General DMT optimality of LR-aided linear MIMO-MAC transceivers with worst-case complexity at most linear in sum-rate

Petros Elia and Joakim Jaldén

***Abstract*—In the setting of multiple-access MIMO channels, the work establishes the DMT optimality of lattice-reduction (LR)-aided regularized *linear* decoders. This is achieved irrespective of the lattice design applied by each user. The decoding algorithms employ efficient solutions to the Nearby Vector Problem with Preprocessing in the presence of a regularized non-Euclidean metric, and in the presence of time-outs.**

The decoders' optimality induces a worst-case computational complexity that is at most linear in the users' sum-rate. This constitutes a substantial improvement over the state of art of DMT optimal decoding, including ML decoders with complexity that is exponential in the sum-rate, or lattice decoders based on solutions to the NP-hard closest vector problem (CVP). The optimality of the efficient decoders is established for all channel statistics, for all channel dimensions, for any number of users, and irrespective of the different rates. The findings directly apply to different computationally intensive multi-user settings such as multi-user MIMO, multi-user cooperative communications, and multi-user MIMO-OFDM.

***Index Terms*—Multiple-Access Methods, Multi-User Receivers, Diversity Multiplexing Tradeoff, Lattice Designs, Lattice Reduction, Linear Receivers, MIMO Communications.**

I. Multiple-access Channel and Lattice Designs

A. Introduction

The general MIMO multiple-access linear channel (MIMO-MAC) model describes different scenarios where independent users utilize multi-dimensional transmit-receive signals in the presence of fading and of each other's interference.

A fundamental performance limit in outage limited communications was presented in [1], for the single-user MIMO case, in the form of the *diversity multiplexing tradeoff* (DMT). This was extended in [2] to the multi-user MIMO-MAC case. The tradeoff has since been widely adopted as a benchmark for transceiver design and analysis. The work in [3] proved, for the single-user case, the existence of DMT optimal lattice-based encoders and decoders, and the result was extended in [4] to the multi-user case.

Recent work by the authors in [5] proved, for the single-user case, the DMT optimality of explicit encoder-decoder structures that employ computationally efficient lattice reduction (LR)-aided linear decoders. The extension of this optimality to the multiple-access case poses challenges relating to the variable densities attributed to the code-channel lattices of different users having different rates. These challenges are addressed herein and the decoders' optimality in the MIMO-MAC case is proven for any optimal or suboptimal lattice code.

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).

P. Elia is with the Mobile Communications Department, EURECOM, Sophia Antipolis, France (email: elia@eurecom.fr) (tel/fax: +33 49300 8132)

J. Jaldén is with the Signal Processing Lab, School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden (email: Joakim.jalden@ee.kth.se)

B. System model

We consider a general (real) multiple-access MIMO channel model [2]

$$\boldsymbol{y} = \sum\nolimits_{i=1}^{k} \boldsymbol{H}_i \boldsymbol{x}_i + \boldsymbol{w} \tag{1}$$

with k independent users and a single receiver, where $\boldsymbol{y}, \boldsymbol{w} \in \mathbb{R}^m$, $\boldsymbol{H}_i \in \mathbb{R}^{m\times n_i}$, and where $\boldsymbol{x}_i \in \mathbb{R}^{n_i}$ is the codeword transmitted by user i. The transmitted codewords $\boldsymbol{x}_i$, for $i = 1, \cdots, k$, are drawn with uniform probability from some codebooks $\mathcal{X}_{r_i} \subset \mathbb{R}^{n_i}$, the channels $\boldsymbol{H}_i$ are random with arbitrary distributions and parameterized by ρ which can be interpreted as the users' SNR, and $\boldsymbol{w}$ is assumed to be i.i.d. Gaussian with unit variance. The users operate for some duration T, at some SNR ρ, and user specific rates[1] $R_i = \frac{1}{T}\log_2|\mathcal{X}_{r_i}|$, which define the user's *multiplexing gain* r_i [1], [2] according to

$$r_i \triangleq \lim_{\rho\to\infty} \frac{R_i(\rho)}{\log_2\rho} = \lim_{\rho\to\infty} \frac{1}{T}\frac{\log|\mathcal{X}_{r_i}(\rho)|}{\log\rho}. \tag{2}$$

In this setting $\boldsymbol{x} \triangleq [\boldsymbol{x}_1^{\mathrm{T}}\, \boldsymbol{x}_2^{\mathrm{T}} \cdots \boldsymbol{x}_k^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^n$, $n = \sum_{i=1}^n n_i$, can be seen to originate from a codebook $\mathcal{X}_{\mathbf{r}} \triangleq \mathcal{X}_{r_1}\times\cdots\times\mathcal{X}_{r_k} \subset \mathbb{R}^n$, where $\mathbf{r} \triangleq [r_1\, r_2 \cdots r_k]^T$ defines the multiplexing gain vector [2].

