# Capacity of Space-Time Wireless Channels: A Physical Perspective

Akbar Sayeed
University of Wisconsin-Madison,
akbar@engr.wisc.edu

Vasanthan Raghavan
University of Wisconsin-Madison,
raghavan@cae.wisc.edu

Jayesh Kotecha
University of Wisconsin-Madison,
jkotecha@ece.wisc.edu

*Abstract* — **Existing results on MIMO channel capacity assume a rich scattering environment in which the channel power scales quadratically with the number of antennas, resulting in linear capacity scaling with the number of antennas. However, such scaling in channel power is physically impossible indefinitely. We thus address the following fundamental question: For a given channel power scaling law, what is the best achievable capacity scaling law? For a channel power scaling, $\rho_c(N) = \mathcal{O}(N^\gamma)$, $\gamma \in (0, 2]$, we argue that the channel capacity cannot scale faster than $C(N) = \mathcal{O}(\sqrt{\rho_c(N)}) = \mathcal{O}(N^{\gamma/2})$. Our approach is based on a family of space-time channels corresponding to different distributions of channel power in the spatial signal space dimensions. We develop the concept of an ideal MIMO channel that achieves the optimal scaling law for a given $\rho_c(N)$. For a given number of antennas, unlike existing results that either emphasize the low or high SNR regimes, we propose a methodology for capacity-optimal signaling at any SNR. The methodology is based on creating the ideal channel from any given physical scattering environment via adaptive-resolution array configurations.**

## I. Introduction

We study the coherent capacity of multiple antenna (MIMO) wireless communication systems from a physical perspective. The pioneering works of [1] and [2] predict a dramatic linear increase in capacity with the number of antennas. However, an implicit assumption is that the channel power $\rho_c(N) = E[\mathrm{trace}(\mathbf{H}\mathbf{H}^H)]$ scales as $\mathcal{O}(N^2)$ due to the assumption of i.i.d. entries of the $N \times N$ channel matrix $\mathbf{H}$. However, such scaling in $\rho_c(N)$ is not possible indefinitely from a physical viewpoint since it implies that the received signal power increases linearly with $N$ for a given transmit signal power. While this assumption may be accurate for small $N$, due to higher coupling of the transmitted power to the receiver relative to single-antenna systems, it cannot be justified for large $N$.

Existing works on attacking the capacity of realistic MIMO channels have focussed on the correlation between entries of $\mathbf{H}$. For example, Chuah et al. addressed capacity scaling of correlated MIMO channels assuming a Kronecker product correlation model [3] and showed that capacity scales linearly, albeit with a smaller scaling factor than an i.i.d. channel. However, $\rho_c(N)$ scales as $\mathcal{O}(N^2)$ under the assumptions in [3]. Evidently, this assumption on $\rho_c(N)$ is prevalent in virtually all existing works.

We thus pose the following fundamental question:

*Given a channel power scaling law, $\rho_c(N) \sim \mathcal{O}(N^\gamma)$, $0 < \gamma \leq 2$, what is the best achievable MIMO capacity scaling law?*

We argue that the coherent channel capacity cannot scale faster than $\mathcal{O}(\sqrt{\rho_c(N)})$ and such a scaling law is achievable. Our approach is based on a unitarily equivalent virtual representation of physical MIMO channels [4] that provides an intuitive and tractable characterization of physical MIMO channels. The entries of the virtual channel matrix $\mathbf{H}_v$ are uncorrelated for any given scattering environment and the correlation in $\mathbf{H}$ is captured by the sparseness of the non-vanishing entries in $\mathbf{H}_v$. Under mild assumptions, for a given $\rho_c(N)$, the number of non-vanishing entries in $\mathbf{H}_v$, $D(N)$ (degrees of freedom (DoF)), also scale as $\mathcal{O}(\rho_c(N))$. Furthermore, the number of propagation paths captured by the MIMO system also scale as $\mathcal{O}(\rho_c(N))$. Thus, for a given $\gamma < 2$, the DoF in the virtual channel matrix scale as $D(N) \sim \mathcal{O}(N^\gamma)$ compared to the maximum $\mathcal{O}(N^2)$. The optimal capacity scaling law $C(N) \sim \mathcal{O}(\sqrt{\rho_c(N)})$ is achieved by an *ideal MIMO channel* that corresponds to an optimal distribution of the channel power in spatial signal space dimensions. The results in this paper generalize our earlier results in [5] which showed that linear capacity scaling is possible for $\rho_c(N) \sim \mathcal{O}(N^2)$ whereas $C(N)/N$ saturates for $\rho_c(N) \sim \mathcal{O}(N)$.

To develop the concept of the ideal MIMO channel, we consider a family of channels, with different distributions of channel power in the $D(N)$ DoF, whose capacity admits the following intuitive interpretation

$$C(N) \approx p(N) \log(1 + \rho_{rx}(N)) \quad (1)$$

where $p(N)$ denotes the number of parallel channels or multiplexing gain, and $\rho_{rx}(N)$ is the receive SNR per parallel channel. On one extreme are *beamforming channels* in which the channel power is distributed to maximize $\rho_{rx}$ (at the expense of $p(N)$), and on the other extreme are *multiplexing channels* which favor $p(N)$ over $\rho_{rx}(N)$. The ideal channel lies in between and corresponds to an optimal distribution of channel power to balance $p(N)$ and $\rho_{rx}(N)$.

Our investigation of optimal capacity scaling via the ideal channel also provides two interesting by-products. First, unlike existing results that either emphasize the low or high SNR regimes, our approach provides key insights into capacity-optimal signaling for any $N$ and any SNR. Second, it suggests a physically motivated methodology to construct the ideal channel from any given scattering environment via adaptive-resolution array configurations.

