



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *IEEE Communications Letters*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Björnson, E., Zetterberg, P., Bengtsson, M., Ottersten, B. (2013)
Capacity Limits and Multiplexing Gains of MIMO Channels with Transceiver Impairments.
IEEE Communications Letters, 17(1): 91-94
<http://dx.doi.org/10.1109/LCOMM.2012.112012.122003>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-117611>

Capacity Limits and Multiplexing Gains of MIMO Channels with Transceiver Impairments

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

Abstract—The capacity of *ideal* MIMO channels has a high-SNR slope that equals the minimum of the number of transmit and receive antennas. This letter analyzes if this result holds when there are distortions from physical transceiver impairments. We prove analytically that such *physical* MIMO channels have a finite upper capacity limit, for any channel distribution and SNR. The high-SNR slope thus collapses to zero. This appears discouraging, but we prove the encouraging result that the *relative* capacity gain of employing MIMO is at least as large as with 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 $M \log_2(\text{SNR}) + \mathcal{O}(1)$, where SNR is the signal-to-noise ratio (SNR). The slope M satisfies $M = \min(N_t, N_r)$, where N_t and N_r are the number of transmit and receive antennas, respectively. M 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

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Manuscript received September 5, 2012. The associate editor coordinating the review of this letter and approving it for publication was D.-A. Toumpakaris.

This work was supported by International Postdoc Grant 2012-228 from The Swedish Research Council and by the HIATUS project (FET 265578).

The authors are with the Signal Processing Lab, ACCESS Linnaeus Center, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden. E. Björnson is also the Alcatel-Lucent Chair on Flexible Radio, SUPELEC 91192 Gif-sur-Yvette, France (e-mail: emil.bjornson@ee.kth.se). B. Ottersten is also with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, L-1359 Luxembourg-Kirchberg, Luxembourg. Digital Object Identifier 10.1109/LCOMM.2012.112012.122003



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

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 generalized MIMO channel with transceiver impairments from [7]. We show that the capacity has a finite high-SNR limit for any channel distribution. The multiplexing gain is thus zero, which is fundamentally different from the ideal case in [1] (detailed above). Similar single-antenna results are given in [5]. The practical MIMO gain—the *relative* capacity increase over single-antenna channels—is however shown to be at least as large as with ideal transceivers.

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

$$\mathbf{y} = \sqrt{\text{SNR}} \mathbf{H} \mathbf{x} + \mathbf{n}, \quad (1)$$

where SNR 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 normalized gain $\mathbb{E}\{\text{tr}(\mathbf{H}^H \mathbf{H})\} = N_t N_r$ and full-rank realizations (i.e., $\text{rank}(\mathbf{H}) = \min(N_t, N_r)$) almost surely—this basically covers all physical channel distributions.

The intended signal \mathbf{x} 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 with a variance that scales with SNR [7].

A generalized MIMO channel is proposed in [6], [7] and verified by measurements. The combined (residual) influence

of impairments in the transmitter hardware is modeled by the *transmitter distortion* $\boldsymbol{\eta}_t \in \mathbb{C}^{N_t}$ and (1) is generalized to

$$\mathbf{y} = \sqrt{\text{SNR}} \mathbf{H} (\mathbf{x} + \boldsymbol{\eta}_t) + \mathbf{n}. \quad (2)$$

Note that $\boldsymbol{\eta}_t$ is the mismatch between the intended signal \mathbf{x} and the signal actually radiated by the transmitter; see Fig. 1. It is well-modeled as uncorrelated Gaussian noise as it is the aggregate residual of many impairments, whereof some are Gaussian and some behave as Gaussian when summed up [7].

Under the normalized power constraint¹ $\text{tr}(\mathbf{Q}) = 1$ with $\mathbf{Q} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ (similar to [1]), the transmitter distortion is

$$\boldsymbol{\eta}_t \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Upsilon}_t(\mathbf{Q})) \quad \text{with} \quad \boldsymbol{\Upsilon}_t = \text{diag}(v_1(q_1), \dots, v_{N_t}(q_{N_t})).$$

The distortion depends on the intended signal \mathbf{x} in the sense that the variance $v_n(q_n)$ is an increasing function of the signal power q_n at the n th transmit antenna (i.e., the n th diagonal element of \mathbf{Q}). We neglect any antenna cross-correlation in $\boldsymbol{\Upsilon}_t$.² In multi-carrier (e.g., OFDM) scenarios, (2) can describe each individual subcarrier. However, there is some distortion leakage between subcarriers that makes q_n less influential on $v_n(q_n)$. For simplicity, we model the leakage as proportional to the average signal power per antenna (i.e., the direct impact of what is done on individual antennas/subcarriers averages out when having many subcarriers). To capture a range of cases we propose

$$v_n(q_n) = \kappa^2 \left((1-\alpha)q_n + \alpha \frac{\sum_{i=1}^{N_t} q_i}{N_t} \right), \quad (3)$$

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 the impact of amplifier non-linearities grows non-linearly in SNR [6]. We assume to operate in the dynamic range where the impact is almost linear.