The joint ML decoder is known to be DMT optimal for each user [2], and is given by

$$\hat{\boldsymbol{x}}_{\mathrm{ML}} = \arg\min_{\hat{\boldsymbol{x}}\in\mathcal{X}_{\mathbf{r}}} \|\boldsymbol{y} - \sum\nolimits_{i=1}^{k} \boldsymbol{H}_i\hat{\boldsymbol{x}}_i\|^2. \tag{3}$$

The *diversity gain* delivered by the set of designs $\mathcal{X}_{r_i}$ under ML decoding is given as a function of $\mathbf{r}$ to be (c.f. [1], [2])

$$d_{\mathrm{ML}}(\mathbf{r}) \triangleq -\lim_{\rho\to\infty} \frac{\log \mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{ML}} \neq \boldsymbol{x}\right)}{\log\rho}. \tag{4}$$

Herein, we consider lattice designs given by $\mathcal{X}_{r_i} = (\rho^{-\frac{r_i T}{n_i}}\Lambda_i) \cap \mathcal{R}_i$, for each user i, where $\Lambda_i \triangleq \{\boldsymbol{G}_i\boldsymbol{z}_i \mid \boldsymbol{z}_i \in \mathbb{Z}^{n_i}\} \subset \mathbb{R}^{n_i}$ is a lattice generated by[2] $\boldsymbol{G}_i \in \mathbb{R}^{n_i\times n_i}$, where $\boldsymbol{z}_i$ is the integer information vector associated to user i, and where $\mathcal{R}_i \subset \mathbb{R}^{n_i}$ is a compact convex shaping region. Specifically $\mathcal{R}_i$ contains $\mathbf{0}$ in its interior, and is independent of ρ. For each $r_i \geq 0$, $\mathcal{X}_{r_i}$ induces a cardinality $|\mathcal{X}_{r_i}| = \rho^{r_i T}$.

[1] We here assume that one use of (1) corresponds to T uses of some underlying channel (see [5]).

[2] The assumption that $\boldsymbol{G}_i$ is square incurs no loss of generality (c.f. [5]).

II. DMT Optimality of Regularized Lattice Decoding in the Multiple-Access Setting

The (general) *regularized lattice decoder* was in [5] given for the single user case, and takes in the multiple-access setting the following form:

$$\hat{\boldsymbol{x}}_{\mathrm{L}} = \arg\min_{\hat{\boldsymbol{x}}\in\Lambda_{\mathbf{r}}} \|\boldsymbol{y} - \sum_{i=1}^{k} \boldsymbol{H}_i\hat{\boldsymbol{x}}_i\|^2 + \sum_{i=1}^{k} \|\hat{\boldsymbol{x}}_i\|^2_{\boldsymbol{T}_i} \quad (5)$$

where

$$\Lambda_{\mathbf{r}} \triangleq \rho^{-\frac{r_1 T}{n_1}}\Lambda_1 \times \cdots \times \rho^{-\frac{r_k T}{n_k}}\Lambda_k,$$

and where $\|\hat{\boldsymbol{x}}_i\|^2_{\boldsymbol{T}_i} \triangleq \hat{\boldsymbol{x}}_i^{\mathrm{T}}\boldsymbol{T}_i\hat{\boldsymbol{x}}_i$ for any[3] positive definite $\boldsymbol{T}_i$, $i = 1, \ldots, k$. Apart from the obvious addition of the regularization terms, we emphasize that (5) differs from (3) also in that the search is performed over the full lattice, and not just the codebook. Although suboptimal in general, searching over the full lattice tends to symmetrize the decoder and will allow for methods used to reduce the decoder complexity (this will be further discussed in Section III). Nevertheless the following result, extending the result in [5], shows that under the assumption of uniformly distributed $\boldsymbol{x}_i$ over $\mathcal{X}_{r_i}$, (5) defines a DMT optimal decoder for the general multiple-access setting.

Theorem 1: For any set of lattice designs $\{\mathcal{X}_{r_i} = (\rho^{-\frac{r_i T}{n_i}}\Lambda_i) \cap \mathcal{R}_i\}_{i=1}^{k}$ employed by the users, and for any fading distribution such that $d_{\mathrm{ML}}(\mathbf{r})$ is continuous at $\mathbf{r}$, the regularized lattice decoder in (5) is DMT optimal, i.e., gives optimal diversity

$$d_{\mathrm{L}}(\mathbf{r}) \triangleq -\lim_{\rho\to\infty} \frac{\log \mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{L}} \neq \boldsymbol{x}\right)}{\log\rho} = d_{\mathrm{ML}}(\mathbf{r}). \quad (6)$$

A. Proof of Theorem 1

The proof extends on the proof for the single-user case in [5] by considering lattices with variable densities across the different dimensions associated to each user, and by showing how in this setting, the regularization factor $\sum_{i=1}^{k}\|\hat{\boldsymbol{x}}_i\|^2_{\boldsymbol{T}_i}$ guarantees DMT optimality, i.e., guarantees that the DMT-performance provided by the lattice designs in the presence of regularized lattice decoding, matches the performance of the same designs in the presence of ML decoding. Note however that we make no assumption that the design achieves the optimal DMT of the underlying channel (see [5]).