The next section reviews the virtual channel representation. Section III develops the notion of the ideal channel and optimal capacity scaling. Section IV proposes a methodology for creating the ideal channel from any given scattering environment. A note about notation: $f(N) \sim \mathcal{O}(g(N)) \leftrightarrow 0 < K_1 \leq \lim_{N\to\infty} \left|\frac{f(N)}{g(N)}\right| \leq K_2 < \infty$, and $f(N) \sim o(g(N)) \leftrightarrow \lim_{N\to\infty} \left|\frac{f(N)}{g(N)}\right| = 0$.

## II. Virtual Modeling of Physical Channels

We consider a single-user narrowband MIMO system with uniform linear arrays (ULAs) of $N_t$ transmit and $N_r$ receive antennas. The transmitted signal $\mathbf{s}$ and the received signal $\mathbf{x}$ are related by

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{n} \tag{2}$$

where $\mathbf{H}$ is the MIMO channel matrix and $\mathbf{n}$ is the AWGN at the receiver. The statistics of entries of $\mathbf{H}$ control the channel capacity. Ideal channel modeling assumes that the entries of $\mathbf{H}$ are i.i.d. Gaussian random variables, which is an unrealistic assumption but facilitates analysis. On the other extreme is a widely used and accurate physical model:

$$\mathbf{H} = \sum_{l=1}^{L} \beta_l \; \mathbf{a}_r(\theta_{r,l})\mathbf{a}_t^H(\theta_{t,l}) \tag{3}$$

where the transmitter and receiver arrays are coupled via $L$ propagation paths with complex path gains $\{\beta_l\}$, Angles of Departure (AoD) $\{\theta_{t,l}\}$ seen at the transmitter, and Angles of Arrival (AoA) $\{\theta_{r,l}\}$ seen at the receiver. In (3), $\mathbf{a}_r(\theta_r)$ and $\mathbf{a}_t(\theta_t)$ denote the receiver response and transmitter steering vectors for receiving/transmitting in the normalized direction $\theta_r/\theta_t$. The variable $\theta$ is related to the physical angle variable $\phi \in [-\pi/2, \pi/2]$ as $\theta = d\sin(\phi)/\lambda$ where $d$ is the antenna spacing and $\lambda$ is the wavelength of propagation. Both $\mathbf{a}_r(\theta_r)$ and $\mathbf{a}_t(\theta_t)$ are periodic in $\theta$ with period 1 [4]. We note that the non-linear dependence of $\mathbf{H}$ on the AoA's and AoD's makes the physical model unsuitable, if not intractable, for capacity analysis.

The recently introduced *virtual channel representation* [4] connects physical models to statistical (e.g., i.i.d.) models by providing a natural and tractable characterization of physical channels. As illustrated in Fig. 1, the virtual representation characterizes the MIMO channel via beamforming in the direction of fixed virtual transmit and receive angles:

$$\mathbf{H} = \sum_{m=1}^{N_r}\sum_{n=1}^{N_t} H_v(m,n)\mathbf{a}_r(\tilde{\theta}_{r,m})\mathbf{a}_t^H(\tilde{\theta}_{t,n}) = \tilde{\mathbf{A}}_r\mathbf{H}_v\tilde{\mathbf{A}}_t^H \tag{4}$$

where $\left\{\tilde{\theta}_{r,m} = \frac{m}{N}\right\}$ and $\left\{\tilde{\theta}_{t,n} = \frac{n}{N}\right\}$ are fixed virtual receive and transmit angles that uniformly sample the unit $\theta$ period and result in unitary (DFT) matrices $\tilde{\mathbf{A}}_r$ and $\tilde{\mathbf{A}}_r$. Thus, $\mathbf{H}$ and $\mathbf{H}_v$ are unitarily equivalent: $\mathbf{H}_v = \tilde{\mathbf{A}}_r^H\mathbf{H}\tilde{\mathbf{A}}_t^H$. The virtual representation is linear and is characterized by the virtual matrix $\mathbf{H}_v$

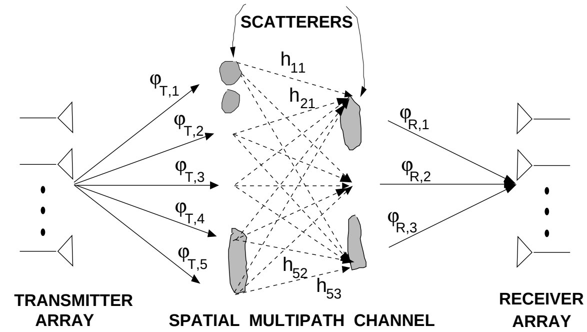


Figure 1: Virtual modeling of physical channels. The virtual angles are fixed a priori and their spacing defines the spatial resolution. The channel is characterized by the virtual channel coefficients, $\{H_v(m,n) = h_{m,n}\}$, that couple the $N_t$ virtual transmit angles, $\{\theta_{t,n}\}$, with the $N_r$ virtual receive angles, $\{\theta_{r,m}\}$.

whose entries represent the coupling between transmit and receive beams in the virtual directions.

Virtual path partitioning relates the virtual coefficients to the physical paths gains [4]

$$H_v(m,n) \approx \left[\sum_{l\in S_{R,m}\cap S_{T,n}} \beta_l\right] \tag{5}$$

where $S_{r,m}$ and $S_{t,n}$ are the spatial resolution bins of size $1/N_r$ and $1/N_t$ corresponding to the $m$-th receive and $n$-th transmit virtual angle. Thus, $H_v(m,n)$ is approximately the sum of gains of all paths whose transmit and receive angles lie within the $(m,n)$-th resolution bin. If there are no paths in a particular resolution bin, the corresponding $H_v(m,n) \approx 0$. It follows that the virtual channel coefficients are approximately independent since disjoint sets of paths contribute to distinct coefficients. We note that the approximation in (5) improves with increasing antennas. The readers are referred to [4, 6, 5] for a detailed discussion of the virtual representation.