III. ANALYSIS OF CHANNEL CAPACITY

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

$$C_{N_t, N_r}(\text{SNR}) = \sup_{\mathbf{f}_{\mathbf{x}}: \text{tr}(\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}) = \text{tr}(\mathbf{Q}) = 1} \mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbb{H}) \quad (4)$$

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

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

$$\sup_{\mathbf{Q}: \text{tr}(\mathbf{Q}) = 1} \mathbb{E}_{\mathbf{H}} \left\{ \log_2 \det \left(\mathbf{I} + \text{SNR} \mathbf{H} \mathbf{Q} \mathbf{H}^H (\text{SNR} \boldsymbol{\Upsilon}_t \mathbf{H}^H + \mathbf{I})^{-1} \right) \right\}$$

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

¹The 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 and we therefore assume that SNR is selected to make the total power usage fulfill all external system constraints.

²Correlation between antennas is predicted in [8], but it is typically small.

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

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

Although the capacity expression in Lemma 1 looks 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 $C_{N_t, N_r}(\infty) = \lim_{\text{SNR} \rightarrow \infty} C_{N_t, N_r}(\text{SNR})$ is finite and bounded as

$$M \log_2 \left(1 + \frac{1}{\kappa^2} \right) \leq C_{N_t, N_r}(\infty) \leq M \log_2 \left(1 + \frac{N_t}{M \kappa^2} \right) \quad (5)$$

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 $N_t \leq N_r$.

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 (5) hold for any channel distribution and are only characterized by the number of antennas and the level of impairments κ .

The bounds in (5) coincide for $N_t \leq N_r$, 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 SNR and α . The lower bound in (5) 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 (5) 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 (5).

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 $\boldsymbol{\Lambda}_M = \text{diag}(\lambda_1, \dots, \lambda_M)$ 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

$$C_{N_t, N_r}(\text{SNR}) = \sum_{i=1}^M \log_2 \left(1 + \frac{\text{SNR} \lambda_i d_i}{\text{SNR} \lambda_i \frac{\kappa^2}{N_t} + 1} \right) \quad (6)$$

for $d_i = \left[\mu - \frac{1}{\lambda_i} \right]_+$ where μ is selected to make $\sum_{i=1}^M d_i = 1$. The capacity is achieved by $\mathbf{Q} = \mathbf{U}_M \text{diag}(d_1, \dots, d_M) \mathbf{U}_M^H$. The upper bound in (5) 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 since $\mathbf{Q} = \frac{1}{M} \mathbf{U}_M \mathbf{U}_M^H$ achieves the upper bound in (5). ■

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 $N_t \geq N_r$, 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 (5) is always asymptotically achievable.

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

A. Numerical Illustrations

Consider a channel with $N_t = N_r = 4$, $\alpha = 1$, 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 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 $C_{4,4}(\infty) = 4 \log_2(1 + \frac{1}{\kappa^2})$ 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 $N_t \in \{4, 12\}$ with $N_r = 4$, $\kappa = 0.05$, and two different channel distributions: deterministic (average capacity with known i.i.d. $\mathcal{CN}(0, 1)$ -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 becomes faster 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 $M \log_2(\text{SNR}) + \mathcal{O}(1)$ [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

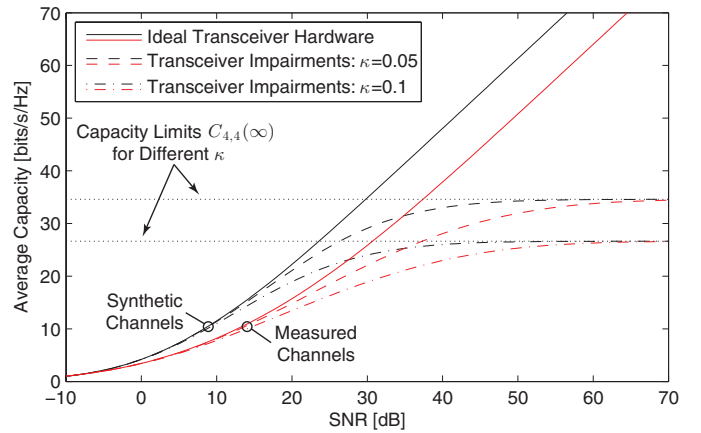


Fig. 2. Average capacity of a 4x4 MIMO channel over different deterministic channel realizations, different levels of transceiver impairments, and $\alpha = 1$.

