MIMO Detection by Lagrangian Dual Maximum-Likelihood Relaxation: Reinterpreting Regularized Lattice Decoding

Jiaxian Pan[†], Wing-Kin Ma[‡], and Joakim Jaldén[§]

Abstract

This paper considers lattice decoding for multi-input multi-output (MIMO) detection under PAM constellations. A key aspect of lattice decoding is that it relaxes the symbol bound constraints in the optimal maximum-likelihood (ML) detector for faster implementations. It is known that such a symbol bound relaxation may lead to a damaging effect on the system performance. For this reason, regularization was proposed to mitigate the out-of-bound symbol effects in lattice decoding. However, minimum mean square error (MMSE) regularization is the only method of choice for regularization in the present literature. We propose a systematic regularization optimization approach by considering a Lagrangian dual relaxation (LDR) of the ML detection problem. As it turns out, the proposed LDR formulation is to find the best diagonally regularized lattice decoder to approximate the ML detector, and all diagonal regularizations, including the MMSE regularization, can be subsumed under the LDR formalism. We show that for the 2-PAM case, strong duality holds between the LDR and ML problems. Also, for general PAM, we prove that the LDR problem yields a duality gap no worse than that of the well-known semidefinite relaxation method. To physically realize the proposed LDR, the projected subgradient method is employed to handle the LDR problem so that the best regularization can be found. The resultant method can physically be viewed as an adaptive symbol bound control wherein regularized lattice decoding is recursively performed to correct the decision. Simulation results show that the proposed LDR approach can outperform the conventional MMSE-based lattice decoding approach.

Index terms− lattice decoding, lattice reduction, MIMO detection, Lagrangian duality, regularization

EDICS: MSP-CODR (MIMO precoder/decoder design), SPC-DETC (Detection, estimation, and demodulation).

[†]Jiaxian Pan is with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong S.A.R., China. E-mail: jxpan@ee.cuhk.edu.hk.

[‡]Wing-Kin Ma is the corresponding author. Address: Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong S.A.R., China. E-mail: wkma@ieee.org.

[§]Joakim Jaldén is with the ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: jalden@kth.se.

I. INTRODUCTION

Many important communication techniques developed over the past few decades, such as multiuser code division multiple access (CDMA), spatial multiplexing and space-time coding in multi-antenna frequency-flat channels, space-frequency coding in multi-antenna orthogonal frequency division multiplexing (OFDM) [1], and most recently, very large-scale antenna systems [2], fall into the scope of linear multi-input multi-output (MIMO) systems. More precisely, in the above noted communication scenarios, one can often formulate the signal model as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{\nu},$$

where $\boldsymbol{y}$ is a multi-dimensional received signal vector, $\boldsymbol{s}$ is a multi-dimensional transmitted symbol vector whose elements are drawn from a symbol constellation set, $\boldsymbol{H}$ is a matrix that characterizes the input-output relation, and $\boldsymbol{\nu}$ is noise. One of the fundamental problems under this generic MIMO setting is the MIMO detection problem, which is to detect the transmitted vector $\boldsymbol{s}$ based on the observation $\boldsymbol{y}$ and $\boldsymbol{H}$. Among various MIMO detection approaches, maximum-likelihood (ML) detection is optimal in minimizing the error probability; more precisely, the probability that the symbol vector $\boldsymbol{s}$ is incorrectly detected is minimized. However, the ML detection problem is known to be NP-hard [3], and even the sphere decoder [4], which is the state-of-the-art solver for the ML detection problem, is proven to exhibit exponential complexity with respect to the problem size [5]. The computational difficulty in solving the ML problem exactly has stimulated a number of works that aim to approximate the ML detector in an efficient manner.

The study of lattice decoding[1] for MIMO detection has received much attention [6]–[10], owing to its good tradeoff between detection accuracy and complexity. In the so-called naive lattice decoding (NLD) method [6], the objective is to find the closest lattice point to the received signal over the whole lattice generated by the channel. In contrast, the ML detector only searches those lattice points within the symbol bounds determined by the constellation used. In other words, NLD relaxes the symbol bound constraints in the ML detector. The reason behind the symbol bound relaxation in NLD is that lattice reduction methods [11], [12] can be employed to help speed up the lattice point search. The improved computational efficiency of NLD has been verified empirically [7], [13]. To further reduce the complexity, one may perform suboptimal lattice decoding methods, such as linear detection or decision feedback (DF) after lattice reduction; these methods are known as the lattice reduction-aided (LRA) methods [14]–[17];

