HAL
open science

Capacity Limits and Multiplexing Gains of MIMO Channels with Transceiver Impairments

Emil Björnson, Per Zetterberg, Mats Bengtsson, Björn Ottersten

To cite this version:

Emil Björnson, Per Zetterberg, Mats Bengtsson, Björn Ottersten. Capacity Limits and Multiplexing Gains of MIMO Channels with Transceiver Impairments. IEEE Communications Letters, 2013, 17 (1), pp.91 - 94. hal-00923417

HAL Id: hal-00923417

https://centralesupelec.hal.science/hal-00923417v1

Submitted on 2 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Capacity Limits and Multiplexing Gains of MIMO Channels with Transceiver Impairments

Emil Björnson, Per Zetterberg, Mats Bengtsson, and Björn Ottersten

arXiv:1209.4093v1 [cs.IT] 18 Sep 2012

***Abstract*—The capacity of *ideal* MIMO channels has an high-SNR slope that equals the minimum of the number of transmit and receive antennas. This letter shows that *physical* MIMO channels behave fundamentally different, due to distortions from transceiver impairments. The capacity has a finite upper limit that holds for any channel distribution and SNR. The high-SNR slope is thus zero, but the *relative* capacity gain of employing multiple antennas is at least as large as for ideal transceivers.**

***Index Terms*—Channel capacity, high-SNR analysis, multi-antenna communication, transceiver impairments.**

I. Introduction

In the past decade, a vast number of papers have studied multiple-input multiple-output (MIMO) communications motivated by the impressive capacity scaling in the high-SNR regime. The seminal article [1] by E. Telatar shows that the MIMO capacity with channel knowledge at the receiver behaves as $\min(N_t, N_r)\log_2(P) + \mathcal{O}(1)$, where N_t and N_r are the total number of transmit and receive antennas, respectively, and P is the signal-to-noise ratio (SNR). The factor $\min(N_t, N_r)$ is the asymptotic gain over single-antenna channels and is called *degrees of freedom* or *multiplexing gain*.

Some skepticism concerning the applicability of these results in cellular networks has recently appeared; modest gains of network MIMO over conventional schemes have been observed and the throughput might even decrease due to the extra overhead [2], [3]. One explanation is the finite channel coherence time that limits the resources for channel acquisition [4] and coordination between nodes [3], thus creating a finite fundamental ceiling for the network spectral efficiency—irrespectively of the power and the number of antennas.

While these results concern large network MIMO systems, there is another non-ideality that also affects performance and manifests itself for MIMO systems of any size: *transceiver impairments* [5]–[10]. Physical radio-frequency (RF) transceivers suffer from amplifier non-linearities, IQ-imbalance, phase noise, quantization noise, carrier-frequency and sampling-rate jitter/offsets, etc. These impairments are conventionally overlooked in information theoretic studies, but this letter shows that they have a non-negligible and fundamental impact on the spectral efficiency in modern deployments with high SNR.

This letter analyzes the capacity of a generalized MIMO channel with transceiver impairments. We generalize results from [5] and [7] and emphasize the high-SNR behavior, since the capacity has a finite limit and thus is fundamentally different from the ideal case in [1]. The main conclusion is that the classic multiplexing gain is zero, but the *relative* improvement over single-antenna channels can be even larger for these physical MIMO channels than with ideal transceivers.

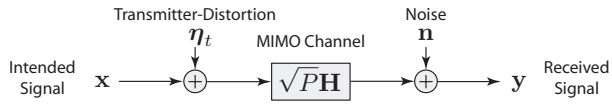

Fig. 1. Block diagram of the generalized MIMO channel considered in this letter. Unlink the classical channel model in [1], the transmitter-distortion generated by physical transceiver implementations is included in the model.

II. Generalized Channel Model

Consider a flat-fading MIMO channel with N_t transmit antennas and N_r receive antennas. The received signal $\mathbf{y} \in \mathbb{C}^{N_r}$ in the classical affine baseband channel model of [1] is

$$\mathbf{y} = \sqrt{P}\mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where P is the SNR, $\mathbf{x} \in \mathbb{C}^{N_t}$ is the intended signal, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is circular-symmetric complex Gaussian noise. The channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is assumed to be a random variable $\mathbb{H}$ having any multi-variate distribution $f_{\mathbb{H}}$ with the normalized gain $\mathbb{E}\{\mathrm{tr}(\mathbf{H}^H\mathbf{H})\} = N_t N_r$ and full-rank realizations (i.e., $\mathrm{rank}(\mathbf{H}) = \min(N_t, N_r)$) almost surely—this basically covers all physical channel distributions.