For the remainder of this section, we consider $N = N_r = N_t$. The (coherent) ergodic capacity of a narrowband MIMO channel, assuming knowledge of $\mathbf{H}$ at the receiver, is given by [1, 2]

$$\begin{aligned} C(N) &= \max_{\mathrm{Tr}(\mathbf{Q})\leq\rho} E_{\mathbf{H}}\left[\log\det\left(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\right] \\ &= \max_{\mathrm{Tr}(\tilde{\mathbf{Q}})\leq\rho} E_{\mathbf{H}_v}\left[\log\det\left(\mathbf{I} + \mathbf{H}_v\tilde{\mathbf{Q}}\mathbf{H}_v^H\right)\right] \end{aligned} \tag{6}$$

where $\rho$ is the transmit SNR, and $\mathbf{Q} = E[\mathbf{s}\mathbf{s}^H]$ is the transmit covariance matrix, and $\tilde{\mathbf{Q}} = \tilde{\mathbf{A}}_T^H\mathbf{Q}\tilde{\mathbf{A}}_T$. It is shown in [7] that a diagonal $\tilde{\mathbf{Q}}$ achieves capacity in the virtual domain. For the rest of the paper, we will work in the virtual domain and use $\mathbf{H}$ to denote $\mathbf{H}_v$.

For any given $N$, define the the channel power as

$$\rho_c(N) = E\left[\mathrm{trace}(\mathbf{H}\mathbf{H}^H)\right] = \sum_{l=1}^{L} E|\beta_l|^2. \tag{7}$$

Suppose that $\mathbf{H}$ has $D(N) \leq N^2$ non-vanishing entries. The $D(N)$ (independent) non-zero entries represent the statistically independent degrees of freedom (DoF) in the MIMO channel. We assume that all the non-zero virtual entries have constant unit power. Thus $\rho_c(N) = D(N)$ and the channel power scaling also reflects the growth rate of the DoF in $\mathbf{H}$. Furthermore, if we assume finite power due to each propagation path and finite number of paths in each resolution bin in (5), then $L(N) \sim \mathcal{O}(\rho_c(N)) = \mathcal{O}(D(N))$: if $\rho_c(N) = D(N)$ scale at a certain rate, then $L(N)$ must also scale at the same rate. This makes physical sense since as $N$ increases (for fixed antenna spacing), the number of physical paths captured by the arrays also increases due to larger array apertures. We note that the i.i.d. model assumes that $\rho_c(N) = D(N) = N^2$. In this paper, we consider slower scaling laws: $\rho_c(N) = D(N) < N^2$ (sparse $\mathbf{H}$). Under our assumptions, it is convenient to model a sparse $N \times N$ virtual matrix with $D(N)$ DoF as

$$\mathbf{H} = \mathbf{M} \odot \mathbf{H}_{iid} \tag{8}$$

where $\odot$ denotes elementwise product, $\mathbf{H}_{iid}$ is an i.i.d. matrix of $\mathcal{CN}(0,1)$ entries, and $\mathbf{M}$ is a mask matrix with $D(N)$ unit entries and zeros elsewhere.

## III. Optimal Capacity Scaling: The Ideal MIMO Channel

Capacity scaling of MIMO channels depends on three fundamental quantities: 1) transmit SNR $\rho$, 2) channel power scaling $\rho_c(N)$, and 3) the distribution of the channel power in the $D(N) = \rho_c(N)$ spatial DoF. For any given $\rho_c(N)$, there is an optimal (ideal) distribution of the DoF that results in the most favorable capacity scaling.

Let us first consider an example to motivate our approach. According to Thm. 1 below, the asymptotic capacity of a MIMO channel of dimension $N$ admits the following interpretation

$$C(N) = p(N) \log(1 + \rho_{rx}(N)) \tag{9}$$

where $p(N)$ is the number of parallel channels (rank of $\mathbf{H}$), and $\rho_{rx}(N)$ is the receive SNR per parallel channel

$$\rho_{rx}(N) = \frac{E[\|\mathbf{Hs}\|^2]}{p(N)} = \rho \frac{\rho_c(N)}{p^2(N)} = \rho \frac{q(N)}{p(N)}, \tag{10}$$

where $q(N) = D(N)/p(N)$ is the number of DoF per parallel channel. Suppose that $D(N) = \rho_c(N) = N$. We consider three channels corresponding to different canonical distributions of the $D(N)$ entries in $\mathbf{H}$ into $p(N)$ and $q(N)$. On one extreme is the **beamforming channel** for which $\mathbf{H}_{bf}$ is an $N \times 1$ matrix with i.i.d. entries: $p(N) = 1$ and all the DoF are distributed to maximize $\rho_{rx}(N) = \rho N$. On the other extreme is the **multiplexing channel** for which $\mathbf{H}_{mux}$ is an $N \times N$ diagonal matrix with i.i.d. diagonal entries: the DoF are distributed to maximize the multiplexing gain so that $p(N) = N$ and $\rho_{rx}(N) = \rho/N$. In between the two extremes is the **ideal channel** for which $\mathbf{H}_{id}$ is a $\sqrt{N} \times \sqrt{N}$ matrix with i.i.d. entries: the DoF are optimally distributed so that $p(N) = q(N) = \sqrt{N}$ and $\rho_{rx}(N) = \rho$. For large $N$, the capacities of the three channels are: $C_{bf}(N) = \log(1 + \rho N)$, $C_{mux} = N \log(1 + \rho/N) \to \rho$, and $C_{id}(N) = \sqrt{N} \log(1 + \rho)$. It is clear that the ideal channel achieves the best capacity scaling $C_{id}(N) = \mathcal{O}(\sqrt{N}) = \mathcal{O}(\sqrt{\rho_c(N)})$ which is in stark contrast to i.i.d. channels for which $\rho_c(N) = N^2$ and the best capacity scaling $C_{iid}(N) = \mathcal{O}(N)$ is achieved with maximum multiplexing gain $N$. We next develop the notion of the ideal channel and argue that it achieves the best possible capacity scaling for any given channel power scaling $\rho_c(N)$. We first define a family of $N$ dimensional mask matrices which distribute $D(N)$ entries into $p(N)$ and $q(N)$ to span channels from $\mathbf{H}_{bf}$ to $\mathbf{H}_{id}$ to $\mathbf{H}_{mux}$.

