structure of that eigenvector; since the magnitude of the entries of $\boldsymbol{x}$ is fixed, this mimicking ought to be in terms of phases only. Formalizing this intuition, it is possible to circumvent the need to exhaustively search the entire field of $4^{N_\mathrm{t}}-1$ possibilities and conveniently identify a subset of only N_t quartet candidates that is sure to contain the one best aligning with the maximum-eigenvalue eigenvector of $\boldsymbol{H}^*\boldsymbol{H}$, denoted henceforth by $\boldsymbol{v}_0$. Precisely, if we let $\varphi_m = \angle(v_{0,m}) + \epsilon$ for $m = 0, \ldots, N_\mathrm{t}-1$, the N_t quartets in the subset can be determined as [19, app. A]

$$\boldsymbol{x}_k = \mathrm{sgn}\big(e^{\mathrm{j}\varphi_{k-1}}\boldsymbol{v}_0\big) \qquad k = 1, \ldots, N_\mathrm{t} \tag{26}$$

where ϵ is a small quantity, positive or negative. If $\boldsymbol{H}$ is rank-1, then this subset is sure to contain the optimum $\boldsymbol{x}_{k^\star}$ [15]; if the rank is higher, then optimality is not guaranteed, but the best value in the above subset is bound to yield excellent performance.

As of the $\frac{E_\mathrm{b}}{N_0}_{\min}$ achieved by $\boldsymbol{x}_{k^\star}$, its explicit evaluation is complicated, yet it can be shown to satisfy [19, app. A]

$$\frac{\pi}{2\,\lambda_0 \log_2 e} \le \frac{E_\mathrm{b}}{N_{0\,\min}} \le \frac{\pi^3 N_\mathrm{t}}{16\,\lambda_0 \|\boldsymbol{v}_0\|_1^2 \log_2 e} \tag{27}$$

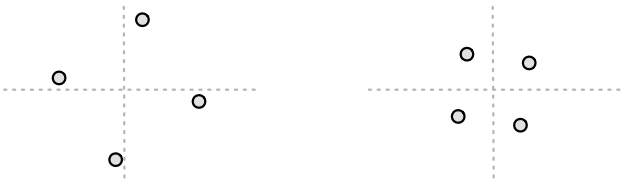
where λ_0 is the maximum eigenvalue of $\boldsymbol{H}^*\boldsymbol{H}$ while $\|\cdot\|_1$ denotes L1 norm.

B. Intermediate SNR

The low-SNR linearity of the mutual information in the received power is the root cause of the optimality of power-based beamforming in that regime. The orientation on the complex plane of the received signals is immaterial; a rotation merely shifts power from the real to the imaginary part, or vice versa. Likewise, the power split among receive antennas is immaterial to the low-SNR mutual information.

At higher SNRs, the linearity breaks down and the mutual information becomes a more intricate function of $\boldsymbol{H}\boldsymbol{x}$, such that proper signal orientations and power balances become important, to keep $\boldsymbol{h}_n\boldsymbol{x}$ away from the ADC quantization boundaries for $n = 0, \ldots, N_\mathrm{r}-1$. This has a dual consequence:

- Transmit beamforming ceases to be generally optimum, even if the channel is rank-1.



(a) $\boldsymbol{h}\mathsf{x}_k$ (b) $\boldsymbol{h}\mathsf{x}_\ell$

Fig. 1: Complex plane representation of the four values of $\boldsymbol{h}\boldsymbol{x}$ for a given $\boldsymbol{h}$ and a given quartet, with the ADC quantization boundaries indicated by dashed lines. Left-hand side, for x_k, which has a larger magnitude but worse orientation. Right-hand side, for x_ℓ, which has a smaller magnitude but better orientation. On this channel, quartet k yields a higher mutual at low SNR while quartet ℓ yields a higher mutual information beyond the low-SNR regime.

- Even within the confines of beamforming, solutions not based on maximizing power are more satisfying.

As exemplified in Fig. 1 for $N_\mathrm{r} = 1$, a beamforming quartet with a better complex-plane disposition at the receiver may be preferable to one yielding a larger magnitude. This is because, after a 1-bit ADC, only 90° rotations and no scalings are possible. The best beamforming quartet must simultaneously ensure large real and imaginary parts for $\boldsymbol{h}_n\boldsymbol{x}$ in a balanced fashion for $n = 0, \ldots, N_\mathrm{r}-1$, and the task of identifying this quartet is a fitting one for learning algorithms [23].