The intended signal in (1) is only affected by a multiplicative channel transformation and additive thermal noise, thus ideal transceiver hardware is implicitly assumed. Physical transceivers suffer from a variety of impairments that are not properly described by (1) [5]–[10]. The influence of impairments is reduced by compensation schemes, leaving a residual distortion that is well-modeled as additive Gaussian noise on the baseband with a variance that scales with P [7].

Building on the analysis and measurements in [6]–[8], the combined influence of impairments in the transmitter hardware is modeled by the generalized MIMO channel

$$\mathbf{y} = \sqrt{P}\mathbf{H}\left(\mathbf{x} + \boldsymbol{\eta}_t\right) + \mathbf{n}, \tag{2}$$

where the *transmitter-distortion* $\boldsymbol{\eta}_t \in \mathbb{C}^{N_t}$ describes the (residual) impairments. This term describes the mismatch between the intended signal $\mathbf{x}$ and the signal actually generated by the transmitter; see the block diagram in Fig. 1.

Under the normalized power constraint[1] $\mathrm{tr}(\mathbf{Q}) = 1$ with $\mathbf{Q} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ (similar to [1]), the transmitter-distortion is

$$\boldsymbol{\eta}_t \sim \mathcal{CN}\big(\mathbf{0}, \boldsymbol{\Upsilon}_t(\mathbf{Q})\big) \quad \text{with} \quad \boldsymbol{\Upsilon}_t = \mathrm{diag}(\upsilon_1(\mathbf{Q}), \ldots, \upsilon_{N_t}(\mathbf{Q})).$$

The distortion depends on the intended signal $\mathbf{x}$ in the sense that the variance $\upsilon_n(\mathbf{Q})$ is an increasing function of the signal power q_n at the nth transmit antenna (i.e., the nth diagonal element of $\mathbf{Q}$). We neglect any cross-correlation in $\boldsymbol{\Upsilon}_t$

[1]The power constraint is only defined on the intended signal, although distortions also contribute a small amount of power. However, this extra power is fully characterized by the SNR, P, and we therefore assume that P is selected to make the total power usage fulfill all external system constraints.

between antennas.[2] In multi-carrier (e.g., OFDM) scenarios, the channel model (2) can describe each individual subcarrier. However, there is some distortion-leakage between subcarriers and for simplicity we model this on a subcarrier basis as leakage between the antennas (the impact of what is done on the different antenna at one subcarrier is likely to average out when having many subcarriers). To capture a range of cases we propose

$$\upsilon_n(\mathbf{Q}) = \kappa^2\Big((1-\alpha)q_n + \frac{\alpha}{N_t}\Big) \tag{3}$$

where the parameter $\alpha \in [0,1]$ enables transition from one ($\alpha=0$) to many ($\alpha=1$) subcarriers. The parameter $\kappa > 0$ is the *level of impairments*.[3] This model is a good characterization of phase noise and IQ-imbalance, while amplifier non-linearities (that make $\upsilon_n(\mathbf{Q})$ increase with P) are neglected [6]—the dynamic range is assumed to always match the output power.

III. Analysis of Channel Capacity

The transmitter knows the channel distribution $f_{\mathbb{H}}$, while the receiver knows the realization $\mathbf{H}$. The capacity of (2) is

$$C_{N_t,N_r}(P) = \sup_{f_{\mathbb{X}}:\, \mathrm{tr}(\mathbb{E}\{\mathbf{x}\mathbf{x}^H\})=\mathrm{tr}(\mathbf{Q})=1} \mathcal{I}(\mathbf{x};\mathbf{y},\mathbb{H}) \tag{4}$$

where $f_{\mathbb{X}}$ is the PDF of $\mathbf{x}$ and $\mathcal{I}(\cdot,\cdot;\cdot)$ is conditional mutual information. Note that $\mathcal{I}(\mathbf{x};\mathbf{y},\mathbb{H}) = \mathbb{E}_{\mathbf{H}}\{\mathcal{I}(\mathbf{x};\mathbf{y}|\mathbb{H}=\mathbf{H})\}$.