**Definition 1 A family of mask matrices.** *Consider a mask matrix* $\mathbf{M}$ *with* $D(N)$ *non-zero entries distributed over* $1 \leq p(N) \leq N$ *columns and* $1 \leq q(N) \leq N$ *non-zero entries in each column. Let* $r(N) = \max(p(N), q(N))$. *Specifically,* $\mathbf{M}$ *is a* $r(N) \times p(N)$ *matrix with non-zero entries given by*

$$\begin{aligned} \mathbf{M}(n + m \mod r(N), n) &= 1 \\ n = 1, \cdots, p(N) \quad &, \quad q_-(N) \leq m \leq q_+(N) \end{aligned} \tag{11}$$

*where* $q_-(N) = \lceil -(q(N)-1)/2 \rceil$ *and* $q_+(N) = \lceil (q(N)-1)/2 \rceil$. $p(N)$ *and* $q(N)$ *satisfy*

$$D(N) = p(N) q(N). \square \tag{12}$$

Note that for $q(N) \geq p(N)$, $\mathbf{M}$ is a $q(N) \times p(N)$ matrix of ones, whereas for $q(N) < p(N)$ $\mathbf{M}$ is a $p(N) \times p(N)$ matrix with $q(N)$ non-vanishing diagonals (a $q(N)$-connected $p(N)$-dimensional matrix in [5]). In all cases, there are $q(N)$ non-zero entries in each column and $p(N)$ non-zero entries in each row. Thus, the corresponding virtual channel matrices $\mathbf{H}$ defined via (8) are *regular* [8] with $\text{rank}(\mathbf{H}) = \min(r(N), p(N)) = p(N)$ for which the uniform-power input over the $p(N)$ parallel channels is optimal [8].

We want to address the following fundamental question: *For a given channel power/DoF scaling law* $\rho_c(N) = D(N)$, *what is the best possible achievable capacity scaling law?* To answer this question, we allow both $p(N)$ and $q(N)$ in Definition 1 to scale under the constraint $D(N) = p(N)q(N)$. To gain insight we also consider specfic scaling laws for $D$, $p$ and $q$.

**Definition 2 Scaling laws for $\rho_c$, $p$ and $q$.** *Let* $D(N) = \rho_c(N) = N^\gamma$, $\gamma \in (0, 2]$; $p(N) = \delta_p N^\alpha$, $\delta_p \in (0, \infty)$, $\alpha \in [0, 1]$; *and* $q(N) = \frac{1}{\delta_p} N^\beta$, $\beta \in [0, 1]$ *in Defintion 1. For a given* $\gamma$, *the feasible range for* $\alpha$ *is*

$$\alpha_{min} = \max(\gamma - 1, 0) \leq \alpha \leq \min(\gamma, 1) = \alpha_{max} \tag{13}$$

*and the corresponding* $\beta = \gamma - \alpha$. $\square$

Note that for $1 \leq \gamma \leq 2$, $\alpha_{min} = \gamma - 1 \geq 0$ and $\alpha_{max} = 1$. For $0 < \gamma \leq 1$, $\alpha_{min} = 0$ and $\alpha_{max} = \gamma$. The family of channels

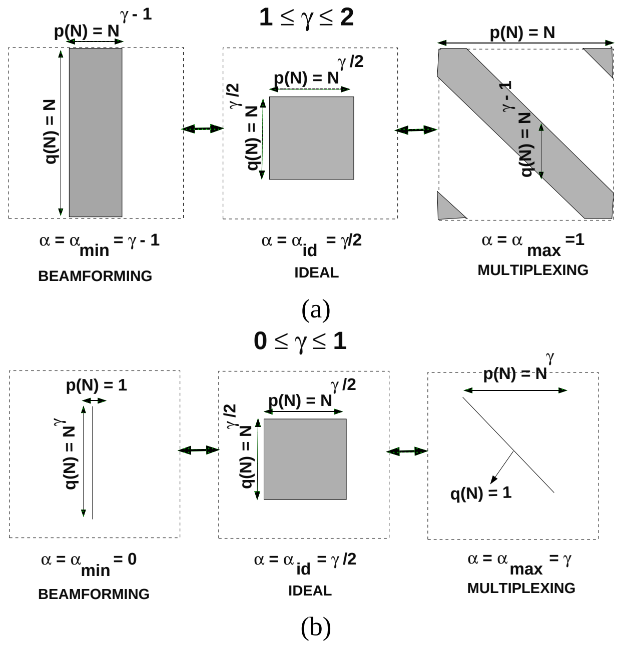


Figure 2: A schematic illustrating the family of channels. (a) $\gamma \in [1, 2]$. (b) $\gamma \in [0, 1]$.

is illustrated in Fig. 2.[1]

We now identify three distinct scaling regimes for $p(N)$ which result in distinct capacity scaling behaviors:

**Beamforming regime** ($\mathbf{H}_{bf}$): $p(N) = o(q(N)) \leftrightarrow \alpha \in [\alpha_{min}, \gamma/2)$. In this regime, $\rho_{rx}(N) \to \infty$.