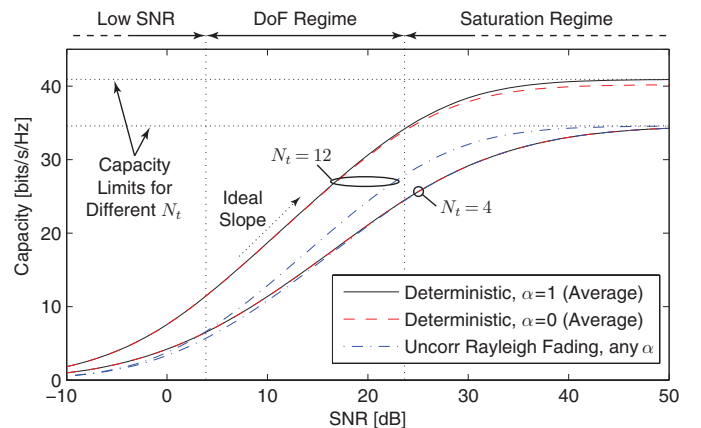


Fig. 3. Capacity of a MIMO channel with $N_r = 4$ and impairments with $\kappa = 0.05$. We consider different N_t , channel distributions, and α -values.

finite upper bound, giving a very different multiplexing gain:

$$\mathcal{M}_{\infty}^{\text{classic}} = \lim_{\text{SNR} \rightarrow \infty} \frac{C_{N_t, N_r}(\text{SNR})}{\log_2(\text{SNR})} = 0. \quad (7)$$

In view of (7), 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 SNR) 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*, $\mathcal{M}(\text{SNR})$, is the ratio of MIMO to SISO capacity at a given SNR. For (2),

$$\mathcal{M}(\text{SNR}) = \frac{C_{N_t, N_r}(\text{SNR})}{C_{1,1}(\text{SNR})}. \quad (8)$$

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 $\mathcal{M}(\text{SNR})$ is as follows.

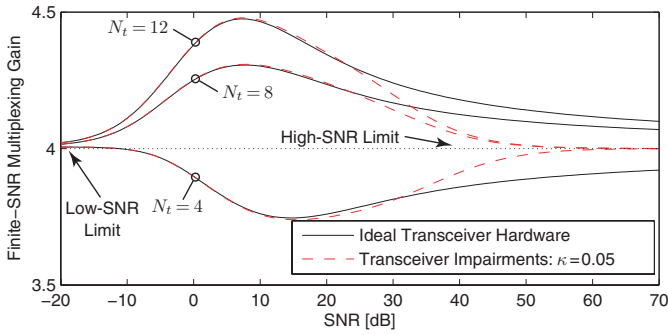


Fig. 4. Finite-SNR multiplexing gain for an uncorrelated Rayleigh fading channel with $N_r = 4$ and $N_t \geq 4$.

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

$$\frac{\mathbb{E}\{\|\mathbf{H}\|_F^2\}}{N_t \mathbb{E}\{|h|^2\}} \leq \lim_{\text{SNR} \rightarrow 0} \mathcal{M}(\text{SNR}) \leq \frac{\mathbb{E}\{\|\mathbf{H}\|_2^2\}}{\mathbb{E}\{|h|^2\}}, \quad (9)$$

$$M \leq \lim_{\text{SNR} \rightarrow \infty} \mathcal{M}(\text{SNR}) \leq M \frac{\log_2(1 + \frac{N_t}{M\kappa^2})}{\log_2(1 + \frac{1}{\kappa^2})}, \quad (10)$$

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 (9) is the same as for ideal transceivers (since $\text{SNR}\mathbf{H}\mathbf{Y}_t\mathbf{H}^H + \mathbf{I} \approx \mathbf{I}$), while (10) shows that physical MIMO channels can achieve $\mathcal{M}(\text{SNR}) > M$ in the high-SNR regime (although ideal transceivers only can achieve $\mathcal{M}(\text{SNR}) = M$).