V. Equiprobable Signaling

The complementary strategy to beamforming is to activate multiple quartets, increasing the rank of $\boldsymbol{\Sigma}_{\boldsymbol{x}}$. Ultimately, all quartets can be activated equiprobably, such that $\boldsymbol{\Sigma}_{\boldsymbol{x}} = 2\boldsymbol{I}$. This renders the signals IID across the transmit antennas, i.e., pure spatial multiplexing.

A. Low SNR

With equiprobable signaling, (23) gives

$$\frac{E_\mathrm{b}}{N_{0\,\min}} = \frac{\pi}{2N_\mathrm{r}\log_2 e} \tag{28}$$

and, combining (27) and (28), the low-SNR advantage of optimum beamforming over equiprobable signaling, denoted by Δ_{BF}, is tightly bounded as

$$\frac{8\,\lambda_0\|\boldsymbol{v}_0\|_1^2}{\pi^2 N_\mathrm{t}N_\mathrm{r}} \le \Delta_{\mathsf{BF}} \le \frac{\lambda_0}{N_\mathrm{r}}. \tag{29}$$

Thus, the low-SNR advantage of beamforming is essentially determined by the maximum eigenvalue of $\boldsymbol{H}^*\boldsymbol{H}$. The advantage is largest in rank-1 channels, and minimal if all eigenvalues are equal.

B. Intermediate SNR

Beamforming is decidedly suboptimum beyond low SNRs, and activating multiple quartets becomes instrumental to surpass the 2-b/s/Hz mark. This is the case even in rank-1 channels, where the activation of multiple quartets allows producing richer signals; this can be seen as the 1-bit counterpart to higher-order constellations. And, given how the curse of dimensionality afflicts the computation of the optimum quartet probabilities, equiprobable signaling is a very enticing way of going about this.

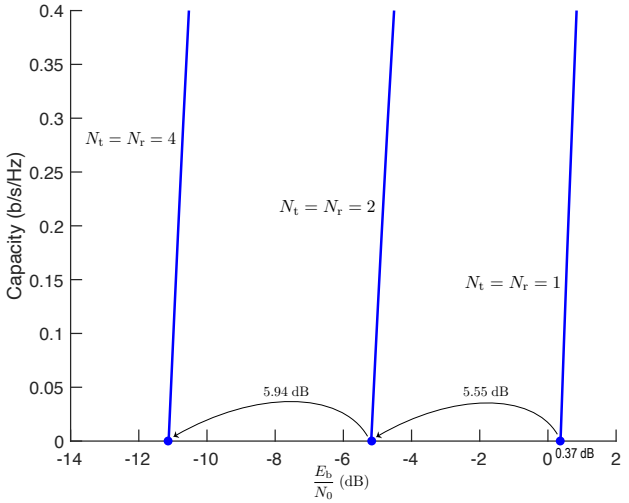


Fig. 2: Capacity as a function of E_b/N_0 for $N_\mathrm{t} = N_\mathrm{r} = 1$, $N_\mathrm{t} = N_\mathrm{r} = 2$, and $N_\mathrm{t} = N_\mathrm{r} = 4$, in a planar-wavefront LOS channel with half-wavelength antenna spacings, $\theta_\mathrm{t} = 0$, $\theta_\mathrm{r} = \pi/6$, and $\phi = \pi/4$.

VI. Channels of Interest

A. LOS with Planar Wavefronts

This channel is rank-1, hence the optimum $\frac{E_\mathrm{b}}{N_{0\,\min}}$ can be achieved with equality by the best beamforming quartet in subset (26). More conveniently for our purposes here, we can rewrite (8) as $\boldsymbol{H} = \sqrt{N_\mathrm{t}N_\mathrm{r}}\boldsymbol{u}\,\boldsymbol{v}^*$ where

$$u_n = \frac{1}{\sqrt{N_\mathrm{r}}}\, e^{-j\pi\frac{2n}{\lambda}d_\mathrm{r}\sin\theta_\mathrm{r}\cos\phi} \tag{30}$$

$$v_m = \frac{1}{\sqrt{N_\mathrm{t}}}\, e^{-j\pi\frac{2m}{\lambda}d_\mathrm{t}\sin\theta_\mathrm{t}}. \tag{31}$$