**Multiplexing regime** ($\mathbf{H}_{mux}$): $q(N) = o(p(N)) \leftrightarrow \alpha \in (\gamma/2, \alpha_{max}]$. In this regime, $\rho_{rx}(N) \to 0$.

**Ideal regime** ($\mathbf{H}_{id}$): $p(N) = \mathcal{O}(q(N)) \leftrightarrow \alpha = \gamma/2$. Define $c = \lim_N q(N)/p(N) = \frac{1}{\delta_p^2} \in (0, \infty)$. In this regime, $\rho_{rx}(N) \to c\rho$.

We note that for $\gamma = 2$, both $p$ and $q$ are $\mathcal{O}(N)$. Thus, $\mathbf{H}_{bf}$, $\mathbf{H}_{id}$, and $\mathbf{H}_{mux}$ all result in the well-known $\mathcal{O}(N)$ capacity scaling, albeit with different slopes.

**Theorem 1 Capacity scaling.** *Suppose the channel power scales as $\rho_c(N) = D(N)$. The asymptotic capacity scaling behavior for the family of channels in Definition 1 is given by*

$$C(N) \approx p(N)\log(1 + \rho_{rx}(N)) \quad (14)$$

$$= p(N)\log(1 + \rho q(N)/p(N)) \quad (15)$$

$$= \delta_p N^{\alpha} \log(1 + \rho c N^{\gamma - 2\alpha}) \quad (16)$$

*where the last equality is for any feasible $(\gamma, \alpha)$ in Definition 2. The scaling law is asymptotically tight in the beamforming and multiplexing regimes for any $\rho$:*

$$C_{bf}(N) \sim p(N)\log\left(\rho \frac{q(N)}{p(N)}\right) \quad (17)$$

$$= \delta_p N^{\alpha} \log\left(\rho c N^{\gamma - 2\alpha}\right) \quad (18)$$

$$C_{mux}(N) \sim \rho q(N) = \frac{\rho}{\delta_p} N^{\gamma - \alpha}. \quad (19)$$

*In the ideal regime, the scaling law is tight for large $c\rho$:*

$$C_{id}(N) \sim p(N)\log(\rho c) = \delta_p N^{\gamma/2} \log(\rho c). \quad (20)$$

**Proof. Case I:** $q(N) > p(N)$. For any $N$, $\mathbf{H}$ is a $q(N) \times p(N)$ i.i.d. matrix and its capacity scales as $C(N) \sim p(N)\log(1 + \rho q(N)/p(N))$ for large $\rho$ [9]. Thus, in the ideal regime the capacity scaling is given by (20) for large $\rho c$ with $c > 1$. In the beamforming regime, for any $\rho$, $\rho_{rx}(N) \to \infty$ and the capacity is given by (18) [10, Thm. 2.9, p. 623]. **Case II:** $q(N) < p(N)$. For any $N$, $\mathbf{H}$ is a $q(N)$-connected $p(N)$-dimensional channel analyzed in [5]. In the ideal regime, the capacity scales as (20) for large $\rho c$ with $c < 1$. In the multiplexing regime, for any $\rho$, $\rho_{rx}(N) \to 0$ and the capacity is given by $C(N) = p(N)E[\mathrm{trace}(\rho \mathbf{H}\mathbf{H}^H)/p(N)] = \rho D(N)/p(N)$ as in (19). **Case III:** $q(N) = p(N)$. $\mathbf{H}$ is a $p(N) \times p(N)$ i.i.d. matrix whose capacity scales as (20) for large $\rho$ with $c = 1$. □.

Theorem 1 shows the importance of the distribution of the $D(N)$ DoF in the spatial dimensions from a capacity scaling viewpoint. The number of parallel channels $p(N)$ (multiplexing gain) and the receive SNR per parallel channel $\rho_{rx}(N)$ play an important role. $\mathbf{H}_{bf}$ and $\mathbf{H}_{mux}$ represent two extremes in which the DoF are distributed to emphasize $\rho_{rx}(N)$ and $p(N)$, respectively, at the expense of the other quantity. $\mathbf{H}_{id}$ represents an ideal distribution in which $\rho_{rx}(N)$ is kept constant. $\mathbf{H}_{bf}$ morphs into $\mathbf{H}_{id}$ in the beamforming regime by increasing $\alpha$ from $\alpha_{min} \to \gamma/2$ ($p(N) \uparrow$; $\rho_{rx}(N) \downarrow$). $\mathbf{H}_{mux}$ morphs into $\mathbf{H}_{id}$ in the multiplexing regime by decreasing $\alpha$ from $\alpha_{max}$ to $\gamma/2$ ($p(N) \downarrow$; $\rho_{rx}(N) \uparrow$). In fact, $\mathbf{H}_{id}$ yields the fastest scaling.

