Fundamental Rate-Reliability-Complexity Limits in Outage Limited MIMO Communications

Petros Elia and Joakim Jaldén

Abstract—**The work establishes fundamental limits between rate, reliability and computational complexity, for the general setting of outage-limited MIMO communications. In the high-SNR regime, the limits are optimized over all encoders, all decoders, and all complexity regulating policies. The work then proceeds to explicitly identify encoder-decoder designs and policies, that meet this optimal tradeoff. In practice, the limits aim to meaningfully quantify different pertinent and interrelated measures, such as the optimal rate-reliability capabilities per unit complexity and power, the optimal diversity gains per complexity costs, or the optimal goodput per flop. Finally the tradeoff's simple nature, renders it useful for insightful comparison of the rate-reliability-complexity capabilities for different encoders-decoders.**

Index Terms—**Diversity-multiplexing tradeoff, complexity, multiple-input multiple-output (MIMO), space-time coders-decoders, fundamental limits, lattice reduction, regularization.**

I. Introduction

A. General system model

We consider the general multiple-input multiple-output (MIMO) communications setting, where the $m \times 1$ vector representation of the received signal $\boldsymbol{y}$ is given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{w}, \tag{1}$$

where $\boldsymbol{x}$ is the $n \times 1$ vector representation of the coded transmitted signals, $\boldsymbol{H}$ the $m \times n$ channel matrix, and where $\boldsymbol{w}$ represents additive noise. $\boldsymbol{H}$ is considered to be random, having an arbitrary distribution, and being parameterized by ρ which is interpreted as the SNR (cf. [1]). $\boldsymbol{w}$ is taken to be i.i.d. Gaussian with fixed variance. We assume that one use of (1) corresponds to T uses of some underlying "physical" channel.

The model applies to several network topologies and scenarios, such as MIMO, MIMO-OFDM, MIMO-MAC, MIMO-ARQ, and cooperative communications, and each such scenario endows $\boldsymbol{H}$ and $\boldsymbol{x}$ with different structures, dimensionalities and statistics. This work specifically considers the non-ergodic, outage-limited setting, in which the above MIMO-related scenarios play a crucial role in improving the error and rate performance, though usually at the expense of much higher encoding-decoding computational complexity.

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 228044. P. Elia acknowledges funding by EU:FP7/2007-2013 grant no. 216076 (SENDORA), by NEWCOM++ (contract n. 216715), and by Mitsubishi RD project Home-eNodeBS. The work of J. Jaldén was supported by the SSF grant ICA08-0046.

P. Elia is with the Mobile Communications Department, EURECOM, Sophia Antipolis, France (email: elia@eurecom.fr)

J. Jaldén is with the ACCESS Linnaeus Center, KTH Signal Processing Lab, Royal Institute of Technology, Stockholm, Sweden (email:jalden@kth.se)

B. Motivation and general results

Error performance and encoding-decoding complexity in telecommunications (cf. [2]–[6]), are widely considered to be two limiting, and interrelated bottlenecks. Joint exposition of these two aspects becomes increasingly necessary, in order to meaningfully quantify the ever increasing complexity costs of reliable communication, in systems that progressively become larger and more dynamic.

A natural question then pertains to establishing and meeting joint fundamental error-performance and complexity limits, optimized over all choices of encoders, decoders and policies. Such limits will be here described, under a high SNR approximation, in the form of an optimal rate-reliability-complexity tradeoff for MIMO communications.

The limits provide answers, within approximation factors corresponding to the high-SNR asymptotics[1], pertaining to the following.

- Description, given ρ, of the best possible achievable rate-reliability-complexity combination, optimized over all transceivers and policies (Theorem 1).
- Description, given ρ, of the union of all achievable rate-reliability-complexity combinations. (Theorem 2).
- Description of the optimal value achieved by a large family of utility measures which quantify the rate-reliability-complexity capabilities of transceivers, and which are decreasing functions of complexity and of error probability. (Theorem 3).

C. Structure of paper

Section II recalls the general transceiver setting, and defines the different performance measures. Section III introduces the asymptotic measures of performance, directly applying the *diversity multiplexing tradeoff* (DMT, [2]) as an asymptotic measure of rate-error performance, and defining the *worst-case complexity exponent* as an asymptotic measure of worst-case complexity. Section IV presents the high-SNR optimal rate-reliability-complexity tradeoff, and the optimal transceiver utility value in its general form, as well as in its simpler, more specific, *homogeneous* variant. Finally Section V concludes.