---

[1]In this paper, we adopt the common terminology in lattice decoding that "decoding" is synonymous with "detection".

see also [18]–[20] for some more recent developments. Though NLD, as well as its LRA methods, can achieve the full receive diversity for spatial multiplexing over i.i.d Gaussian fading channels [16], [21], it was revealed that NLD fails to achieve the optimal diversity-multiplexing tradeoff (DMT) under general MIMO system models [6], [8]. The reason behind this is that NLD completely ignores the symbol bounds. The inability for NLD to impose symbol bounds has motivated the study of regularized lattice decoding [6], [9]. In regularized lattice decoding, a quadratic penalization term is added to the lattice decoding metric for the purpose of penalizing the lattice points that are far away from the origin. Since the penalization of the lattice points outside the symbol bounds is much heavier than that of the lattice points within the symbol bounds, regularized lattice decoding is more likely to produce a valid lattice point within the symbol bounds. Therefore, the penalization can serve as a means of controlling the symbol bounds implicitly. Interestingly, lattice decoding with any positive-definite regularization can achieve the optimal DMT in general [9]. Moreover, the same optimal DMT result applies to its LRA counterparts. Despite these exciting advances, no existing works attempt to shed light on how the regularization should be optimally designed. Other than the well-known minimum mean square error (MMSE) regularization, no other choice of regularization is offered in the literature.

In this paper, we develop a new MIMO detection approach by considering a Lagrangian dual relaxation (LDR) of the ML detection problem. Unlike many existing relaxed ML MIMO detection approaches where the problem domain is usually the continuous real vector space, the proposed LDR formulation has a discrete set of all integers as the problem domain. As it turns out, the LDR formulation preserves the structure of lattice decoding, and all diagonally regularized lattice decoders, including NLD and MMSE-regularized lattice decoding (MMSE LD), can be seen as particular instances of the LDR formulation. In addition, the LDR problem can be regarded as a problem of finding the best diagonal regularization (in a Lagrangian sense) for the corresponding regularized lattice decoder to approximate the ML detector. Surprisingly, when particularizing the LDR formulation to the 2-PAM constellation case, we show that strong duality holds between the LDR problem and the ML detection problem. This means that solving the regularized lattice decoder with a proper diagonal regularization found by LDR yields the same optimal solution as the ML detector. We also prove that LDR yields relaxation tightness no worse than a previously developed ML relaxation, namely, bound constrained semidefinite relaxation [22], [23], which is a representative relaxed ML detection method. In order to obtain the best regularization, we resort to the projected subgradient method [24], [25] for handing the LDR problem which is a nondifferentiable convex optimization problem. Interestingly, the projected subgradient method involves solving a sequence of regularized lattice decoding problems whose regularizations are iteratively updated for alleviating symbol

bound violation. Hence, the projected subgradient method can be interpreted as an adaptive symbol bound control via regularization. Based on the exact implementation of the projected subgradient method, we devise a new MIMO detector which is called the LDR LD detector. The practical implementation of the LDR LD detector is carefully designed for complexity reduction. To further reduce the complexity, we also consider efficient approximate LDR LD detectors by combining the idea of the projected subgradient method and conventional suboptimal lattice decoding methods. Simulation results show that the LDR LD detector achieves almost the same symbol error rate (SER) performance as the ML detector. Moreover, the approximate LDR LD detectors can yield significantly better SER performance than their conventional counterparts with comparable computational complexities.

The rest of this paper is organized as follows. We introduce the MIMO system model and briefly review lattice decoding in Section II. Then, in Section III, we propose our LDR formulation for the ML detection problem, analyze its relationship with the ML problem, and introduce the idea of the projected subgradient method for solving the LDR problem. This is followed by Section IV, where we explain in detail the practical implementations of the LDR methods. Simulation results are presented in Section V to demonstrate the performance of the proposed methods. Section VI concludes this paper.

The results presented in this article can be reproduced by the source code we provided at http://www.ee.cuhk.edu.hk/~wkma/mimo/.