Irrespective of the array orientations, $\lambda_0 = N_\mathrm{t}N_\mathrm{r}$ and $\|\boldsymbol{v}_0\|_1^2 = N_\mathrm{t}$ such that (27) reverts to

$$\frac{\pi}{2\,N_\mathrm{t}N_\mathrm{r}\log_2 e} \le \frac{E_\mathrm{b}}{N_{0\,\min}} \le \frac{\pi^3}{16\,N_\mathrm{t}N_\mathrm{r}\log_2 e}, \tag{32}$$

which depends symmetrically on N_t and N_r. The significance of $\frac{E_\mathrm{b}}{N_{0\,\min}}$ as the key measure of the low-SNR performance can be appreciated in Fig. 2, which depicts the low-SNR capacity as a function of $\frac{E_\mathrm{b}}{N_0}$ for $N_\mathrm{t} = N_\mathrm{r} = 1$, 2, and 4 in an exemplary LOS setting. Adding antennas essentially displaces the capacity by the amount by which $\frac{E_\mathrm{b}}{N_{0\,\min}}$ changes.

Shown in Fig. 3 is how $\frac{E_\mathrm{b}}{N_{0\,\min}}$ improves with the number of antennas ($N_\mathrm{t} = N_\mathrm{r}$) for the same setting. Also shown are the values for equiprobable signaling, decidedly undesirable in this case as per (29), The low-SNR advantage of beamforming accrues steadily with the numbers of antennas and the bounds in (27) tightly bracket the optimum $\frac{E_\mathrm{b}}{N_{0\,\min}}$. The gap to full-resolution beamforming is small.

Moving up to intermediate SNRs, the beamforming and equiprobable-signaling performance on another setting is presented in Fig. 4. Also included is the actual capacity with $p_1, \ldots, p_{4^{N_\mathrm{t}}-1}$ optimized via Blahut-Arimoto. Up to when the 2-b/s/Hz ceiling is approached, beamforming performs splendidly. Past that level, and no matter the

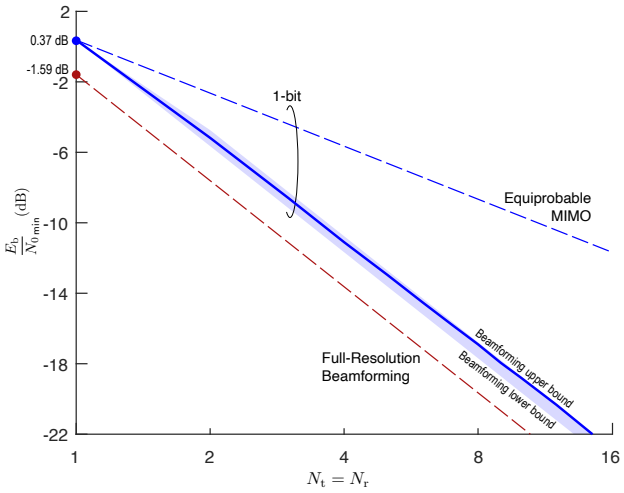


Fig. 3: Minimum E_b/N_0 as a function of $N_t = N_r$ for a planar-wavefront LOS channel with half-wavelength antenna spacings, $\theta_t = 0$, $\theta_r = \pi/6$, and $\phi = \pi/4$: 1-bit beamforming (exact values in solid, interval spanned by the bounds in shaded) vs equiprobable signaling. Also shown is the performance with full resolution.