II. Transceiver design and decoding policy: rate, reliability and complexity

A. Transceiver design and decoding policy

Consider a sequence of transceiver designs $\mathcal{X}_\rho, \mathcal{D}_\rho$, parameterized by ρ, where $\mathcal{X}_\rho \subset \mathbb{R}^n$ denotes the codebook that maps

[1] For increasing ρ, the approximation factor vanishes to a value smaller than any polynomial function of ρ, i.e., smaller than any ρ^ϵ, for any $\epsilon > 0$.

information into transmitted signals, and where $\mathcal{D}_\rho$ denotes the decoder(s) that extract information from the received signals. Let the transmitted codewords $\boldsymbol{x}$ be picked, with uniform probability, from the codebook $\mathcal{X}_\rho$. Transmission has duration T, SNR ρ, rate

$$R = \frac{1}{T}\log_2|\mathcal{X}_\rho|,$$

and an enforced power constraint such that

$$\frac{1}{|\mathcal{X}_\rho|}\sum_{\boldsymbol{x}\in\mathcal{X}_\rho}\|\boldsymbol{x}\|^2 = T. \tag{2}$$

For simplicity we write $\mathcal{X}$,$\mathcal{D}$, and we let the parameterization be implied.

Consider a policy $\mathcal{P}$ (short for $\mathcal{P}_{\rho\mathcal{X}\mathcal{D}}$), which generally trades-off error performance with complexity, by forcing the decoder to limit the number of numerical operations (i.e., flops), up to a maximum designated number of flops. Once this limiting number of flops is reached, the decoder quits and declares an error. This limiting number of flops may or may not be chosen as a function of the instantaneous $\boldsymbol{H},\boldsymbol{y}$, and will generally depend on ρ.

B. Rate, reliability and complexity

The error probability $P_{\mathcal{X}\mathcal{D}\mathcal{P}}$ introduced by the specific $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$, is simply

$$P_{\mathcal{X}\mathcal{D}\mathcal{P}} := \mathrm{P}\left(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}\ :\ \hat{\boldsymbol{x}}_{\mathcal{X}\mathcal{D}\mathcal{P}} \neq \boldsymbol{x}\right), \tag{3}$$

where $\hat{\boldsymbol{x}}_{\mathcal{X}\mathcal{D}\mathcal{P}}$ denotes the vector decoded by $\mathcal{D}$, under the restrictions of $\mathcal{P}$. For a given $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$ and a given realization of problem inputs $\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}$, then $N_{\mathcal{X}\mathcal{D}\mathcal{P}}(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w})$ will denote the overall instantaneous introduced complexity, in flops. Then worst-case complexity is simply given by

$$C_{\mathcal{X}\mathcal{D}\mathcal{P}} := \sup_{\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}} N_{\mathcal{X}\mathcal{D}\mathcal{P}}(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}). \tag{4}$$

A pertinent measure of performance for any $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$ then becomes the corresponding set of achievable combinations $(R,\rho,P_{\mathcal{X}\mathcal{D}\mathcal{P}},C_{\mathcal{X}\mathcal{D}\mathcal{P}})$, or an equivalent one-to-one mapping of this set. In the rate-of-change setting of interest, a meaningful general mapping is chosen to have output of the form $\left(R,\rho,\frac{\log P_{\mathcal{X}\mathcal{D}\mathcal{P}}}{\log Z},\frac{\log C_{\mathcal{X}\mathcal{D}\mathcal{P}}}{\log L}\right)$, where Z,L can regulate the refinements of these rates-of-change.

III. Error and complexity exponents

A. Quantifying error performance: DMT

As a measure of rate-reliability performance, we adopt the refinement of the diversity-multiplexing tradeoff, identified by Zheng and Tse in [2], as a fundamental performance limit in outage-limited MIMO communications.