**Corollary 1** *For any given channel power scaling law $\rho_c(N) = D(N)$, the ideal channel characterized by*

$$p_{id}(N) = \mathcal{O}\left(\sqrt{\rho_c(N)}\right),\ q_{id}(N) = \mathcal{O}\left(\sqrt{\rho_c(N)}\right) \quad (21)$$

*achieves the fastest asymptotic capacity scaling*

$$C_{id}(N) \sim \sqrt{\rho_c(N)}\log(1 + \rho c) \quad (22)$$

*For $\rho_c(N) = N^{\gamma}$, $\alpha_{id} = \frac{\gamma}{2}$; both $p_{id}$ and $q_{id}$ scale as $\mathcal{O}\left(N^{\gamma/2}\right)$.*

**Proof.** From (18) and (20) it follows that $C_{bf}(N) = o(C_{id}(N))$ for $\alpha < \alpha_{id} = \gamma/2$. Similarly from (19) and (20) it follows that $C_{mux}(N) = o(C_{id}(N))$ for $\alpha > \alpha_{id} = \gamma/2$. □

We illustrate the capacity scaling behavior in Thm. 1 in Fig. 3 for $\gamma = 1$. $C_{bf}(N)$ (for $\alpha = \alpha_{min} = 0$), $C_{id}(N)$ and $C_{mux}(N)$

[1] We assume that $p(N)$ and $q(N)$ are integers; for a given $\alpha, \delta_p$, non-integer values of $(p, q)$ can be realized by appropriate time sharing between two channels: $(\lfloor p \rfloor, \lceil q \rceil)$ and $(\lceil p \rceil, \lfloor q \rfloor)$.

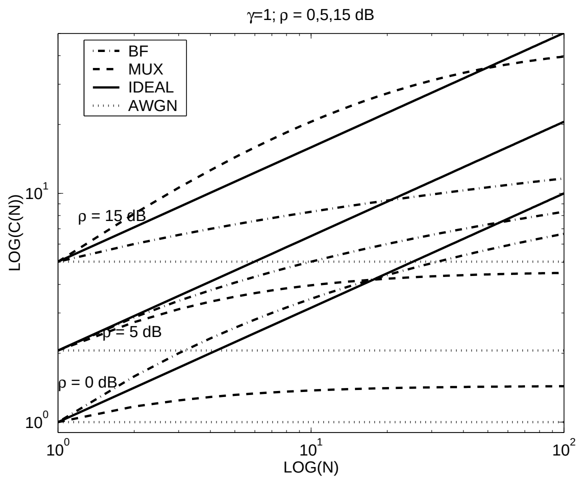


Figure 3: $C_{bf}$, $C_{id}$, and $C_{mux}$ as a function of $N$ for $\gamma = 1$. The three sets of plots correspond to $\rho = 0, 5, 15$dB.

(for $\alpha = \alpha_{max} = 1$) are plotted on a log-log scale as a function of $N$ for three SNRs: $\rho = 0, 5, 15$dB. The capacity of the AWGN channel at the corresponding SNRs is also plotted for comparison. As evident, at low SNR, $C_{mux} < C_{id} < C_{bf}$ initially but $C_{id}$ eventually dominates as $N$ gets large. On the other hand, at high SNR, $C_{bf} < C_{id} < C_{mux}$ initially but $C_{id}$ eventually dominates as $N$ gets large.

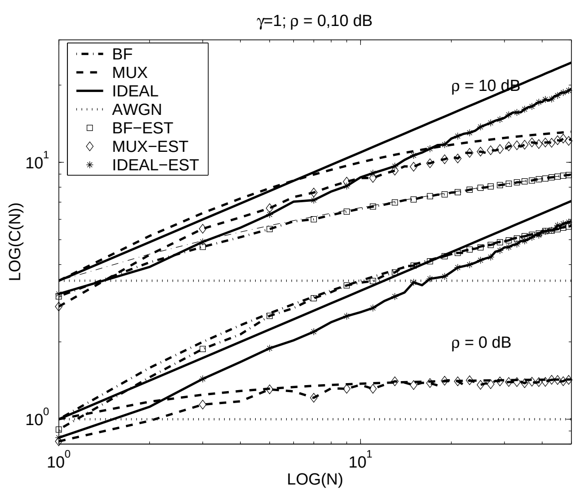


Figure 4: Theoretical and simulated values of $C_{bf}$, $C_{id}$, and $C_{mux}$ as a function of $N$ for $\gamma = 1$. $\rho = 0, 10$dB.

In Fig. 4, we compare the theoretical capacity scaling law in Thm. 1 with actual capacity estimated numerically by averaging over 200 channel realizations for every $N$. Since, $\gamma = 1$, $\mathbf{H}_{bf}$, $\mathbf{H}_{id}$ and $\mathbf{H}_{mux}$ are the same as in the motivating example. As evident, estimated $C_{id}$ exhibits the largest offset since its $\rho_{rx}$ remains fixed, whereas there is a close agreement between theoretical and estimated values of $C_{bf}$ and $C_{mux}$ for large $N$ since $\mathbf{H}_{bf}$ goes into the high SNR regime whereas $\mathbf{H}_{mux}$ goes into the low SNR regime. The figure demonstrates that the theoretical capacity scaling law quite accurately predicts the trends in actual capacity scaling.