Lemma 1. The capacity $C_{N_t,N_r}(P)$ can be expressed as

$$\sup_{\mathbf{Q}:\, \mathrm{tr}(\mathbf{Q})=1} \mathbb{E}_{\mathbf{H}}\Big\{\log_2\det\big(\mathbf{I}+P\mathbf{H}\mathbf{Q}\mathbf{H}^H(P\mathbf{H}\boldsymbol{\Upsilon}_t\mathbf{H}^H+\mathbf{I})^{-1}\big)\Big\} \tag{5}$$

and is achieved by $\mathbf{x} \sim \mathcal{CN}(\mathbf{0},\mathbf{Q})$ for some feasible $\mathbf{Q} \succeq \mathbf{0}$.

Proof: For any realization $\mathbb{H} = \mathbf{H}$ and fixed P, (2) is a classical MIMO channel but with the noise covariance $(P\mathbf{H}\boldsymbol{\Upsilon}_t\mathbf{H}^H + \mathbf{I})$. Eq. (5) and the sufficiency of using a Gaussian distribution on $\mathbf{x}$ then follows from [1]. ■

Although the capacity expression in (5) appears similar to that of the classical MIMO channel in (1) and [1], it behaves very differently—particularly in the high-SNR regime.

Theorem 1. The asymptotic capacity limit $C_{N_t,N_r}(\infty) = \lim_{P\to\infty} C_{N_t,N_r}(P)$ is finite and bounded as

$$M\log_2\left(1+\frac{1}{\kappa^2}\right) \le C_{N_t,N_r}(\infty) \le M\log_2\left(1+\frac{N_t}{M\kappa^2}\right) \tag{6}$$

where $M = \min(N_t,N_r)$. The lower bound is asymptotically achieved by $\mathbf{Q} = \frac{1}{N_t}\mathbf{I}$. The two bounds coincide if $N_t \le N_r$.

Proof: The proof is given in the appendix. ■

This theorem shows that physical MIMO systems have a finite capacity limit in the high-SNR regime—this is fundamentally different from the unbounded asymptotic capacity for ideal transceivers [1]. Furthermore, the bounds in (6) hold for any channel distribution and are only characterized by the number of antennas and the level of impairments κ.

The bounds in (6) coincide for $N_t \le N_r$, while only the upper bound grows with the number of transmit antennas when $N_t > N_r$. Informally speaking, the lower and upper bounds are tight when the high-SNR capacity-achieving $\mathbf{Q}$ is isotropic in a subspace of size N_t and size $\min(N_r,N_t)$, respectively. The following corollaries exemplify these extremes.

Corollary 1. Suppose the channel distribution is right-rotationally invariant (e.g., $\mathbb{H} \sim \mathbb{H}\mathbf{U}$ for any unitary matrix $\mathbf{U}$). The capacity is achieved by $\mathbf{Q} = \frac{1}{N_t}\mathbf{I}$ for any P and α. The lower bound in (6) is asymptotically tight for any N_t.

Proof: The right-rotational invariance implies that the N_t dimensions of $\mathbf{H}^H\mathbf{H}$ are isotropically distributed, thus the concavity of $\mathbb{E}\{\log\det(\cdot)\}$ makes an isotropic covariance matrix optimal. The lower bound in (6) is asymptotically tight as it is constructed using this isotropic covariance matrix. ■

This corollary covers Rayleigh fading channels that are uncorrelated at the transmit side, but also other channel distributions with isotropic spatial directivity at the transmitter.

The special case of a deterministic channel matrix enables stronger adaptivity of $\mathbf{Q}$ and achieves the upper bound in (6).