In this setting, both the error probability $P_{\mathcal{X}\mathcal{D}\mathcal{P}}$ introduced by the specific $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$, as well as the cardinality of $\mathcal{X}$, are parameterized by $Z=\rho$. Specifically the code cardinality

$$|\mathcal{X}| = 2^{RT},$$

is described by the *multiplexing gain*

$$r \triangleq \lim_{\rho\to\infty}\frac{R}{\log_2\rho} = \lim_{\rho\to\infty}\frac{1}{T}\frac{\log|\mathcal{X}|}{\log\rho}, \tag{5}$$

and the associated error performance delivered by the transceiver and policy, is described by the *diversity gain* [2]

$$d_{\mathcal{X}\mathcal{D}\mathcal{P}}(r) := -\lim_{\rho\to\infty}\frac{\log P_{\mathcal{X}\mathcal{D}\mathcal{P}}}{\log\rho}. \tag{6}$$

B. Regulating and quantifying complexity performance: worst-case complexity exponent

In the described setting, for a given ρ,R, we consider the one-to-one mapping of

$$C_{\mathcal{X}\mathcal{D}\mathcal{P}} \ \leftrightarrow\ \frac{\log\left(C_{\mathcal{X}\mathcal{D}\mathcal{P}}\right)}{\log L}$$

where L is some chosen function of $|\mathcal{X}|$. A general asymptotic worst-case complexity measure then takes the form

$$\lim_{\rho\to\infty}\frac{\log C_{\mathcal{X}\mathcal{D}\mathcal{P}}}{\log L}. \tag{7}$$

Similar to the DMT in [2] which meaningfully measures the high-SNR $P_{\mathcal{X}\mathcal{D}\mathcal{P}}$ as a polynomial power of ρ, our chosen measure of complexity will also be an exponent over ρ, keeping in line with pertinent complexity behavior $C_{\mathcal{X}\mathcal{D}\mathcal{P}}$ of most known transceivers. In this scale of interest where $L=\rho$, the worst-case complexity exponent takes the following form

$$c_{\mathcal{X}\mathcal{D}\mathcal{P}}(r) := \lim_{\rho\to\infty}\frac{\log\left(\sup_{\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}} N_{\mathcal{X}\mathcal{D}\mathcal{P}}(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w})\right)}{\log\rho}. \tag{8}$$

We briefly note that $c_{\mathcal{X}\mathcal{D}\mathcal{P}}(r)$ is set by the structural properties of the design $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$ as well as the statistical properties of $\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}$. We also note that the worst-case exponent of any reasonable decoder is bounded as

$$0 \leq c_{\mathcal{X}\mathcal{D}\mathcal{P}}(r) \leq rT.$$

The upper bound is easily seen to be tight, because $c_{\mathcal{X}\mathcal{D}}(r) = rT$ is the exponent corresponding to full-search uninterrupted ML decoders[2] in the presence of a canonical code with multiplexing gain r, i.e., $|\mathcal{X}| \doteq \rho^{rT}$, where the $\doteq$ notation is used when $f(\rho)\doteq\rho^x$ iff (cf. [2])

$$\lim_{\rho\to\infty}\frac{\log f(\rho)}{\log\rho} = x\,. \tag{9}$$

The symbols $\dot{\geq}$ and $\dot{\leq}$ are defined similarly.

The tradeoff is now put together.

IV. Performance-complexity tradeoff

We proceed to establish the fundamental limits, optimized over all achievable rate-reliability-complexity combinations of any transceiver and policy, up to a factor that vanishes in the limit of high ρ.

Towards this we describe the decoder and encoder structures, that together with a specific policy, meet a natural upper bound to this tradeoff, for all values of r. We start with the decoder, but for now disregard the policy.

[2]We here note that strictly speaking, $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$ may potentially introduce a complexity exponent larger than rT. In such a case though, $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$ may be substituted by a lookup table implementation of $\mathcal{X}$ and an unrestricted ML decoder. This encoder-decoder will jointly introduce a complexity exponent equal to rT, thus maintaining the bound. It is also noted that the number of flops per visited codeword is independent of ρ.

1) The candidate decoder – the DMT optimal LLL based LR-aided, regularized linear decoder: We focus on the efficient and DMT optimal, LLL-based lattice-reduction (LR)-aided regularized linear decoder, presented in its general form in [1], [7], [8] for different settings, drawing from works such as [9], [10]. We clarify that the decoder applies to lattice codes, and for completeness recall the decoder's three main steps. In the first step, the decoder performs *regularization* via MMSE-GDFE like preprocessing, thus inducing a regularized metric (cf. [1])

$$\hat{\boldsymbol{x}}_{\mathrm{L}} = \arg\min_{\hat{\boldsymbol{x}}\in\Lambda_r} \|\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{x}}\|^2 + \|\hat{\boldsymbol{x}}\|^2_{\boldsymbol{T}}. \tag{10}$$

In the above, Λ_r is the scaled lattice corresponding to the code, and $\boldsymbol{T}$ is a positive definite matrix. The above metric penalizes far away elements of Λ_r that are generally non codewords. The second step includes *lattice-reduction* using the LLL algorithm [11], and the last step is an efficient *linear detection* using, for example, the *rounding off* algorithm.