II. BACKGROUND

A. MIMO Detection Problem Formulation

We consider a generic real-valued linear MIMO signal model

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{\nu}, \tag{1}$$

where $\boldsymbol{y} \in \mathbb{R}^M$ is a received signal vector, $\boldsymbol{H} \in \mathbb{R}^{M \times N}$ a channel matrix, $\boldsymbol{s} \in \mathbb{R}^N$ a transmitted symbol vector, and $\boldsymbol{\nu} \in \mathbb{R}^M$ a noise vector. Here, N and M denote the input and output problem sizes, respectively. We assume that the elements of $\boldsymbol{\nu}$ are independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance σ_ν^2, and that each symbol s_i is drawn from a $(u+1)$-PAM constellation set

$$\mathcal{S} = \{\pm 1, \pm 3, \ldots, \pm u\}$$

in an i.i.d. uniform manner, where u is the constellation symbol bound. As mentioned in the introduction, (1) has been widely adopted as a signal model for detection problems arising in various communication scenarios; readers are referred to the literature, such as [4], [7], for more details. Also, we should note

that any complex-valued linear MIMO model with QAM constellations can be equivalently represented by the above real-valued PAM MIMO model through a simple complex-to-real conversion process; again, see the literature [4], [7].

The ML MIMO detector, which is optimal in minimizing the vector error probability of detecting $\boldsymbol{s}$ given knowledge of the channel $\boldsymbol{H}$, is the solution of the following optimization problem

$$\min_{\boldsymbol{s}} \quad \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 \tag{2}$$
$$\text{s.t.} \quad s_i \in \{\pm 1, \pm 3, \ldots, \pm u\}, \ i = 1, \ldots, N.$$

where $\|\cdot\|$ denotes the 2-norm. The ML problem (2) is known to be NP-hard for a general realization $(\boldsymbol{y}, \boldsymbol{H})$. This implies that all the existing exact ML solvers, including the well-known sphere decoders [4], [7], would be computationally prohibitive when N is large.

B. Lattice Decoding

Lattice decoding aims at tackling the ML problem efficiently using lattice reduction-based techniques. A simple version of lattice decoding is naive lattice decoding (NLD) [6], where, instead of dealing with the ML problem (2), one considers a lattice decoding problem

$$\min_{\boldsymbol{s}} \quad \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 \tag{3}$$
$$\text{s.t.} \quad \boldsymbol{s} \in 2\mathbb{Z}^N + \mathbf{1},$$

where $\mathbf{1}$ denotes an all-one vector of appropriate length, and $\mathbb{Z}$ is the set of all integers. The NLD problem is an unbounded relaxation of the ML problem—it ignores the symbol bound constraints $-u \leq s_i \leq u$ in the original ML problem, but keeps the symbols s_i in the discrete set $2\mathbb{Z} + 1$. The reason for doing this is to facilitate the use of an efficient processing technique, namely, *lattice reduction*.

In lattice reduction, the channel matrix $\boldsymbol{H}$ is transformed to another channel matrix

$$\tilde{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{U}, \tag{4}$$

where $\tilde{\boldsymbol{H}}$ is called a lattice-reduced channel, and $\boldsymbol{U}$ is a unimodular matrix; i.e., $\boldsymbol{U} \in \mathbb{Z}^{N \times N}$ and $|\det(\boldsymbol{U})| = 1$. Loosely speaking, the transformation (4) is designed such that the transformed channel matrix $\tilde{\boldsymbol{H}}$ may be better conditioned with short and roughly orthogonal column vectors [12], [26]. A popular algorithm for the lattice reduction process (4) is the Lenstra-Lenstra-Lovász (LLL) reduction [11], [12]. Since a unimodular $\boldsymbol{U}$ satisfies $\boldsymbol{U}^{-1}\mathbb{Z}^N = \mathbb{Z}^N$, the symbol vector $\boldsymbol{s}$ can be transformed to $\tilde{\boldsymbol{s}} = \boldsymbol{U}^{-1}\boldsymbol{s}$ and the domain of $\tilde{\boldsymbol{s}}$ is $\tilde{\boldsymbol{s}} \in 2\mathbb{Z}^N + \boldsymbol{U}^{-1}\mathbf{1}$. Subsequently, we can equivalently recast the NLD

problem (3) as

$$\min_{\tilde{s}} \quad \|\boldsymbol{y} - \tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}}\|^2$$
$$\text{s.t.} \quad \tilde{\boldsymbol{s}} \in 2\mathbb{Z}^N + \boldsymbol{U}^{-1}\mathbf{1}, \tag{5}$$

which is also a lattice decoding problem, but with a "better" channel $\tilde{\boldsymbol{H}}$. The equivalent lattice decoding problem (5) is then solved by applying a sphere decoder for unbounded integers. Empirical and theoretical results in [7] and [10], [13] show that lattice reduction can boost the speed for a sphere decoder to solve problem (5). Alternatively, one may apply low-complexity methods to approximate problem (5). For example, one may use an inexact sphere decoder that stops decoding once the complexity reaches some prefixed runtime limit [27]. Another well-known approximation is the simple decision feedback (DF) method, which leads to the zero-forcing (ZF) lattice reduction aided (LRA)-DF detector [14], [15].