*A. Dependence of the Ideal Channel on SNR.* It can be shown that for any given $N$, $p(N)$, $q(N)$, and $D(N)$, the capacity formula (15) is accurate as $\rho \downarrow 0$ or as $\rho \uparrow \infty$. While not accurate in the medium SNR range, it does accurately predict an important trend regarding optimal allocation of $D(N)$ DoF in $p(N)$ and $q(N)$ as a function of $\rho$ for any given $N$. First note that

$$C'(p) = \frac{dC(p)}{dp} = \log\left(1 + \frac{\rho q}{p}\right) - \frac{2}{1 + \frac{p}{\rho q}}, \qquad (23)$$

where $\rho q/p = \rho_{rx}$. It follows that for sufficiently small $\rho_{low}$, $C'(p) < 0\ ,\ \rho \le \rho_{low}$, and for sufficiently large $\rho_{high}$, $C'(p) > 0\ ,\ \rho \ge \rho_{high}$, for all $1 \le p \le N$. Thus, for $\rho \le \rho_{low}$, $p_{min} = p_{bf} = 1$ is optimal and for $\rho \ge \rho_{high}$, $p_{max} = p_{mux} = N$ is optimal. The ideal channel with $p_{id} = q_{id} = \sqrt{N}$ is robust in the sense that at low SNRs $C_{mux} < C_{id} < C_{bf}$ whereas at high SNRs $C_{bf} < C_{id} < C_{mux}$. This is illustrated in Fig. 5. More

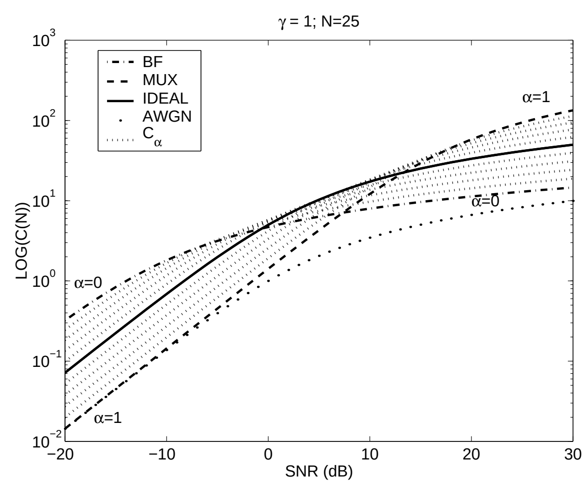


Figure 5: $C_{bf}$, $C_{id}$, $C_{mux}$ and $C_\alpha$ as a function of $\rho$ (dB) for $N = 25$ and $\gamma = 1$.

importantly, for each $\rho \in (\rho_{low}, \rho_{high})$, there exists an optimal $p_\rho$ which yields the highest capacity. This is also illustrated in Fig. 5 in which in addition to $C_{bf}$, $C_{mux}$ and $C_{ideal}$ we also plot

$$C_\alpha = N^\alpha \log\left(1 + \rho N^{\gamma - 2\alpha}\right) \qquad (24)$$

corresponding to $D(N) = N^\gamma$, $p(N) = N^\alpha$, and $q(N) = N^{\gamma-\alpha}$. $C_\alpha$ is plotted for 10 equally spaced values of $\alpha \in [0, 1]$ for $\gamma = 1$ and $N = 25$. It is evident that $\rho_{low} \approx -6$dB, $\rho_{high} \approx 16$dB and for each intermediate SNR, there is a $C_\alpha$ curve that yields the maximum capacity. This shows that as we move from low to high SNRs, the optimal distribution of channel DoFs morphs from the beamforming configuration ($\alpha = \alpha_{min}$) into the multiplexing configuration ($\alpha = \alpha_{max}$) via the ideal channel ($\alpha = \gamma/2$). It is worth noting that $C'(p) = 0$ in (23) at

$\rho q/p \approx 4$. Thus, $c_{opt} = \frac{q}{p}|_{opt} \approx 4/\rho$. The second derivative of $C(p)$

$$C''(p) = \frac{d^2C(p)}{dp^2} = \frac{2}{p}\frac{(\frac{p}{\rho q}-1)}{(\frac{p}{\rho q}+1)^2} \quad (25)$$

reveals an inflection point of $C(p)$ at $\rho_{rx} = \rho q/p = 1$: $C(p)$ is strictly concave ($C''(p) < 0$) for $\rho_{rx} < 1$ and strictly convex for $\rho_{rx} > 1$.

## IV. Creating the Ideal Channel Via Adaptive Array Configuration

In this section, we outline a methodology for realizing an ideal channel in practice for any given physical scattering environment. It requires the use of adaptive-resolution antenna array configurations for transforming a given scattering environment into an ideal MIMO channel.

Consider a given $N$ and an arbitrary scattering environment characterized by $L(N)$ paths with Rayleigh, equal power, amplitudes, whose AoA's and AoD's are uniformly and independently distributed within the spatial angular spreads: $(\phi_{r,l}, \phi_{t,l}) \sim \text{unif}\left([\phi_{r,min}, \phi_{r,max}] \times [\phi_{t,min}, \phi_{t,max}]\right)$. Suppose that number of paths grow as $L(N) \sim \mathcal{O}(D(N)) = N^{\gamma}$. To realize the ideal channel, we consider all possible array configurations with $p(N)$ transmit and $r(N) = \max(p(N), q(N))$ receive antennas that satisfy: $(p(N), q(N)) \in \{(p,n) : p(N) = N^{\alpha},\ q(N) = N^{\gamma-\alpha},\ \alpha \in [\alpha_{min}, \alpha_{max}]\}$.

Consider $\gamma = 1$ for concreteness. For given transmit and receive angular spreads there exist corresponding critical antenna spacings $d_{tx}$ and $d_{rx}$ [4] so that arrays with $p(N) = r(N) = N$ uniformly spaced transmit and receive antennas realize the finest resolution channel $\mathbf{H}_{mux}$ in the multiplexing regime with $p_{max}(N) = r_{max}(N) = N$. The finest resolution array apertures are: $A_{tx} = Nd_{tx}$ and $A_{rx} = Nd_{rx}$. Any feasible $(p(N), q(N))$ MIMO channel can be generated by using arrays that sample $A_{tx}$ and $A_{rx}$ with uniformly spaced antennas: $d_{tx}(N) = A_{tx}/p(N) = d_{tx}N/p(N)$, $d_{rx}(N) = A_{rx}/r(N) = d_{rx}N/r(N)$. The corresponding $r(N) \times p(N)$ matrix $\mathbf{H}_{p,q}$ is related to $\mathbf{H}_{mux} = \mathbf{H}_{N,1}$ as