Under standard assumptions on continuity, and in the presence of a policy $\mathcal{P}_{rT}$ that lets the decoder run its course irrespective of the complexity, the above decoder was shown in [1] to be DMT optimal, i.e., that

$$d_{\mathcal{X},\mathcal{D}_{\mathrm{LRR}},\mathcal{P}_{rT}}(r) = \sup_{\mathcal{D}} d_{\mathcal{X},\mathcal{D},\mathcal{P}_{rT}},$$

irrespective of the lattice design $\mathcal{X}$, and irrespective of the fading statistics.

It is the case though that the decoder's LLL step introduces worst-case complexity that is infinite [12]. This problem is successfully addressed by the policy discussed below.

2) The LR-based policy $\mathcal{P}_{\mathrm{LR}}$: To limit the above infinite complexity, the work in [1] proposed a policy that capitalizes on the fact that to achieve DMT optimality, it is not required to LLL reduce every conceivable channel. Instead, in the event that too many flops occur, the policy instructs the implementation of the LLL algorithm to halt, and the decoder to declare an error. Special emphasis is given to guaranteeing that the event of halting is not more common than the event of error, thus avoiding degradation of the asymptotic error performance. Specifically the halting policy, to be denoted as $\mathcal{P}_{\mathrm{LR}}$, was defined on the basis of the bound on the number K of LLL cycles that are necessary for reduction of matrix $\boldsymbol{M}$ which generates the composite code-channel lattice. This bound is given by [12], [13] to be

$$K \leq n^2 \log_{\frac{2}{\sqrt{3}}} \kappa(\boldsymbol{M}) + n, \tag{11}$$

where $\kappa(\boldsymbol{M})$ denotes the 2-norm condition number of $\boldsymbol{M}$. Based on this bound, $\mathcal{P}_{\mathrm{LR}}$ deploys the LLL algorithm only if

$$\kappa(\boldsymbol{M}) \leq \rho^{\frac{1}{2}(d_{\mathrm{ML}}(r)+1)+\epsilon}, \quad \epsilon > 0, \tag{12}$$

where $d_{\mathrm{ML}}(r) \triangleq d_{\mathcal{X},\mathcal{D}_{\mathrm{ML}},\mathcal{P}_{rT}}(r)$ describes the DMT achieved by the uninterrupted ML decoder. By showing that

$$\mathrm{P}\left(\kappa(\boldsymbol{M}) \geq \rho^{\frac{1}{2}(d_{\mathrm{ML}}(r)+1)+\epsilon}\right) \dot{\leq} \rho^{-d_{\mathrm{ML}}(r)},$$

i.e., that the event of halting is less common than the event of error under full ML decoding, it was proven in [1] that, over any range of multiplexing gains r, the combination of $\mathcal{D}_{\mathrm{LRR}}$ and $\mathcal{P}_{\mathrm{LR}}$ achieves DMT optimal decoding of any lattice design $\mathcal{X}_\Lambda$, and does so with worst-case complexity of $O(\log\rho)$. This implies a worst-case complexity that is at most linear in the rate[3], at high SNR. It also constitutes substantial improvement over sphere decoding implementations where the worst-case complexity reported (see for example [14] for fast decodable codes [15]–[17]) is also exponential in R, albeit with a smaller exponent than full search.

3) The overall worst-case complexity exponent jointly introduced by lattice encoding, $\mathcal{D}_{\mathrm{LRR}}$ and $\mathcal{P}_{\mathrm{LR}}$: With the above in mind, we proceed to establish the overall computational complexity jointly introduced by lattice encoding and by the different components of $\mathcal{D}_{\mathrm{LRR}}$, in the presence of $\mathcal{P}_{\mathrm{LR}}$.

a) Decoder and policy: We first quickly note that the regularization and linear-decoding steps, introduce complexity that is essentially independent of $\rho, \boldsymbol{H}$, and bounded above by $O(n^2)$, thus inducing a zero complexity exponent.