Corollary 2. Suppose $\alpha = 1$ and the channel $\mathbf{H}$ is deterministic and full rank. Let $\mathbf{H}^H\mathbf{H} = \mathbf{U}_M\boldsymbol{\Lambda}_M\mathbf{U}_M^H$ denote a compact eigendecomposition where $\boldsymbol{\Lambda}_M = \mathrm{diag}(\lambda_1,\ldots,\lambda_M)$ contains the non-zero eigenvalues and the semi-unitary $\mathbf{U}_M \in \mathbb{C}^{N_t\times M}$ contains the corresponding eigenvectors. The capacity is

$$C_{N_t,N_r}(P) = \sum_{i=1}^{M}\log_2\left(1+\frac{\lambda_i d_i}{\lambda_i\frac{\kappa^2}{N_t}+1}\right) \tag{7}$$

for $d_i = \big[\mu - \frac{1}{\lambda_i}\big]_+$ where μ is selected to make $\sum_{i=1}^M d_i = 1$. The capacity is achieved by $\mathbf{Q} = \mathbf{U}_M\mathrm{diag}(d_1,\ldots,d_M)\mathbf{U}_M^H$. The upper bound in (6) is asymptotically tight for any N_t.

Proof: The capacity-achieving $\mathbf{Q}$ is derived as in [1], using the Hadamard inequality. The capacity limit follows as $\mathbf{Q} = \frac{1}{M}\mathbf{U}_M\mathbf{U}_M^H$ achieves the upper bound. ■

Although the capacity behaves differently under impairments, the optimal waterfilling power allocation in Corollary 2 is the same as for ideal transceivers (also noted in [7]). When $N_t \ge N_r$, the capacity limit $M\log_2(1+\frac{N_t}{M\kappa^2})$ is improved by increasing N_t, because a deterministic $\mathbf{H}$ enables selective transmission in the N_r non-zero channel dimensions while the transmitter-distortion is isotropic over all N_t dimensions.

We conclude the analysis by elaborating on the fact that the lower bound in (6) is always asymptotically achievable.

Corollary 3. If the channel distribution $f_{\mathbb{H}}$ is unknown at the transmitter, the worst-case mutual information $\min_{f_{\mathbb{H}}}\mathcal{I}(\mathbf{x};\mathbf{y},\mathbb{H})$ is maximized by $\mathbf{Q} = \frac{1}{N_t}\mathbf{I}$ (for any α) and approaches $M\log_2\big(1+\frac{1}{\kappa^2}\big)$ as $P\to\infty$.

A. Numerical Illustrations

Consider a channel with $N_t = N_r = 4$ and varying SNR. Fig. 2 shows the average capacity over different deterministic channels, either generated synthetically with independent $\mathcal{CN}(0,1)$-entries or taken from the channel measurements in [11]. The level of impairments is varied as $\kappa \in \{0.05,\, 0.1\}$.

Ideal and physical transceivers behave similarly at low and medium SNRs in Fig. 2, but fundamentally different at high SNRs. While the ideal capacity grows unboundedly, the capacity with impairments approaches the capacity limit $C_{4,4}(\infty) =$

[2] The paper [8] predicts such a correlation, but it is typically small.

[3] The error vector magnitude, $\mathrm{EVM} = \frac{\mathbb{E}\{\|\boldsymbol{\eta}_t\|^2\}}{\mathbb{E}\{\|\mathbf{x}\|^2\}}$, is a common measure for quantifying RF transceiver impairments. Observe that the EVM equals κ^2 for the considered $\upsilon_n(\mathbf{Q})$ in (3). EVM requirements in the range $\kappa \in [0.08, 0.175]$ occur in Long Term Evolution (LTE) [9, Section 14.3.4].

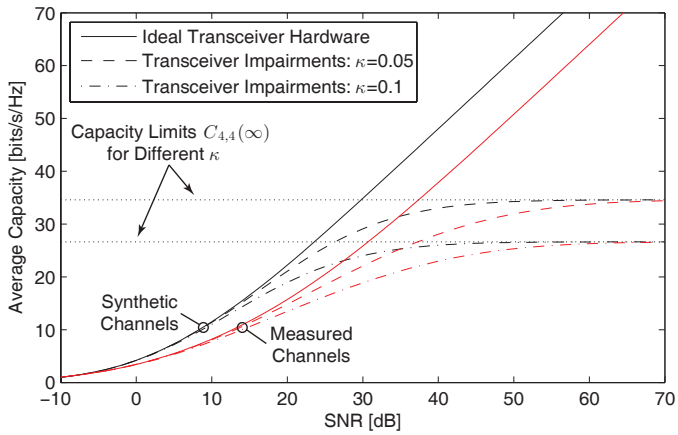

Fig. 2. Average capacity of a 4x4 MIMO channel over different deterministic channel realizations and different levels of transceiver impairments.