Consider the following, guaranteed to exist, spherical region

$$\mathcal{B}_i \triangleq \{\boldsymbol{d} \in \mathbb{R}^{n_i} \mid \|\boldsymbol{d}\|^2 \leq \gamma_i\}, \quad (7)$$

where the radius $\gamma_i > 0$ is chosen to be independent of ρ, and also chosen to guarantee that $\boldsymbol{d}_1 + \boldsymbol{d}_2 \in \mathcal{R}_i$ for any $\boldsymbol{d}_1, \boldsymbol{d}_2 \in \mathcal{B}_i$. For

$$\nu_{\mathbf{r}} \triangleq \min_{\boldsymbol{d}\in\bigotimes_{i=1}^{k}(\Lambda_{r_i}\cap\mathcal{B}_i):\boldsymbol{d}\neq\mathbf{0}} \tfrac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2, \quad (8)$$

where $\boldsymbol{H} \triangleq [\boldsymbol{H}_1\, \boldsymbol{H}_2\, \cdots\, \boldsymbol{H}_k]$, the first task will be to show that for any $\mathbf{r} > \mathbf{0}$,

$$\limsup_{\rho\to\infty} \frac{\log \mathrm{P}\left(\nu_{\mathbf{r}} \leq 1\right)}{\log\rho} \leq -d_{\mathrm{ML}}(\mathbf{r}). \quad (9)$$

[3] $\boldsymbol{T}_i$ may be optimized for improved performance.

To see this, assume that $\boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i$ and $\nu_{\mathbf{r}} \leq 1$, where the latter implies the existence of $\boldsymbol{d} \in \bigotimes_{i=1}^{k}(\mathcal{B}_i \cap \Lambda_{r_i})$, $\boldsymbol{d} \neq \mathbf{0}$, such that $\nu_{\mathbf{r}} = \frac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2 \leq 1$ (see (8)). The fact that $\boldsymbol{d}, \boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i$ and $\boldsymbol{d}, \boldsymbol{x} \in \bigotimes_{i=1}^{k}\Lambda_{r_i}$ implies that $\hat{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{d} \in \bigotimes_{i=1}^{k}(\mathcal{R}_i \cap \Lambda_{r_i})$, i.e., $\hat{\boldsymbol{x}} \in \mathcal{X}_{\mathbf{r}}$. The ML decoder will choose $\hat{\boldsymbol{x}}$ over $\boldsymbol{x}$ with probability $\mathrm{P}\left(\boldsymbol{x} \to \hat{\boldsymbol{x}} | \boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i, \boldsymbol{H}\right) = Q\left(\frac{1}{2}\|\boldsymbol{H}\boldsymbol{d}\|\right)$, where since $\nu_{\mathbf{r}} = \frac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2 \leq 1$, then

$$\mathrm{P}\left(\boldsymbol{x} \to \hat{\boldsymbol{x}} | \boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i, \boldsymbol{H}\right) \geq Q(1) > 0. \quad (10)$$

Furthermore by applying standard counting techniques it may be shown that for $\mathbf{r} > 0$, and for $\boldsymbol{x}$ uniformly distributed over $\bigotimes_{i=1}^{k}(\mathcal{R}_i \cap \Lambda_{r_i})$, it holds that

$$\lim_{\rho\to\infty} \mathrm{P}\left(\boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i\right) = \frac{V(\bigotimes_{i=1}^{k}\mathcal{B}_i)}{V(\bigotimes_{i=1}^{k}\mathcal{R}_i)} > f \quad (11)$$

for some $f > 0$ independent of ρ.

For $\hat{\boldsymbol{x}}_{\mathrm{ML}} \in \mathcal{X}_{\mathbf{r}}$ being the output of the ML decoder, and given that $\boldsymbol{x}$ is independent of $\boldsymbol{H}$ and thus also independent of $\nu_{\mathbf{r}}$, it is the case that

$$\mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{ML}} \neq \boldsymbol{x}\right) \geq \mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{ML}} \neq \boldsymbol{x} | \boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i, \nu_{\mathbf{r}} \leq 1\right) \times \mathrm{P}\left(\boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i\right) \mathrm{P}\left(\nu_{\mathbf{r}} \leq 1\right). \quad (12)$$

Equation (10) says that $\mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{ML}} \neq \boldsymbol{x} | \boldsymbol{x} \in \bigotimes_{i=1}^{k}\mathcal{B}_i, \nu_{\mathbf{r}} \leq 1\right) \doteq \rho^0$, which combines with (11), (12) to give that

$$\mathrm{P}\left(\nu_{\mathbf{r}} \leq 1\right) \dot{\leq} \mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{ML}} \neq \boldsymbol{x}\right)$$

which proves (9).

As in the case of (8), define

$$\nu_{\mathbf{r}+\boldsymbol{\zeta}} \triangleq \min_{\boldsymbol{d}\in\bigotimes_{i=1}^{k}(\Lambda_{r_i+\zeta_i}\cap\mathcal{B}_i):\boldsymbol{d}\neq\mathbf{0}} \tfrac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2$$

where $\boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 & \zeta_2 & \cdots & \zeta_k \end{bmatrix}$, and where ζ_i, $i = 1, \cdots, k$ and δ are any constants that satisfy

$$\frac{2\zeta_i T}{n_i} = \frac{2\zeta_j T}{n_j} > \delta > 0. \quad (13)$$

At this stage make the assumptions that

$$\nu_{\mathbf{r}+\boldsymbol{\zeta}} \geq 1 \quad (14)$$

and that

$$\|\boldsymbol{w}\|^2 \leq \rho^{\delta}. \quad (15)$$

The first task is to show that these two conditions are sufficient for a correct decision by the regularized lattice decoder in (5), provided that ρ is sufficiently large.