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 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 REMARKS

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 independent of the

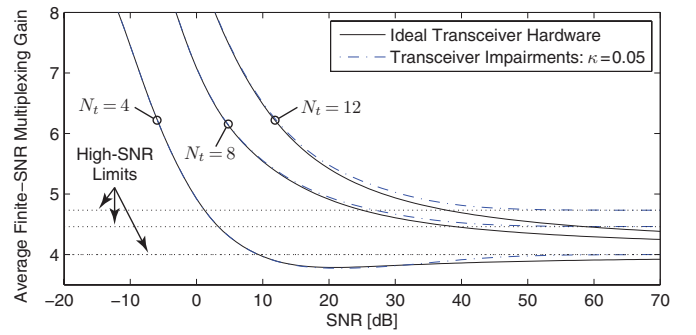


Fig. 5. Average finite-SNR multiplexing gain of deterministic channels (generated with independent $\mathcal{CN}(0, 1)$ -entries) with $N_r = 4$ and $N_t \geq 4$.

channel distribution. 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. (7)). Nevertheless, the MIMO capacity grows roughly linearly with $M = \min(N_t, N_r)$ (see Theorem 2) over the whole SNR range, thus showing the encouraging result that also physical systems can achieve great gains from employing MIMO and 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 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.

Finally, 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 increases with SNR (since it reduces the noise term and $\log_2 \det(\cdot)$ is concave) and satisfies

$$\begin{aligned} & \log_2 \det(\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{Q}\mathbf{H}^H (\text{SNR}\mathbf{H}\mathbf{Y}_t\mathbf{H}^H + \mathbf{I})^{-1}) \\ &= \log_2 \det(\mathbf{I} + \text{SNR}\mathbf{U}_M^H(\mathbf{Q} + \mathbf{Y}_t)\mathbf{U}_M\mathbf{\Lambda}_M) \\ & \quad - \log_2 \det(\mathbf{I} + \text{SNR}\mathbf{U}_M^H\mathbf{Y}_t\mathbf{U}_M\mathbf{\Lambda}_M) \rightarrow \\ & \log_2 \det(\mathbf{U}_M^H(\mathbf{Q} + \mathbf{Y}_t)\mathbf{U}_M\mathbf{\Lambda}_M) - \log_2 \det(\mathbf{U}_M^H\mathbf{Y}_t\mathbf{U}_M\mathbf{\Lambda}_M) \\ &= \log_2 \det(\mathbf{I} + \mathbf{U}_M^H\mathbf{Q}\mathbf{U}_M(\mathbf{U}_M^H\mathbf{Y}_t\mathbf{U}_M)^{-1}) \quad (11) \\ &= \log_2 \det(\mathbf{I} + \mathbf{Y}_t^{-1/2}\mathbf{Q}\mathbf{Y}_t^{-H/2}\mathbf{\Pi}_{\mathbf{Y}_t^{H/2}}\mathbf{U}_M) \\ &= \sum_{i=1}^M \log_2(1 + \mu_i(\mathbf{Y}_t^{-1/2}\mathbf{Q}\mathbf{Y}_t^{-H/2}\mathbf{\Pi}_{\mathbf{Y}_t^{H/2}}\mathbf{U}_M)) \quad (12) \end{aligned}$$

as $\text{SNR} \rightarrow \infty$. The first equality follows from expanding the logarithm and from the rule $\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$. This enables letting $\text{SNR} \rightarrow \infty$ and achieve an expression where

the impact of Λ_M cancels out. We then identify the projection matrix $\Pi_{\Upsilon_t^{H/2}\mathbf{U}_M} = \Upsilon_t^{H/2}\mathbf{U}_M(\mathbf{U}_M^H\Upsilon_t\mathbf{U}_M)^{-1}\mathbf{U}_M^H\Upsilon_t^{1/2}$ onto $\mathbf{U}_M^H\Upsilon_t^{1/2}$. The i th strongest eigenvalue is denoted $\mu_i(\cdot)$.

As the convergence $\text{SNR} \rightarrow \infty$ is uniform, we can achieve bounds by showing that all realizations have 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 $\Upsilon_t = \frac{\kappa^2}{N_t}\mathbf{I}$ and makes (11) independent of \mathbf{H} . Since (12) is a Schur-concave function in the eigenvalues, an upper bound is achieved by replacing $\mu_i(\cdot)$ with the average eigenvalue $\frac{1}{M}\text{tr}(\Upsilon_t^{-1/2}\mathbf{Q}\Upsilon_t^{-H/2}\Pi_{\Upsilon_t^{H/2}\mathbf{U}_M}) \leq \frac{1}{M}\text{tr}(\Upsilon_t^{-1/2}\mathbf{Q}\Upsilon_t^{-H/2}) = \frac{N_t\kappa^2}{M}$, where the inequality follows from removing the projection matrix (since $\Pi_{\Upsilon_t^{H/2}\mathbf{U}_M} \preceq \mathbf{I}$). 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-multiplexing tradeoff for correlated Rayleigh and Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3965–3979, 2006.