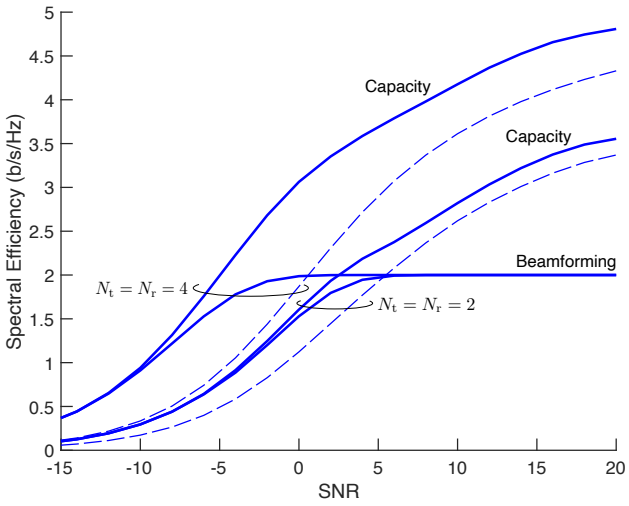


Fig. 4: Spectral efficiency as a function of SNR for $N_t = N_r = 2$ and $N_t = N_r = 4$, in a planar-wavefront LOS channel with half-wavelength antenna spacings, $\theta_t = \pi/4$, $\theta_r = \pi/6$, and $\phi = \pi/4$. In solid, capacity and beamforming; in dashed, equiprobable signaling.

rank-1 nature of the channel, equiprobable signaling is highly superior, tracking the capacity to within a roughly constant shortfall.

B. LOS with Spherical Wavefronts

The scope of channels in this class is very large, depending on the array topologies and orientations; for concreteness, we focus on ULAs and draw insights whose generalization would be welcome follow-up work. A key property of ULA-spawned channels in this class is [20]

$$\boldsymbol{H}^*\boldsymbol{H} \approx \frac{N_{\mathsf{max}}}{\eta}\boldsymbol{D}_{\mathrm{tx}}^*\boldsymbol{F}\mathrm{diag}(\underbrace{1,\ldots,1}_{\eta N_{\mathsf{min}}},0,\ldots,0)\boldsymbol{F}^*\boldsymbol{D}_{\mathrm{tx}}$$

where the approximation sharpens with the numbers of antennas while $\boldsymbol{F}$ is a unitary Fourier matrix, $\boldsymbol{D}_{\mathrm{tx}}$ and η

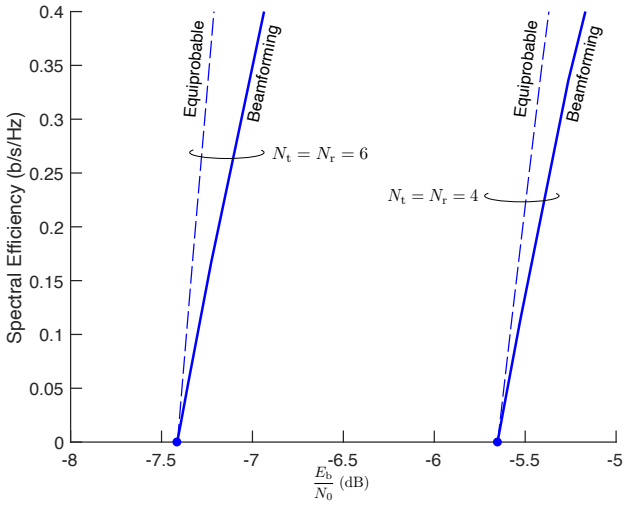


Fig. 5: Spectral efficiency as a function of E_b/N_0 for $N_t = N_r = 4$ and $N_t = N_r = 6$ in an LOS channel with $\eta = 1$. In solid, beamforming performance; in dashed, capacity with equiprobable signaling.

are as introduced in Sec. II-B, and $N_{\mathsf{min}} = \min(N_t, N_r)$. Thus, $\lambda_0 \approx N_{\mathsf{max}}/\eta$ and $\|\boldsymbol{v}_0\|_1^2 \approx N_t$. By specializing (27), the $\frac{E_b}{N_{0\,\min}}$ attained by beamforming is seen to satisfy

$$\frac{\pi\eta}{2N_{\mathsf{max}}\log_2 e} \lesssim \frac{E_b}{N_{0\,\min}} \lesssim \frac{\pi^3\eta}{16N_{\mathsf{max}}\log_2 e}, \qquad (33)$$