Towards this end, let $c \triangleq \max_{\boldsymbol{r}\in\bigotimes_{i=1}^{k}\mathcal{R}_i}\|\boldsymbol{r}\|^2_{\boldsymbol{T}}$, where $\boldsymbol{T} \triangleq \mathrm{diag}[\boldsymbol{T}_1, \cdots, \boldsymbol{T}_k]$ is block diagonal, and note that c is independent of the transmitted codeword $\boldsymbol{x}$ and ρ, and that $c < \infty$ because the $\mathcal{R}_i$ are bounded. This, in conjunction with the condition in (15), implies that $\boldsymbol{x}$ induces a regularized metric of

$$\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \|\boldsymbol{x}\|^2_{\boldsymbol{T}} = \|\boldsymbol{w}\|^2 + \|\boldsymbol{x}\|^2_{\boldsymbol{T}} \leq \rho^{\delta} + c. \quad (16)$$

The task will be to compare (16) with the metric for any $\hat{\boldsymbol{x}} \in \Lambda_{\mathbf{r}}, \hat{\boldsymbol{x}} \neq \boldsymbol{x}$. Towards this note that the assumption in (14) implies

$$\tfrac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2 \geq 1 \quad \forall\, \boldsymbol{d} \in \bigotimes_{i=1}^{k} (\Lambda_{r_i+\zeta_i} \cap \mathcal{B}_i) : \boldsymbol{d} \neq \mathbf{0}\,, \tag{17}$$

by the definition in (8). Since $\Lambda_{r_i} = \rho^{\frac{\zeta_i T}{n_i}} \Lambda_{r_i+\zeta_i}$, then scaling (17) by $\rho^{\frac{\zeta_i T}{n_i}}$ results in

$$\tfrac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2 \geq \rho^{\frac{2\zeta_i T}{n_i}} \quad \forall\, \boldsymbol{d} \in \bigotimes_{i=1}^{k} (\Lambda_{r_i} \cap \rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i)\,,\, \boldsymbol{d} \neq \mathbf{0}. \tag{18}$$

It is now the case that $\boldsymbol{x} \in \frac{1}{2}\bigotimes_{i=1}^{k} \rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i$ for all $\rho \geq \rho_1$, given some sufficiently large ρ_1, because for bounded $\mathcal{R}_i$ then $\mathcal{R}_i \subset \frac{1}{2}\rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i$, for all i. Note that ρ_1 can be chosen independent of the transmitted $\boldsymbol{x}$.

First, consider the case where $\hat{\boldsymbol{x}} \in \frac{1}{2}\bigotimes_{i=1}^{k} (\rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i \cap \Lambda_{r_i})$ and note that for any such $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$, then $\boldsymbol{d} = \boldsymbol{x} - \hat{\boldsymbol{x}} \in \bigotimes_{i=1}^{k} (\rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i \cap \Lambda_{r_i})$, and

$$\tfrac{1}{4}\|\boldsymbol{H}(\boldsymbol{x} - \hat{\boldsymbol{x}})\|^2 = \tfrac{1}{4}\|\boldsymbol{H}\boldsymbol{d}\|^2 \geq \rho^{\frac{2\zeta_i T}{n_i}} \tag{19}$$

directly from (18). Combining (13),(14), and (19), gives that there is some $\rho_2 \geq \rho_1$, independent of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, for which the triangle inequality guarantees that

$$\|\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}\|^2 = \|\boldsymbol{H}(\boldsymbol{x} - \hat{\boldsymbol{x}}) + \boldsymbol{w}\|^2 \geq \rho^{\frac{2\zeta_i T}{n_i}}$$

for all $\rho \geq \rho_2$. Consequently,

$$\|\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}\|^2 + \|\hat{\boldsymbol{x}}\|^2_{\boldsymbol{T}} \geq \rho^{\frac{2\zeta_i T}{n_i}} \tag{20}$$

for any $\hat{\boldsymbol{x}} \in \Lambda_r$ where $\hat{\boldsymbol{x}} \in \frac{1}{2}\bigotimes_{i=1}^{k} \rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i$ and $\rho \geq \rho_2$.