$$\mathbf{H}_{p,q} = \mathbf{U}_r^T \mathbf{H}_{mux} \mathbf{U}_p \quad (26)$$

$$\mathbf{U}_p = \begin{bmatrix} \mathbf{1}_{N/p} & \mathbf{0}_{N/p} & \cdots & \mathbf{0}_{N/p} \\ \vdots & \mathbf{1}_{N/p} & \ddots & \vdots \\ \mathbf{0}_{N/p} & \mathbf{0}_{N/p} & \cdots & \mathbf{1}_{N/p} \end{bmatrix} \quad (27)$$

where $\mathbf{U}_p$ is an $N \times p(N)$ matrix. Note that channel power is preserved in the above transformation: $E[\text{trace}(\mathbf{H}_{p,q}\mathbf{H}_{p,q}^H)] = E[\text{trace}(\mathbf{H}_{mux}\mathbf{H}_{mux}^H)]$. In particular, the three canonical channels correspond to: $\mathbf{H}_{bf} \sim (p_{min}, q_{max})$; $\mathbf{H}_{mux} \sim (p_{max}, q_{min})$; and $\mathbf{H}_{id} \sim (p_{id}, q_{id})$.

We illustrate with a numerical example. Suppose that there are $L(N) = N = 25$ ($\gamma = 1$) uniformly distributed paths within maximum angular spreads. In this case, $\mathbf{H}_{mux} = \mathbf{H}_{25,1}$ is a 25 x 25 matrix corresponding to 25-dimensional arrays with $\lambda/2$ spacing. The $25 \times 1$ $\mathbf{H}_{bf} = \mathbf{H}_{1,25}$ and the $5 \times 5$ $\mathbf{H}_{id} = \mathbf{H}_{5,5}$ are generated via (27). Note that $\mathbf{H}_{id}$ corresponds to $5\lambda/2$ spacing. The capacities of these channels (corresponding to uniform power input) were numerically computed via 200 independent realizations of physical path gains. These capacities, along with their theoretical values (as in Fig. 5), are plotted in Fig. 6 and agree quite well. We note that the uniform-power input is not optimal for the simulated channel and we expect the agreement to be even more striking if optimal inputs are used.

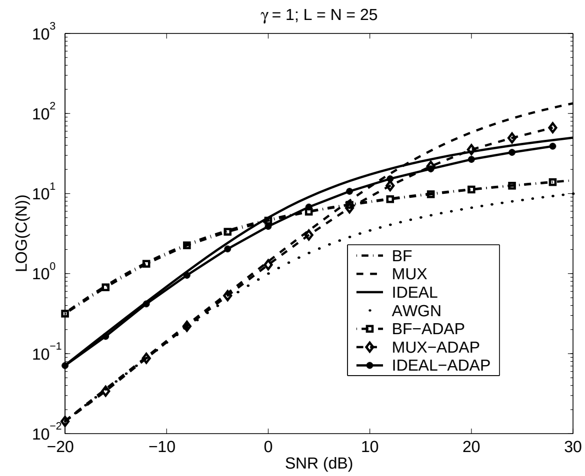


Figure 6: Theoretically ideal values of $C_{bf}$, $C_{id}$, $C_{mux}$ as a function of $\rho$, as well as simulated capacities of $\mathbf{H}_{bf}$, $\mathbf{H}_{id}$, and $\mathbf{H}_{mux}$ constructed from a random scattering environment via adaptive arrays. $L = N = 25$ and $\gamma = 1$.

## References


[1] Í. E. Telatar, "Capacity of Multi-antenna Gaussian channels," *AT& T-Bell Labs Internal Tech. Memo.*, June 1995.

[2] G. J. Foschini, M. J. Gans "On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas," *Wireless Per. Comm.*, vol. 6, no. 3, pp. 311-335, 1998.

[3] Chen-Nee Chuah, D. N. C. Tse, J. M. Kahn, R. A. Valanzuela, "Capacity Scaling in MIMO Wireless Systems under Correlated Fading," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 637-650, Mar. 2002.

[4] A. M. Sayeed, "Deconstructing Multi-antenna fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2563-2579, Oct. 2002.

[5] K. Liu, V. Raghavan, A. M. Sayeed, "Capacity Scaling and Spectral Efficiency in Wideband Correlated MIMO Channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2504-2526, Oct. 2003.

[6] A. M. Sayeed, V. V. Veeravalli, "The Essential Degrees of Freedom in Space-Time Fading Channels," *Proc. PIMRC'02, pp. 1512-1516, Portugal, Lisbon, 2002.*,

[7] V. V. Veeravalli, Y. Liang, A. M. Sayeed, "Correlated MIMO Rayleigh fading channels: Capacity, Optimal Signaling, and Scaling Laws," *submitted to the IEEE Trans. Inform. Th.*, Sept. 2003.

[8] A. M. Tulino, A. Lozano, S. Verdu, "Capacity-achieving Input Covariance Matrix for Correlated Multi-Antenna Channels," *2003 Allerton Conference for Computing, Control and Communications, pp. 242-251.*

[9] P. B. Rapajic, D. Popescu, "Information Capacity of a Random Signature Multiple-input Multiple-output Channel," *IEEE Trans. Inform. Th.*, vol. 48, no. 8, pp. 1245-1248, Aug. 2000.

[10] Z. D. Bai, "Methodologies in Spectral Analysis of Large Dimensional Random Matrices, A Review," *Stat. Sinica*, vol. 9, pp. 611-677, 1999.