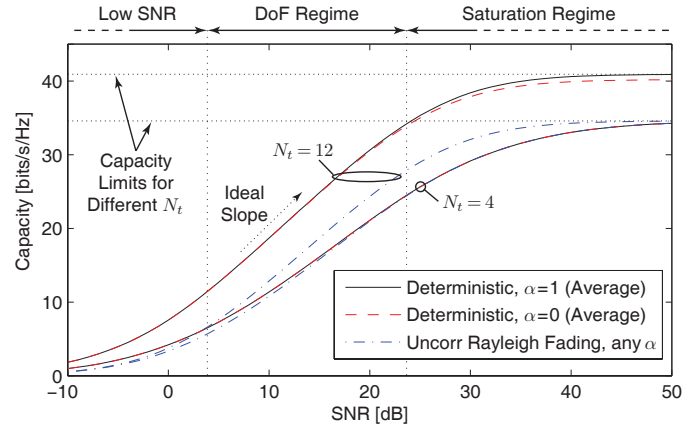

Fig. 3. Capacity of a MIMO channel with $N_r = 4$ and impairments with $\kappa = 0.05$. We consider different N_t, channel distributions, and α-values.

$4\log_2(1+\frac{1}{\kappa^2})$ in Theorem 1. The difference between the uncorrelated synthetic channels and the realistically correlated measured channels vanishes asymptotically. Therefore, only the level of impairments, κ, decides the capacity limit.

Next, we illustrate the case $N_t \geq N_r$ and different α. Fig. 3 considers $N_t \in \{4,\, 12\}$, while having $N_r = 4$, $\kappa = 0.05$, and two different channel distributions: deterministic (we average over independent $\mathcal{CN}(0,1)$-entries) and uncorrelated Rayleigh fading. We show $\alpha \in \{0,\, 1\}$ in the deterministic case, while the random case gives $\mathbf{Q} = \frac{1}{N_t}\mathbf{I}$ and same capacity for any α.

These channels perform similarly and have the same capacity limit when $N_t = 4$. The convergence to the capacity limit is improved for the random distribution when N_t increases, but the value of the limit is unchanged. Contrary, the capacity limits in the deterministic cases increase with N_t (and with α since it makes the distortion more isotropic). Fig. 3 shows that there is a medium SNR range where the capacity exhibits roughly the same M-slope as achieved asymptotically for ideal transceivers. Following the terminology of [3], this is the *degrees-of-freedom (DoF) regime* while the high-SNR regime is the *saturation regime*; see Fig. 3. This behavior appeared in [3] for large cellular networks due to limited coherence time, but we demonstrate its existence for any physical MIMO channel (regardless of size) due to transceiver impairments.

IV. Gain of Multiplexing

The MIMO capacity with ideal transceivers behaves as $M\log_2(P)+\mathcal{O}(1)$ [1], thus it grows unboundedly in the high-SNR regime and scales linearly with the so-called multiplexing gain $M = \min(N_t, N_r)$. On the contrary, Theorem 1 shows that the capacity of physical MIMO channels has a finite upper bound, giving a very different multiplexing gain:

$$\mathcal{M}_{\infty}^{\text{classic}} = \lim_{P\to\infty} \frac{C_{N_t,N_r}(P)}{\log_2(P)} = 0. \tag{8}$$

In view of (8), one might think that the existence of a non-zero multiplexing gain is merely an artifact of ignoring the transceiver impairments that always appear in practice. However, the problem lies in the classical definition, because also physical systems can gain in capacity from employing multiple antennas and utilizing spatial multiplexing. A practically more relevant measure is the relative capacity improvement (at a finite P) of an $N_t \times N_r$ MIMO channel over the corresponding single-input single-output (SISO) channel.

Definition 1. The *finite-SNR multiplexing gain*, $\mathcal{M}(P)$, is the ratio of MIMO to SISO capacity at a given P. For (2) we get

$$\mathcal{M}(P) = \frac{C_{N_t,N_r}(P)}{C_{1,1}(P)}. \tag{9}$$

This ratio between the MIMO and SISO capacity quantifies the exact gain of multiplexing. The concept of a finite-SNR multiplexing gain was introduced in [12] for ideal transceivers, while the refined Definition 1 can be applied to any channel model. The asymptotic behavior of $\mathcal{M}(P)$ is as follows.