Second, consider the case where $\hat{\boldsymbol{x}} \notin \frac{1}{2}\bigotimes_{i=1}^{k} \rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i$ in which case (7) states that $\|\hat{\boldsymbol{x}}\|^2 \geq \frac{1}{4}(\min_i\{\gamma_i\})^k \rho^{\frac{2\zeta_i T}{n_i}}$ which in turn implies $\|\hat{\boldsymbol{x}}\|^2_{\boldsymbol{T}} \geq \frac{1}{4}\rho^{\frac{2\zeta_i T}{n_i}}(\min_i\{\gamma_i\})^k \lambda_{\min}(\boldsymbol{T})$ where $\lambda_{\min}(\boldsymbol{T}) > 0$ denotes the minimum eigenvalue of $\boldsymbol{T}$. It follows that

$$\|\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}\|^2 + \|\hat{\boldsymbol{x}}\|^2_{\boldsymbol{T}} \geq \frac{1}{4}\rho^{\frac{2\zeta_i T}{n_i}}(\min_i\{\gamma_i\})^k \lambda_{\min}(\boldsymbol{T}) \tag{21}$$

for any $\hat{\boldsymbol{x}} \notin \bigotimes_{i=1}^{k} \rho^{\frac{\zeta_i T}{n_i}} \mathcal{B}_i$.

For the transmitted codeword $\boldsymbol{x}$, (16) implies that

$$\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \|\boldsymbol{x}\|^2_{\boldsymbol{T}} \leq a(\rho) \triangleq \rho^{\delta} + c \tag{22}$$

which is compared to

$$b(\rho) \triangleq \min\big(1, \tfrac{1}{4}(\min_i\{\gamma_i\})^k \lambda_{\min}(\boldsymbol{T})\big)\rho^{\frac{2\zeta_i T}{n_i}} \tag{23}$$

to show that, given (13), there is some $\rho_3 \geq \rho_2$, again independent of $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$, for which $a(\rho) < b(\rho)$ for all $\rho > \rho_3$. As a result, for any other $\hat{\boldsymbol{x}} \in \Lambda_{\mathbf{r}} \backslash \{\boldsymbol{x}\}$, (20) and (21) imply that

$$\|\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{x}}\|^2 + \|\hat{\boldsymbol{x}}\|^2_{\boldsymbol{T}} \geq b(\rho) > a(\rho) \tag{24}$$

for all $\rho \geq \rho_3$. It has thus been shown that the transmitted codeword yields the minimum metric, i.e., that $\hat{\boldsymbol{x}}_{\mathrm{L}} = \boldsymbol{x}$, as long as $\rho \geq \rho_3$ and as long as the conditions in (14), (15) are satisfied. Equivalently for an error to occur when $\rho \geq \rho_3$ it must be that $\nu_{\mathbf{r}+\boldsymbol{\zeta}} < 1$ or $\|\boldsymbol{w}\| > \rho^{\delta}$, and thus

$$\mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{L}} \neq \boldsymbol{x}\right) \leq \mathrm{P}\left(\nu_{\mathbf{r}+\boldsymbol{\zeta}} < 1\right) + \mathrm{P}\left(\|\boldsymbol{w}\| > \rho^{\delta}\right)\,, \tag{25}$$

for $\rho \geq \rho_3$. The exponential tail of the Gaussian distribution guarantees that $\mathrm{P}\left(\|\boldsymbol{w}\| > \rho^{\delta}\right) \doteq \rho^{-\infty}$, and the term becomes asymptotically irrelevant. As a result, in conjunction with (9), it is the case that for any $\boldsymbol{\zeta} > 0$ satisfying (13) then $\limsup_{\rho\to\infty} \frac{\log \mathrm{P}(\hat{\boldsymbol{x}}_{\mathrm{L}} \neq \boldsymbol{x})}{\log \rho} \leq -d_{\mathrm{ML}}(\mathbf{r} + \boldsymbol{\zeta})$. Given the continuity of $d_{\mathrm{ML}}(\mathbf{r})$ at $\mathbf{r}$, i.e., given that $\lim_{\boldsymbol{\zeta}\to\mathbf{0}} d_{\mathrm{ML}}(\mathbf{r}+\boldsymbol{\zeta}) = d_{\mathrm{ML}}(\mathbf{r})$, as we choose ζ_i arbitrarily small, it may be concluded that

$$\limsup_{\rho\to\infty} \frac{\log \mathrm{P}\left(\hat{\boldsymbol{x}}_{\mathrm{L}} \neq \boldsymbol{x}\right)}{\log \rho} \leq -d_{\mathrm{ML}}(\mathbf{r})\,, \tag{26}$$

for any $\mathbf{r} \geq 0$. The optimality of the ML decoder completes the proof. □

III. Computationally Efficient Decoding: C-approximate solutions to the CVP

The decoder in (5) may be rewritten in the form of a closest vector problem (CVP). To this end, note that by incorporating the lattice design, the signal model in (1) may be written as

$$\boldsymbol{y} = \boldsymbol{H}\Theta\mathbf{G}\boldsymbol{z} + \boldsymbol{w} \tag{27}$$

where Θ is a power normalizing diagonal matrix having the elements of the ith n_i-tuple on the diagonal being equal to $\rho^{-\frac{r_i T}{n_i}}$, where $\boldsymbol{z} = [\boldsymbol{z}_1^T \ \cdots \ \boldsymbol{z}_k^T]^T$ and $\boldsymbol{z}_i$ is the original integer information vector for user i, and where $\mathbf{G}$ is the composite lattice generator matrix corresponding to $\Lambda_{\mathbf{r}}$. Let