which indicates that a smaller η is preferable at low SNR, meaning antennas as tightly spaced as possible and array orientations as endfire as possible: wavefront curvatures trim the beamforming gains, and reducing η mitigates the extent of such curvatures. With growing η, the low-SNR performance degrades, but beamforming retains an edge over equiprobable signaling for $\eta < 1$ or $N_t > N_r$. Alternatively, for $\eta = 1$ and $N_t = N_r$, (33) is no better than the equiprobable-signaling $\frac{E_b}{N_{0\,\min}}$ in (28). In fact, for this important configuration whose eigenvalues are equal [24], *any* transmission strategy achieves this $\frac{E_b}{N_{0\,\min}}$; indeed, using $\boldsymbol{H}^*\boldsymbol{H} = N_r\boldsymbol{I}$ and $\|\mathsf{x}_k\|^2 = 2N_t$ in (24), (28) emerges irrespective of $p_1,\ldots,p_{4^{N_t}-1}$.

At intermediate SNRs, the full-resolution wisdom is that the performance depends only on η and it improves monotomically with η up to $\eta = 1$, where capacity is achieved by IID signaling. These insights, underpinned by the approximate equality of the ηN_{min} nonzero eigenvalues of $\boldsymbol{H}^*\boldsymbol{H}$, cease to hold in the 1-bit realm due to the transmitter's inability of accessing those singular values via precoding. Indeed, when the only ability is to manipulate the quartet probabilities (see Fig. 6):

- The performance does not depend only on η, but further on θ_t, θ_r, ϕ, D, and d_t and d_r.
- The optimum configuration need not correspond to $\eta = 1$.

The main takeaway for our purpose, though, is that equiprobable signaling closely tracks the capacity.

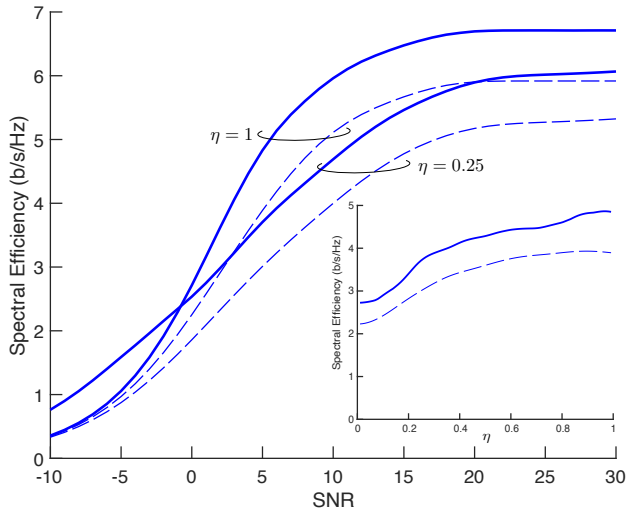


Fig. 6: Main plot: spectral efficiency as a function of SNR for $N_{\mathrm{t}} = N_{\mathrm{r}} = 4$, both with optimized (solid lines) and uniform (dashed lines) quartet probabilities. Inset: spectral efficiency as a function of η for $\mathsf{SNR} = 5$ dB. The channel is LOS and the arrays are broadside, $d_{\mathrm{t}} = d_{\mathrm{r}}$.

VII. Conclusion

The computation of the 1-bit capacity becomes quickly unwieldy with the number of antennas and the derivation of general 1-bit precoding solutions is a formidable task. Fortunately, such general precoding can be skirted via a judicious switching of beamforming and equiprobable signaling, with the added benefits that these strategies are amenable to analytical characterizations and that their requirements in terms of channel-state information at the transmitter are minimal: $\log_2 4^{N_{\mathrm{t}}-1} = 4\,(N_{\mathrm{t}}-1)$ bits for beamforming, none for equiprobable signaling.

Although the transition from beamforming to equiprobable signaling could be finessed by progressively activating quartets as the SNR grows, a direct switching at some appropriate point suffices to operate within a few dB of capacity at both low and intermediate SNRs.

Of much interest would be to assess the impact of channel estimation at the receiver, by extending existing results for full-resolution DACs and 1-bit ADCs [8], [25]. Also pertinent would be to establish the bandwidth over which a frequency-flat representation suffices for each channel model, and to extend the analyses to account for the intersymbol interference caused by spatial widening, i.e., by the distinct delays between the various transmit and receive antennas.

Acknowledgment

Work supported by the European Research Council under grant agreement 694974, by Projects RTI2018-102112 and 101040, and by ICREA.