Regarding the lattice reduction step, we recall the hard bound

$$\begin{aligned} K &\leq n^2 \log_{\frac{2}{\sqrt{3}}} \kappa(\boldsymbol{M}) + n \\ &\dot{\leq} n^2 \log_{\frac{2}{\sqrt{3}}} \rho^{\frac{1}{2}(d_{\mathrm{ML}}(r)+1)+\epsilon} + n, \quad \epsilon > 0, \end{aligned}$$

on the number of LLL flops enforced by $\mathcal{P}_{\mathrm{LR}}$. This bound implies that

$$\exists z \in \mathbb{R}^+ : \mathrm{P}\left(N(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}) > z\log\rho\right) = 0,$$

which in turn means that

$$\mathrm{P}\left(N(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}) > \rho^c\right) \doteq \rho^{-\infty}, \ \forall c > 0. \tag{13}$$

In conjunction with the equivalent representation (drawing from [18], which presents some $c(r)$ of different $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$)

$$c(r) = \sup\{c\colon -\lim_{\rho\to\infty} \frac{\log P(N(\boldsymbol{H},\boldsymbol{x},\boldsymbol{w}) \geq \rho^c)}{\log\rho} \leq d(r)\}$$

of a worst-case complexity exponent $c(r)$ that allows for $d(r)$, we conclude that the LLL algorithm under $\mathcal{P}_{\mathrm{LR}}$, also introduces an effective complexity exponent equal to zero. Consequently the entire $\mathcal{D}_{\mathrm{LRR}}, \mathcal{P}_{\mathrm{LR}}$ introduces a minimal complexity exponent, equal to zero.

b) Lattice encoding: Moving on to encoding, it is again easy to see that any lattice code $\mathcal{X}_\Lambda$ comes with encoding complexity that is bounded as $O(n^2)$, thus minimally adding to the overall complexity exponent of any transceiver/policy.

We are now able to combine the complexities from the encoder and the decoder, and to provide the following.

Lemma 1: A lattice code $\mathcal{X}_\Lambda$, in conjunction with the decoder-policy $\mathcal{D}_{\mathrm{LRR}}, \mathcal{P}_{\mathrm{LR}}$, jointly accept a minimum, over all encoders, decoders and policies, effective complexity exponent, i.e.,

$$c_{\mathcal{X}_\Lambda,\mathcal{D}_{\mathrm{LRR}},\mathcal{P}_{\mathrm{LR}}} = \inf_{\mathcal{X},\mathcal{D},\mathcal{P}} c_{\mathcal{X},\mathcal{D},\mathcal{P}} = 0. \tag{14}$$

[3]The result is extended in [8] to the MIMO-MAC case, to show that this optimality holds with worst-case complexity that is at most linear in the users' sum-rate.

4) The overall error performance: With respect to the error performance of $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$, we utilize the result in [1] which proves that the DMT optimality of $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$, holds irrespective of the lattice code that it is applied to, i.e., that for *any* fixed lattice code $\mathcal{X}_\Lambda$, then

$$d_{\mathcal{X}_\Lambda,\mathcal{D}_{\text{LRR}},\mathcal{P}_{\text{LR}}}(r) = \sup_{\mathcal{D},\mathcal{P}} d_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r). \quad (15)$$

Disregarding for now issues on code design, we proceed to formalize the performance-complexity optimality of $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$.

5) The overall effective complexity/error exponent jointly induced by lattice encoding, $\mathcal{D}_{\text{LRR}}$ and $\mathcal{P}_{\text{LR}}$: Combining (14) and (15) gives the following.

Lemma 2: The high-SNR rate-reliability-complexity tradeoff achieved by the $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$, is better or equal to the tradeoff achieved by any other decoder-policy, irrespective of the lattice code $\mathcal{X}_\Lambda$ applied, i.e.,

$$\begin{aligned}&\big(d_{D_{\mathcal{X}_\Lambda,\mathcal{D}_{\text{LRR}},\mathcal{P}_{\text{LR}}}}(r), c_{D_{\mathcal{X}_\Lambda,\mathcal{D}_{\text{LRR}},\mathcal{P}_{\text{LR}}}}(r)\big)\\ &\qquad = \big(\sup_{\mathcal{D},\mathcal{P}} d_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r), \inf_{\mathcal{D},\mathcal{P}} c_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r)\big).\end{aligned}$$

Here it is stressed that this achievable tradeoff may be suboptimal, as it is limited by the reliability capabilities of the specific code $\mathcal{X}_\Lambda$.