Theorem 2. Let h denote the SISO channel. The finite-SNR multiplexing gain, $\mathcal{M}(P)$, for (2) and any α satisfies

$$\frac{\mathbb{E}\{\|\mathbf{H}\|_F^2\}}{N_t\,\mathbb{E}\{|h|^2\}} \leq \lim_{P\to 0}\mathcal{M}(P) \leq \frac{\mathbb{E}\{\|\mathbf{H}\|_2^2\}}{\mathbb{E}\{|h|^2\}}, \tag{10}$$

$$M \leq \lim_{P\to\infty}\mathcal{M}(P) \leq M\frac{\log_2(1+\frac{N_t}{M\kappa^2})}{\log_2(1+\frac{1}{\kappa^2})}, \tag{11}$$

where $\|\cdot\|_F$ and $\|\cdot\|_2$ denote the Frobenius and spectral norm, respectively. The upper bounds are achieved for deterministic channels (with full rank and $\alpha = 1$). The lower bounds are achieved for right-rotationally invariant channel distributions.

Proof: The low-SNR behavior is achieved by Taylor approximation: $\mathbf{Q} = \frac{1}{N_t}\mathbf{I}$ gives the lower bound, while the per-realization-optimal $\mathbf{Q} = \mathbf{u}\mathbf{u}^H$ (where $\mathbf{u}$ is the dominating eigenvector of $\mathbf{H}^H\mathbf{H}$) gives the upper bound. The high-SNR behavior follows from Theorem 1 and its corollaries. ■

This theorem indicates that transceiver impairments have little impact on the relative MIMO gain, which is a very positive result for practical applications. The low-SNR behavior in (10) is the same as for ideal transceivers (since $P\mathbf{H}\boldsymbol{\Upsilon}_t\mathbf{H}^H + \mathbf{I} \approx \mathbf{I}$), while (11) shows that physical MIMO channels can achieve $\mathcal{M}(P) > M$ in the high-SNR regime (although ideal transceivers only can achieve $\mathcal{M}(P) = M$).

A. Numerical Illustrations

The finite-SNR multiplexing gain is shown in Figs. 4 and 5 for uncorrelated Rayleigh fading and deterministic channels, respectively, with $N_t \in \{4,\, 8,\, 12\}$, $N_r = 4$, $\kappa = 0.05$, $\alpha = 1$.

The limits in Theorem 2 are confirmed by the simulations. Although the capacity behavior is fundamentally different for

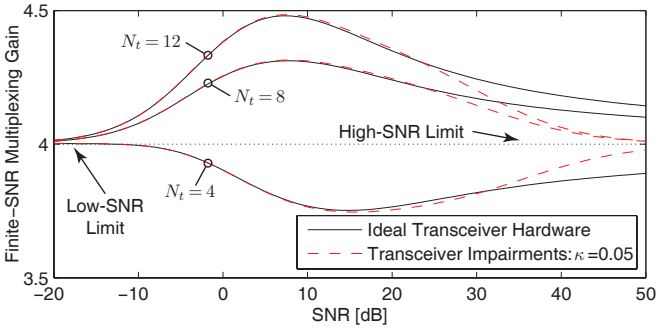

Fig. 4. Finite-SNR multiplexing gain for an uncorrelated Rayleigh fading channel with $N_r = 4$ and $N_t \geq 4$.