$$\boldsymbol{y}' \triangleq \boldsymbol{Q}^T \begin{bmatrix} \boldsymbol{y} \\ \mathbf{0} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \boldsymbol{H} \\ \sqrt{\boldsymbol{T}} \end{bmatrix} = \boldsymbol{Q}\boldsymbol{R}$$

where $\boldsymbol{Q} \in \mathbb{R}^{(m+n)\times n}$ and $\boldsymbol{R} \in \mathbb{R}^{n\times n}$ are the factors of the QR decomposition of the extended channel matrix, i.e., $\boldsymbol{Q}$ has orthogonal columns and $\boldsymbol{R}$ is upper triangular. It is now straightforward to show that (5) is equivalent to

$$\hat{\boldsymbol{z}}_{\mathrm{L}} = \arg\min_{\hat{\boldsymbol{z}} \in \mathbb{Z}^n} \|\boldsymbol{y}' - \boldsymbol{M}\hat{\boldsymbol{z}}\|^2 \tag{28}$$

where $\boldsymbol{M} \triangleq \boldsymbol{R}\boldsymbol{\Theta}\boldsymbol{G}$, and where $\hat{\boldsymbol{x}}_{\mathrm{L}} = \boldsymbol{\Theta}\boldsymbol{G}\hat{\boldsymbol{z}}_{\mathrm{L}}$.

Obtaining $\hat{\boldsymbol{x}}_{\mathrm{L}}$ in (5), or equivalently $\hat{\boldsymbol{z}}_{\mathrm{L}}$ in (28), thus requires the solution of a CVP in the lattice generated by $\boldsymbol{M}$. However, the CVP is unfortunately NP-hard in general [6], even after preprocessing [7], [8]. Specifically [7] proves that for a large family of *recursive cube search algorithms*, such as the fastest currently known CVP algorithms ([9], [7]), there exist lattices that induce decoding complexity that is exponential in dimension, irrespective of the amount of preprocessing. The work in [8] extends the result to all algorithms, and shows that it is not possible to find optimal polynomial time solutions to the general CVP, (i.e., irrespective of the input lattice), even in the presence of any form of preprocessing.

In our case this NP-hardness implies that even if LR preprocessing methods [10] are used when obtaining the exact solution to (28), it is unlikely that there will be any general search techniques with a (worst-case) complexity that grows sub-exponentially in the problem dimension n, unless the code itself provides a structure that simplifies decoding, such as for example in the case of orthogonal designs [11].

However, such constraints would severely limit the available lattice dimensions and rates. For most high-performance lattice designs no such efficient solutions to (28), or (5), are known.

A. DMT optimality of C-approximate lattice decoding

The above motivates the study of suboptimal implementations of the regularized lattice decoder, in the form of approximate solutions to the induced CVP. Specifically we are interested in finding a *C-approximate solution* to (5), i.e., in employing a *C-approximation algorithm* [12], which for a fixed $C > 1$ and for arbitrary inputs $\boldsymbol{y} \in \mathbb{R}^m$ and $\boldsymbol{H} \in \mathbb{R}^{m\times n}$, decides on $\hat{\boldsymbol{x}} \in \Lambda_{\mathbf{r}}$ that satisfies

$$\xi(\hat{\boldsymbol{x}}) \leq C\xi(\hat{\boldsymbol{x}}_{\rm L}) \quad \text{where} \quad \xi(\hat{\boldsymbol{x}}) = \|\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{x}}\|^2+\|\hat{\boldsymbol{x}}\|^2_{\boldsymbol{T}}\,. \tag{29}$$

In the above, recall that $\hat{\boldsymbol{x}}_{\rm L}$ is by definition the vector which provides the minimum metric in (5). The following result states that any C-approximation algorithm for (5) is sufficient for DMT optimal decoding.

Theorem 2: For any set of lattice designs $\{\mathcal{X}_{r_i} = (\rho^{-\frac{r_i T}{n_i}}\Lambda_i) \cap \mathcal{R}_i\}_{i=1}^k$ employed by the users, and for any fading distribution such that $d_{\rm ML}(\mathbf{r})$ is continuous at $\mathbf{r}$, all C-approximate implementations of the regularized lattice decoder are DMT optimal over the MIMO-MAC, provided C is independent of ρ, i.e.,

$$d_{\rm A}(\mathbf{r}) = d_{\rm ML}(\mathbf{r})\,, \tag{30}$$

where

$$d_{\rm A}(\mathbf{r}) \triangleq -\lim_{\rho\to\infty}\frac{\log \mathrm{P}\left(\hat{\boldsymbol{x}}_{\rm A} \neq \boldsymbol{x}\right)}{\log\rho}\,, \tag{31}$$

for $\boldsymbol{x}$ uniformly distributed over $\mathcal{X}_{\mathbf{r}}$, and where $\hat{\boldsymbol{x}}_{\rm A}$ is any C-approximate solution to (5).