What remains now is to combine the optimal components $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$, with suitable code designs.

6) Employing DMT optimal codes, to meet the rate-reliability-complexity tradeoff: We have just seen in Lemma 2 that, given any lattice design $\mathcal{X}_\Lambda$, the combination $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$ achieves the highest allowable tradeoff over any transceiver-policy that includes $\mathcal{X}_\Lambda$. Consequently what remains is to identify *lattice* code designs that optimize both $c_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r)$ and $d_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r)$, in the presence of $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$. Optimizing of $c_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r)$ has already been achieved in Lemma 1 which proved that any lattice design $\mathcal{X}_\Lambda$ gives $c_{\mathcal{X}_\Lambda,\mathcal{D}_{\text{LRR}},\mathcal{P}_{\text{LR}}} = \inf_{\mathcal{X},\mathcal{D},\mathcal{P}} c_{\mathcal{X},\mathcal{D},\mathcal{P}} = 0$. Hence what remains is to find a lattice design that optimizes $d_{\mathcal{X}_\Lambda,\mathcal{D},\mathcal{P}}(r)$, in the presence of $\mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$. This in turn is further simplified in the presence of (15), and the task is now limited to simply finding DMT optimal lattice codes, i.e., codes that asymptotically meet the outage region

$$\mathcal{O} = \{\boldsymbol{H} : \frac{1}{T}\log\det\big(I + \beta\boldsymbol{H}\boldsymbol{H}^\dagger\big) < R\}, \text{ some fixed } \beta,$$

of the equivalent MIMO channel to achieve asymptotically optimal performance (cf. [2])

$$d_{\text{opt}}(r) := \sup_{\mathcal{X},\mathcal{D},\mathcal{P}} d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r) = P(\boldsymbol{H} \in \mathcal{O}). \quad (16)$$

The existence of such lattice codes has been proven in [19], for the quasi-static Rayleigh fading channel, and a unified family of such codes was explicitly constructed in [20] using *cyclic division algebras* (CDA). Further such codes have, over the last few years, been described for a plethora of MIMO models. These codes are based on different variants of CDA codes (cf. [21], [22]), and have been shown, under basic continuity conditions, to provide DMT optimality for all channel dimensions, and most often for all fading statistics. Such codes can, for example, be found in [20], [23]–[29], and they DMT-optimally apply to several MIMO scenarios, including MIMO, MIMO-OFDM, MIMO-MAC (Rayleigh fading), MIMO-ARQ, as well as to most existing cooperative communication protocols.

For all the above MIMO scenarios, we have now the final result, which holds under basic continuity conditions.

A. The optimal tradeoff

Theorem 1: The high-SNR optimal, over all encoders, decoders and policies, rate-reliability-complexity behavior is given by

$$\begin{aligned}&\text{opt}_{\mathcal{X},\mathcal{D},\mathcal{P}}\big(d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r), c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)\big)\\ &= \big(d_{\mathcal{X}_{\text{CDA}},\mathcal{D}_{\text{LRR}},\mathcal{P}_{\text{LR}}}(r), c_{\mathcal{X}_{\text{CDA}},\mathcal{D}_{\text{LRR}},\mathcal{P}_{\text{LR}}}(r)\big) = \big(d_{\text{opt}}(r), 0\big) \quad (17)\end{aligned}$$

and is achieved for all multiplexing gains, all channel dimensions and (in most known cases) all fading statistics, by the CDA-based designs $\mathcal{X}_{\text{CDA}}$, the LR-aided regularized linear decoder $\mathcal{D}_{\text{LRR}}$, and the LR-based policy $\mathcal{P}_{\text{LR}}$.

Equivalently the result shows that the achievable rate-reliability-complexity combination

$$\big(R = r\log\rho, P \doteq \rho^{-d_{\text{opt}}(r)}, C \doteq \rho^0\big) \quad (18)$$

is optimal, up to a factor that vanishes in the limit of high ρ. We quickly note that $\mathcal{X}_{\text{CDA}}, \mathcal{D}_{\text{LRR}}, \mathcal{P}_{\text{LR}}$ is currently the only known tradeoff-optimal design.

Directly from the above, we have the following.