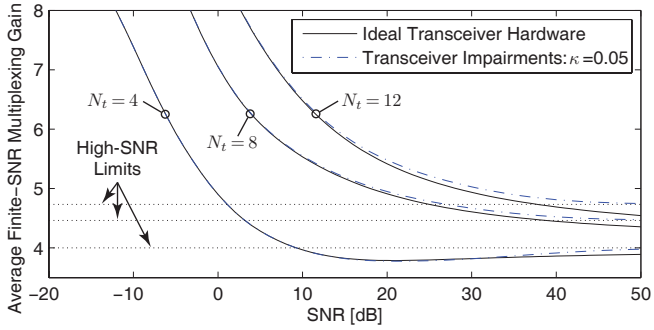

Fig. 5. Average finite-SNR multiplexing gain of deterministic channels (generated with independent $\mathcal{CN}(0,1)$-entries) with $N_r = 4$ and $N_t \geq 4$.

physical and ideal transceivers, the finite-SNR multiplexing gain is remarkably similar—not unexpected since the asymptotic limits in Theorem 2 are almost the same for any level of transceiver impairments. The main difference is in the high-SNR regime, where (a) there is a faster convergence to the limits under impairments and (b) deterministic channels achieve an asymptotic gain higher than M when $N_t > N_r$.

V. Concluding Implications

Unlike conventional capacity analysis, the capacity of physical MIMO systems saturates in the high-SNR regime (see Theorem 1) and the finite capacity limit is channel independent. This fundamental result is explained by the distortion from transceiver impairments and that its power is proportional to the signal power. The classic multiplexing gain is thus zero (see Eq. (8)). Nevertheless, the MIMO capacity grows roughly linearly with $M = \min(N_t, N_r)$ (see Theorem 2) over the whole SNR range, thus showing that also physical systems can achieve great gains from spatial multiplexing.

Technological advances can reduce transceiver impairments, but there is currently an opposite trend towards small low-cost low-power transceivers where the inherent *dirty RF effects* are inevitable and the transmission is instead adapted to them.

The point-to-point MIMO capacity limit in Theorem 1 is also an upper bound for scenarios with extra constraints; for example, network MIMO which is characterized by distributed power constraints and limited coordination both between transmit antennas and between receive antennas. The capacity in such scenarios therefore saturates in the high-SNR regime—even in small networks where the analysis in [3] is not applicable. Note that the finite-SNR multiplexing gain decreases when adding extra constraints [10] and that impairments limit the asymptotic accuracy of channel acquisition schemes.

Appendix: Proof of Theorem 1

As a preliminary, consider any full-rank channel realization $\mathbf{H}$. Let $\mathbf{H}^H\mathbf{H} = \mathbf{U}_M\mathbf{\Lambda}_M\mathbf{U}_M^H$ denote a *compact* eigendecomposition (with $\mathbf{U}_M \in \mathbb{C}^{N_t \times M}$, $\mathbf{\Lambda}_M \in \mathbb{C}^{M \times M}$; see Corollary 2). The mutual information is increasing in P and satisfies

$$\begin{aligned}
&\log_2 \det\big(\mathbf{I}+P\mathbf{H}\mathbf{Q}\mathbf{H}^H(P\mathbf{H}\mathbf{\Upsilon}_t\mathbf{H}^H+\mathbf{I})^{-1}\big) \\
&\quad = \log_2 \det\big(\mathbf{I}+P\mathbf{U}_M^H(\mathbf{Q}+\mathbf{\Upsilon}_t)\mathbf{U}_M\mathbf{\Lambda}_M\big) \\
&\qquad - \log_2 \det\big(\mathbf{I}+P\mathbf{U}_M^H\mathbf{\Upsilon}_t\mathbf{U}_M\mathbf{\Lambda}_M\big) \quad \rightarrow \\
&\log_2 \det\big(\mathbf{U}_M^H(\mathbf{Q}+\mathbf{\Upsilon}_t)\mathbf{U}_M\mathbf{\Lambda}_M\big) - \log_2 \det\big(\mathbf{U}_M^H\mathbf{\Upsilon}_t\mathbf{U}_M\mathbf{\Lambda}_M\big) \\
&\quad = \log_2 \det\big(\mathbf{I}+\mathbf{U}_M^H\mathbf{Q}\mathbf{U}_M(\mathbf{U}_M^H\mathbf{\Upsilon}_t\mathbf{U}_M)^{-1}\big) \qquad (12) \\
&\quad = \log_2 \det\Big(\mathbf{I}+\mathbf{\Upsilon}_t^{-1/2}\mathbf{Q}\mathbf{\Upsilon}_t^{-H/2}\mathbf{\Pi}_{\mathbf{\Upsilon}_t^{H/2}\mathbf{U}_M}\Big) \\
&\quad = \sum_{i=1}^{M} \log_2\Big(1+\mu_i(\mathbf{\Upsilon}_t^{-1/2}\mathbf{Q}\mathbf{\Upsilon}_t^{-H/2}\mathbf{\Pi}_{\mathbf{\Upsilon}_t^{H/2}\mathbf{U}_M})\Big) \qquad (13)
\end{aligned}$$