Proof: The proof follows from the proof of Theorem 1, it can be found in [5] and is sketched here for completeness. In particular, for $a(\rho), b(\rho)$ in (22), (23) then (13) implies that $\lim_{\rho\to\infty}\frac{b(\rho)}{a(\rho)} = \infty$, and that we can select $\rho_4 \geq \rho_3$ such that $b(\rho) \geq Ca(\rho)$ for all $\rho \geq \rho_4$. As the metric for $\boldsymbol{x}$ is upper bounded by $a(\rho)$, and the metric of any other vector is lower bounded by $b(\rho)$, it follows that under the assumptions of (14),(15), and for $\rho \geq \rho_4$, then $\hat{\boldsymbol{x}}_{\rm A} = \boldsymbol{x}$. The remaining proof is similar to the proof of Theorem 1. □

We note that, in many cases, it is more straightforward to find C-approximate solutions to (28) rather than to (5), although as shown in [5] any C-approximate solutions to (28) is also a C-approximate solution to (5). We shall for this reason focus on (28) in what follows.

B. DMT optimality of LR-aided lattice decoding: employing solutions for the C-approximate CVP with Preprocessing

The NP-hardness of the CVP motivates the use of C-approximate rather than exact solutions. As was shown by Theorem 2, any such solution also provides a DMT optimal decoder. However, the CVP may not be efficiently approximated for arbitrary $C > 1$. For example, it has been shown in [6] that approximating the CVP to within almost-polynomial factors in n, is still NP-hard[4]. Nevertheless, for sufficiently large values of C, tractable solutions do exist. To this end, consider the CVP with a reduced lattice basis, i.e.,

$$\hat{\boldsymbol{z}}'_{\rm L} = \arg\min_{\hat{\boldsymbol{z}}'\in\mathbb{Z}^n}\|\boldsymbol{y}' - \boldsymbol{M}\boldsymbol{U}\hat{\boldsymbol{z}}'\|^2 \tag{32}$$

where $\boldsymbol{U} \in \mathbb{R}^{n\times n}$ is a unimodular matrix, i.e., $\boldsymbol{U}\mathbb{Z}^n = \mathbb{Z}^n$. The problem in (32) is clearly equivalent to (28) with $\hat{\boldsymbol{z}}_{\rm L} = \boldsymbol{U}\hat{\boldsymbol{z}}'_{\rm L}$. Minimizing the norm in (32) over arbitrary $\hat{\boldsymbol{z}}' \in \mathbb{R}^n$, followed by simple rounding to the nearest integer point leads to the approximate solition

$$\hat{\boldsymbol{z}}'_{\rm A} = \lceil(\boldsymbol{M}\boldsymbol{U})^{-1}\boldsymbol{y}'\rfloor \tag{33}$$

where $\lceil\cdot\rfloor$ denotes per-component rounding to $\mathbb{Z}^n$. This approach was in [15] referred to as the *rounding off* algorithm, and shown to provide a C_1-approximate solution with $C_1 \triangleq 1 + 2n(9/2)^{\frac{n}{2}}$, given a reduced basis $\boldsymbol{M}\boldsymbol{U}$ obtained by the LLL algorithm[5] [10]; it is equivalent to the LLL-based LR-aided linear implementation of the MMSE-GDFE decoder when $\boldsymbol{T} = \boldsymbol{I}$ [16]–[20]. Similarly, the LLL-based LR-aided SIC implementation was in [15] referred to as the *nearest plane algorithm*, and shown to provide a C_2-approximate solution to the CVP with $C_2 \triangleq 2^{\frac{n}{2}}$ [15]. We note that C_1, C_2 are independent of ρ, and that Theorem 2 therefore applies. For completeness, we give the following corollary to Theorem 2.

Corollary 2a: The efficient LLL-based LR-aided linear implementations of the regularized lattice decoders provide DMT optimal decoding of any set of lattice designs operating in a multiple-access MIMO setting.

Proof: The corollary follows by the equivalence of the LR-aided linear decoder and the rounding off algorithm in [15], or of the LR-aided SIC decoder and the nearest plane algorithm in [15], in conjunction with Theorem 2. □

Interestingly, Corollary 2a applies also to a time-limited implementation of the Schnorr-Euchner (SE) sphere decoder [21], [22] operating on (32), provided the sphere decoder tree-search is allowed to reach the first leaf-node (see [5] for more details).

C. Decoding complexity

The LR-aided decoders discussed above, first LLL reduce the lattice basis $\boldsymbol{M}$, and then provide a C-approximate solution to the CVP according to (33) or by using an SIC based procedure. The complexity of obtaining an approximate solution in the reduced basis is only $O(n^2)$ [16], [17], [23] and independent of ρ, while the preprocessing relying on the LLL reduction is more complex. Thus, the decoder complexity is dominated by the basis reduction, i.e., finding $\boldsymbol{U}$.

The work in [24] shows that even though the worst-case complexity of the LLL algorithm is often cited as polynomial

[4]We note that C-approximate solutions for arbitrary $C > 1$ are not guaranteed even if oracle preprocessing of the lattice basis is allowed. Specifically [13] (resp. [14]) have shown that the CVP with preprocessing (CVPP) is NP-hard to approximate to within any factor less than $\sqrt{5/3}$ (resp. $\sqrt{3}$), or equivalently that under the assumption that P$\neq$ NP there exist lattices for which the CVP cannot be approximated to within $\sqrt{5/3}$ (resp. $\sqrt{3}$) in polynomial time, no matter how the lattice is represented.