Theorem 2: In the high SNR regime, the union of all achievable rate-reliability-complexity combinations, considering all reasonable $\mathcal{X},\mathcal{D},\mathcal{P}$, is given by

$$\begin{aligned}&\{\big(R = r\log\rho, P \doteq \rho^{-d(r)}, C \doteq \rho^{c(r)}\big)\},\\ &\qquad\qquad 0 \le d(r) \le d_{\text{opt}}(r), 0 \le c(r) \le rT.\end{aligned}$$

Proof: For a given R, any of the above reliability-complexity pairs can be achieved by employing an $\mathcal{X},\mathcal{D},\mathcal{P}$ that is optimal with respect to (18), modifying though $\mathcal{P}$ to introduce the appropriate amount of extra complexity and errors[4]. ∎

a) Optimal limits on general reliability-complexity functions: Another measure of the rate-reliability-complexity capabilities of different transceivers can take the form of general utility functions. Towards this we define the following.

Definition 1: Let Γ be a weighting function that is increasing in $d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)$, decreasing in $c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)$, and which reflects the different costs assigned separately to erroneous detection, and complexity. Then for any given $\mathcal{X},\mathcal{D},\mathcal{P}$, the Γ-*general rate-reliability-complexity limit* takes the form

$$D_{\mathcal{X},\mathcal{D},\mathcal{P}}(r) := \Gamma\big(d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r), c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)\big). \quad (19)$$

Towards motivating useful and meaningful implementation of the limit, we identify the following simple manifestation as one of many special cases of the general limit.

Definition 2: The *homogeneous rate-reliability-complexity limit* for a given $\mathcal{X},\mathcal{D},\mathcal{P}$, and a given weighting factor $\gamma \ge 0$, takes the form

$$D_{\mathcal{X},\mathcal{D},\mathcal{P}}(r) := d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r) - \gamma c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r), \quad (20)$$

[4] Constructing such modification is trivial. We also note that the worst case $(r, d(r) = 0, c(r) = rT)$ corresponds to a full-search transceiver that provides subexponential decay of the probability of error, for increasing SNR.

and describes the diversity gain minus the normalized complexity cost.

It is interesting to interpret the rate-reliability-complexity limit $D_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)$, as a limit that describes the high-SNR error capabilities of $\mathcal{X}$,$\mathcal{D}$,$\mathcal{P}$, per unit power and complexity. Equivalently, the limit may be described as a measure of diversity gain per complexity order.

The following result, which holds under basic continuity conditions and for the same scenarios as Theorem 1, describes the optimizing value achieved by a large family of measures Γ.

Theorem 3: The optimal, over all encoders, decoders and policies, Γ-general rate-reliability-complexity limit $D(r)$, is given by

$$D(r) = \Gamma\big(\sup_{\mathcal{X},\mathcal{D},\mathcal{P}} d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r), 0\big) = \Gamma\big(d_{\text{opt}}(r), 0\big) \tag{21}$$

and is achieved for all multiplexing gains, and all channel dimensions by the CDA-based designs $\mathcal{X}_{\text{CDA}}$, the LR-aided regularized linear decoder $\mathcal{D}_{\text{LRR}}$, and the LR-based policy $\mathcal{P}_{\text{LR}}$.

Proof: The proof is direct by noting that

$$D(r) = \sup_{\mathcal{X},\mathcal{D},\mathcal{P}} \Gamma\big(d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r), c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)\big) \tag{22}$$

$$\leq \Gamma\big(\sup_{\mathcal{X},\mathcal{D},\mathcal{P}} d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r), \inf_{\mathcal{X},\mathcal{D},\mathcal{P}} c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r)\big), \tag{23}$$

and then by applying Theorem 1. ■

The following holds for the more intuitive, cost-symmetric version of the limit.

Corollary 3a: The optimal, over all encoders, decoders and policies, homogeneous rate-reliability-complexity limit, is given by

$$D(r) = \sup_{\mathcal{X},\mathcal{D},\mathcal{P}} d_{\mathcal{X},\mathcal{D},\mathcal{P}}(r) - \gamma c_{\mathcal{X},\mathcal{D},\mathcal{P}}(r) = d_{\text{opt}}(r). \tag{24}$$

V. Conclusions

The tradeoff and its achievability, provide worst-case guarantees on the complexity required for provably optimal performance in outage-limited MIMO communications. The guarantees hold over a surprisingly broad setting, and they come with reduced transmission energy and delay, as well as reduced algorithmic power consumption and hardware. The tradeoff concisely quantifies these guarantees and the capabilities of different transceivers, as well as quantifies the role of policies in simplifying algorithms which would otherwise introduce unbounded complexity.