as $P \rightarrow \infty$. The first equality follows from expanding the logarithm and using the rule $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$. This enables taking the limit $P \rightarrow \infty$ and achieve an expression where the impact of $\mathbf{\Lambda}_M$ cancels out. We then identify the projection matrix $\mathbf{\Pi}_{\mathbf{\Upsilon}_t^{H/2}\mathbf{U}_M} = \mathbf{\Upsilon}_t^{H/2}\mathbf{U}_M(\mathbf{U}_M^H\mathbf{\Upsilon}_t\mathbf{U}_M)^{-1}\mathbf{U}_M^H\mathbf{\Upsilon}_t^{1/2}$ onto $\mathbf{U}_M^H\mathbf{\Upsilon}_t^{1/2}$. The ith strongest eigenvalue is denoted $\mu_i(\cdot)$.

The convergence as $P \rightarrow \infty$ is uniform, thus we can achieve bounds by showing that all realizations has the same asymptotic bound. A lower bound is given by any feasible $\mathbf{Q}$; we select $\mathbf{Q} = \frac{1}{N_t}\mathbf{I}$ as it gives $\mathbf{\Upsilon}_t = \frac{\kappa^2}{N_t}\mathbf{I}$ and makes (12) independent of $\mathbf{H}$. Since (13) is a Schur-concave function in the eigenvalues, an upper bound is achieved by replacing $\mu_i(\cdot)$ with the average eigenvalue $\frac{1}{M}\mathrm{tr}(\mathbf{\Upsilon}_t^{-1/2}\mathbf{Q}\mathbf{\Upsilon}_t^{-H/2}\mathbf{\Pi}_{\mathbf{\Upsilon}_t^{H/2}\mathbf{U}_M}) \leq \frac{1}{M}\mathrm{tr}(\mathbf{\Upsilon}_t^{-1/2}\mathbf{Q}\mathbf{\Upsilon}_t^{-H/2}) = \frac{N_t\kappa^2}{M}$, where the inequality follows from removing the projection matrix (since $\mathbf{\Pi}_{\mathbf{\Upsilon}_t^{H/2}\mathbf{U}_M} \preceq \mathbf{I}$). Note that the upper and lower bounds coincide when $N_t \leq N_r$, thus $\mathbf{Q} = \frac{1}{M}\mathbf{I}$ is asymptotically optimal in this case.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecom.*, vol. 10, no. 6, pp. 585–595, 1999.

[2] A. Barbieri, P. Gaal, S. Geirhofer, T. Ji, D. Malladi, Y. Wei, and F. Xue, "Coordinated downlink multi-point communications in heterogeneous cellular networks," in *Proc. ITA*, 2012.

[3] A. Lozano, R. Heath, and J. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, submitted, arXiv:1204.0011.

[4] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.

[5] T. Koch, A. Lapidoth, and P. Sotiriadis, "Channels that heat up," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3594–3612, 2009.

[6] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Springer, 2008.

[7] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *Proc. ITG/IEEE WSA*, 2010.

[8] N. Moghadam, P. Zetterberg, P. Händel, and H. Hjalmarsson, "Correlation of distortion noise between the branches of MIMO transmit antennas," in *Proc. IEEE PIMRC*, 2012.

[9] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, 2nd ed. Wiley, 2011.

[10] E. Björnson, P. Zetterberg, and M. Bengtsson, "Optimal coordinated beamforming in the multicell downlink with transceiver impairments," in *Proc. IEEE GLOBECOM*, 2012.

[11] N. Jaldén, P. Zetterberg, B. Ottersten, and L. Garcia, "Inter- and intrasite correlations of large-scale parameters from macrocellular measurements at 1800 MHz," *EURASIP J. Wirel. Commun. Netw.*, 2007.

[12] R. Narasimhan, "Finite-SNR diversity performance of rate-adaptive MIMO systems," in *Proc. IEEE GLOBECOM*, 2005.