References

[1] H. Hamada *et al.*, "300-GHz-band 120-Gb/s wireless front-end based on InP-HEMT PAs and mixers," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 9, pp. 2316–2335, 2020.

[2] B. Murmann, "The race for the extra decibel: A brief review of current ADC performance trajectories," *IEEE Solid-State Circuits Magazine*, vol. 7, no. 3, pp. 58–66, 2015.

[3] J. A. Nossek and M. T. Ivrlač, "Capacity and coding for quantized MIMO systems," in *Int'l Conf. Wireless Commun. and Mobile Computing*, 2006, pp. 1387–1392.

[4] A. Mezghani and J. A. Nossek, "Analysis of Rayleigh-fading channels with 1-bit quantized output," in *IEEE Int'l Symp. Inform. Theory (ISIT)*, 2008, pp. 260–264.

[5] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, 2009.

[6] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2156–2168, 2014.

[7] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Processing*, vol. 63, no. 20, pp. 5498–5512, 2015.

[8] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Processing*, vol. 65, no. 15, pp. 4075–4089, 2017.

[9] A. Mezghani, R. Ghiat, and J. Nossek, "Transmit processing with low resolution D/A-converters," in *IEEE Int'l Conf. Electronics, Circuits and Systems-(ICECS'09)*, 2009, pp. 683–686.

[10] B. Usman, H. Jedda, A. Mezghani, and J. Nossek, "MMSE precoder for massive MIMO using 1-bit quantization," in *IEEE Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 3381–3385.

[11] H. Jedda, J. A. Nossek, and A. Mezghani, "Minimum BER precoding in 1-bit massive MIMO systems," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016.

[12] J. Guerreiro, R. Dinis, and P. Montezuma, "Use of 1-bit digital-to-analogue converters in massive MIMO systems," *Electr. Letters*, vol. 52, no. 9, pp. 778–779, 2016.

[13] K. Gao, N. Estes, B. Hochwald, J. Chisum, and N. Laneman, "Power-performance analysis of a simple one-bit transceiver," in *Inform. Theory and Applic. Workshop (ITA)*, 2017.

[14] K. Gao, N. Laneman, and B. Hochwald, "Capacity of multiple one-bit transceivers in a Rayleigh environment," in *IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2018, pp. 1–6.

[15] K. Gao, J. N. Laneman, and B. Hochwald, "Beamforming with multiple one-bit wireless transceivers," in *Inform. Theory and Applic. Workshop (ITA)*, 2018.

[16] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Processing*, vol. 65, no. 17, pp. 4624–4634, 2017.

[17] Y. Nam, H. Do, Y.-S. Jeon, and N. Lee, "On the capacity of MISO channels with one-bit ADCs and DACs," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2132–2145, 2019.

[18] A. Bazrafkan and N. Zlatanov, "Asymptotic capacity of massive MIMO with 1-bit ADCs and 1-bit DACs at the receiver and at the transmitter," *IEEE Access*, vol. 8, pp. 152 837–152 850, 2020.

[19] A. Lozano, "1-bit MIMO for terahertz channels," *arXiv: 2109.04390*, 2021.

[20] H. Do, N. Lee, and A. Lozano, "Reconfigurable ULAs for line-of-sight MIMO transmission," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 2933–2947, 2020.

[21] P. Larsson, "Lattice array receiver and sender for spatially orthonormal MIMO communication," in *Proc. IEEE Veh. Technol. Conf.*, May 2005, pp. 192–196.

[22] R. W. Heath, Jr. and A. Lozano, *Foundations of MIMO Communication*. Cambridge University Press, 2018.

[23] R. Nikbakht and A. Lozano, "Terahertz transmit beamforming with 1-bit DACs and ADCs," in *European Signal Processing Conf. (EUSIPCO)*, 2021.

[24] T. Haustein and U. Krauger, "Smart geometrical antenna design exploiting the LOS component to enhance a MIMO system based on Rayleigh-fading in indoor scenarios," in *Proc. IEEE Personal Indoor Mobile Radio Commun.*, Sep. 2003, pp. 1144–1148.

[25] M. T. Ivrlac and J. A. Nossek, "On MIMO channel estimation with single-bit signal-quantization," in *ITG Smart Antenna Workshop*, 2007.