References

[1] J. Jaldén and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models," 2009, submitted to *IEEE Trans. Inform. Theory*, available on arXiv:cs/0905.4023 [cs.IT].

[2] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[3] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[4] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[5] A. H. Banihashemi and A. K. Khandani, "On the complexity of decoding lattices using the Korkin-Zolotarev reduced basis," *IEEE Trans. Inform. Theory*, vol. 44, no. 1, pp. 162–171, Jan. 1998.

[6] D. Micciancio, "The hardness of the closest vector problem with preprocessing," *IEEE Trans. Inform. Theory*, vol. 47, no. 3, pp. 1212–1215, Mar. 2001.

[7] J. Jaldén and P. Elia, "LR-aided MMSE lattice decoding is DMT optimal for all approximately universal codes," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[8] P. Elia and J. Jaldén, "General DMT optimality of LR-aided linear MIMO-MAC transceivers with worst-case complexity at most linear in sum-rate," in *Proc. IEEE Information Theory Worshop (ITW)*, Cairo, Egypt, Jan. 2010.

[9] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE Global Conf. Communications (GLOBECOM)*, Taipei, Taiwan, Nov. 2002.

[10] C. Windpassinger and R. F. H. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction," in *Proc. IEEE Information Theory Worshop (ITW)*, Paris, France, Mar. 2003.

[11] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Matematische Annalen*, vol. 261, no. 4, pp. 1432–1807, Dec. 1982.

[12] J. Jaldén, D. Seethaler, and G. Matz, "Worst- and average-case complexity of LLL lattice reduction in MIMO wireless systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, Nevada, USA, Apr. 2008.

[13] H. Daudée and B. Vallée, "An upper bound on the average number of iterations of the LLL algorithm," *Theoretical Computer Science*, vol. 123, no. 1, Jan. 1994.

[14] E. Biglieri, Y. Hong, and E. Viterbo, "On fast-decodable space-time block codes," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 524–530, Feb. 2009.

[15] O. Tirkkonen and R. Kashaev, "Combined information and performance optimization of linear MIMO modulations," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, Jun. 2002.

[16] J. Paredes, A. B. Gershman, and M. Gharavi-Alkhansari, "A 2×2 space-time code with non-vanishing determinant and fast maximum likelihood decoding," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Honolulu, Hawaii, USA, Apr. 2007.

[17] M. Samuel and M. P. Fitz, "Reducing the detection complexity by using 2×2 multi-strata space-time codes," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007.

[18] J. Jaldén and P. Elia, "On the sphere-decoding complexity of linear diversity-multiplexing optimal codes," 2010, in preparation.

[19] H. El Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 968–985, Jun. 2004.

[20] P. Elia, K. R. Kumar, S. A. Pawar, P. Vijay Kumar, and H.-F. Lu, "Explicit space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3869–3884, Sep. 2006.

[21] B. A. Sethuraman, B. Sundar Rajan, and V. Shashidhar, "Full-diversity, high-rate, space-time block codes from division algebras," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2596–2616, Oct. 2003.

[22] J.-C. Belfiore and G. Rekaya, "Quaternionic lattices for space-time coding," in *Proc. IEEE Information Theory Worshop (ITW)*, Paris, France, Mar. 2003.

[23] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the MIMO amplify-and-forward cooperative channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 647–663, Feb. 2007.

[24] H.-F. Lu, "Constructions of multiblock space-time coding schemes that achieve the diversity multiplexing tradeoff," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3790–3796, Aug. 2008.

[25] K. Raj Kumar and G. Caire, "Space-time codes from structured lattices," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 547–556, Feb. 2009.

[26] P. Elia and P. Vijay Kumar, "Space-time codes that are approximately universal for the parallel, multi-block and cooperative DDF channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, 2009.

[27] H.-F. Lu and C. Hollanti, "Diversity-multiplexing tradeoff-optimal code constructions for symmetric MIMO multiple-access channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[28] P. Elia, K. Vinodh, M. Anand, and P. Vijay Kumar, "D-MG tradeoff and optimal codes for a class of AF and DF cooperative communication protocols," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, Jul. 2009.

[29] S. A. Pawar, K. Raj Kumar, P. Elia, P. Vijay Kumar, and B. A. Sethuraman, "Space-time codes achieving the DMD tradeoff of the MIMO-ARQ channel," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, Jul. 2009.