[5]Note here that regularization guarantees that $\boldsymbol{M}$ is always full rank, rendering the LLL algorithm applicable, regardless of the channel realization and the system dimensionality.

in the dimension of the lattice, it is in fact infinite if applied to arbitrary, real valued, $\boldsymbol{M} \in \mathbb{R}^{n\times n}$. This implies that the worst-case complexity of the LLL-based LR-aided decoder is, strictly speaking, unbounded if applied to arbitrary channels. However, in order to achieve DMT optimal performance it is not required to LLL reduce every conceivable channel. Instead the decoder may be allowed to time-out and declare an error when the number of floating point operations exceeds a given threshold, given that the time-out event is not more common, in asymptotic terms, than the probability of decoding error.

To gain insight into such a time-out mechanism, note that the number K of LLL cycles required to reduce a given lattice basis $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ may be bounded according to [24], [25]

$$K \leq n^2 \log_s \kappa(\boldsymbol{M}) + n \tag{34}$$

where $s = 2/\sqrt{3}$ and where $\kappa(\boldsymbol{M})$ denotes the 2-norm condition number of $\boldsymbol{M}$. Each iteration requires $O(n^2)$ floating point operations [10], which may be reduced to $O(n)$ if only an effectively LLL-reduced basis is required [26]. In light of (34) we may thus limit the application of the LLL algorithm to bases $\boldsymbol{M}$ with bounded condition number $\kappa(\boldsymbol{M})$, or allow the decoder the option to time out, stop, and declare an error. Details of the time-out condition can be found in [5].

Employing the above time-out approach, and quantifying the natural connection between channels that induce error and channels that induce high-complexity, allows for the following.

Corollary 2b: For a very general class of multiple-access channels, and for any given number of users, DMT optimal decoding of any set of lattice designs is feasible at a worst-case complexity that is at most $O(\log \rho)$.

Proof: The proof follows the steps in [5] and is omitted here due to lack of space.

In the scale of interest, this worst-case complexity can be seen to be $O(n^4)$ and at most linear in the sum-rate of the users[6], and is thus substantially reduced in comparison to the worst-case complexity of other proven DMT optimal decoders, which have complexity that is exponential in the sum-rate and dimensionality, or induce solutions to the NP-hard CVP.

IV. Conclusion

The work established the DMT optimality of efficient LLL-based LR-aided linear (or SIC) implementations of the regularized lattice decoders, for a very general multiple-access MIMO setting. In conjunction with DMT optimal MIMO-MAC lattice-designs [27], the current work establishes that DMT optimality in the computationally demanding MIMO-MAC setting can be achieved with computationally efficient decoders, at a complexity that is at most linear in rate.

References

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

[2] D. N. C. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859–1874, Sep. 2004.

[6]Note that when a DFE rather than a linear receiver is used, the term reduces to $O(n^3)$ [26]. Also note that $O(n^3)$ quantifies the complexity of the decoders used, but by no means does it imply that the work here deals with the case of asymptotically high n.

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

[4] Y.-H. Nam and H. E. Gamal, "On the optimality of lattice coding and decoding in multiple access channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007.

[5] 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].

[6] I. Dinur, G. Kindler, and S. Safra, "Approximating CVP to within almost-polynomial factors is NP-hard," in *39th Annual Symposium on Foundations of Computer Science*, Palo Alto, California, Nov. 1998.

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

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

[9] R. Kannan, "Minkowski's convex body theorem and integer programming," *Mathematics of Operation Research*, vol. 12, no. 3, pp. 415–440, Aug. 1987.

[10] 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.

[11] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[12] J. Hromkovič, *Algorithms for Hard Problems: Introduction to Combinatorial Optimization, Randomization, Approximation and Heuristics*, 2nd ed. Springer, 2002.

[13] U. Feige and D. Micciancio, "The inapproximability of lattice and coding problems with preprocessing," *Journal of Computer and System Sciences*, vol. 69, no. 1, pp. 45–67, 2004.

[14] O. Regev, "Improved inapproximability of lattice and coding problems with preprocessing," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2031–2037, Sep. 2004.

[15] L. Babai, "On Lovász' lattice reduction and the nearest lattice point problem," *Combinatorica*, vol. 6, no. 1, pp. 1–13, Mar. 1986.

[16] 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.

[17] 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.

[18] A. D. Murugan, H. E. Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: rediscovering the sequential decoder," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 933– 953, Mar. 2006.

[19] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2401, Oct. 2003.

[20] ——, "MMSE-GDFE lattice decoding for underdetermined linear channels," in *Proc. Conf. on Information Science and Systems*, Princeton, New Jersey, USA, 2004.

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

[22] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, pp. 181–191, 1994.

[23] D. Wübben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer, "Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction," in *Proc. IEEE Int. Conf. Communications (ICC)*, Paris, France, Jun. 2004.

[24] 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.

[25] 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.

[26] C. Ling and H. Howgrave-Graham, "Effective LLL reduction for lattice decoding," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007.

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