C. Regularized Lattice Decoding

While lattice reduction provides an attractive way to handle the NLD problem (3), it also transforms the symbol bound set $\{\boldsymbol{s} \mid -u \leq s_i \leq u, \ \forall i\}$ to a complicated polyhedron $\{\tilde{\boldsymbol{s}} = \boldsymbol{U}^{-1}\boldsymbol{s} \mid -u \leq s_i \leq u, \ \forall i\}$. There is no known way for a sphere decoder to efficiently manage such a polyhedron bound constraint, and, for this reason, the latter is ignored in NLD. As a consequence, NLD may output an out-of-bound symbol decision. Unfortunately, the out-of-bound symbol events can be detrimental to the system error rate performance. It is shown that NLD fails to achieve the optimal diversity-multiplexing tradeoff (DMT) in general MIMO system models [6], [8]. This drawback has motivated endeavors that study a regularized version of the NLD problem [6], [9]

$$\min_{\boldsymbol{s}} \quad \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T\boldsymbol{T}\boldsymbol{s}$$
$$\text{s.t.} \quad \boldsymbol{s} \in 2\mathbb{Z}^N + \mathbf{1}, \tag{6}$$

where $\boldsymbol{T}$ is a positive-semidefinite regularization matrix. The addition of the regularization term $\boldsymbol{s}^T\boldsymbol{T}\boldsymbol{s}$ penalizes symbol vectors $\boldsymbol{s}$ that are far away from the origin, thereby attempting to constrain the optimal solutions of the regularized lattice decoding problem (6) within the symbol bounds in an implicit manner. The analysis in [9] shows that regularization can alleviate the negative effect of having no explicit symbol bound constraints, and the regularized lattice decoding and its associated LRA methods with a positive-definite $\boldsymbol{T}$ can achieve the same DMT as the true ML detector. A well-known choice of $\boldsymbol{T}$ is the MMSE regularization matrix $\boldsymbol{T} = \sigma_\nu^2/\sigma_s^2\boldsymbol{I}$, where σ_s^2 and σ_ν^2 denote the symbol and noise variances, respectively, and $\boldsymbol{I}$ is the identity matrix. The corresponding lattice decoding and LRA methods are called MMSE LD and MMSE LRA-DF. However, this popular MMSE regularization is the only available

choice of regularizations in the existing literature. Little attention has been paid to the optimization of the regularization matrix $\boldsymbol{T}$ for the purpose of improving the lattice decoding performance. In addition, the best regularization in the sense of approximating the ML problem is unknown in the literature.

III. LAGRANGIAN DUAL ML RELAXATION

In this section, we consider a Lagrangian dual relaxation (LDR) formulation of the ML detection problem (2). An important motivation behind our endeavor is that LDR can provide us with the tightest approximation to the ML problem in a Lagrangian sense. Hence, by studying the LDR problem, we may be able to derive an approximate ML algorithm that yields good solution accuracy. Also, we will see that the LDR formulation shows relationship to regularized lattice decoding.

A. The LDR Formulation

Let us rewrite the ML problem (2) as

$$
\min_{\boldsymbol{s} \in 2\mathbb{Z}^N + \boldsymbol{1}} \quad \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2
$$
$$
\text{s.t.} \quad s_i^2 \leq u^2, \; i = 1, \ldots, N. \tag{7}
$$

Here, it is important to note that we define the problem domain of (7) as the discrete set $2\mathbb{Z}^N + \boldsymbol{1}$. Such an attempt is significantly different from that in many existing relaxed ML MIMO detection methods, such as semidefinite relaxation [22], [23], [28]–[30], where the problem domain is often the N-dimensional real space $\mathbb{R}^N$. Our goal is to derive the Lagrangian dual problem of the above ML formulation. Let

$$
\mathcal{L}(\boldsymbol{s}, \boldsymbol{\lambda}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T \left(\boldsymbol{s}^2 - u^2 \boldsymbol{1}\right) \tag{8}
$$

denote the Lagrangian function of (7), where $\boldsymbol{\lambda} \succeq \boldsymbol{0}$ is the Lagrangian dual variable for the constraints $s_i^2 \leq u^2$ (note that the notation $\boldsymbol{\lambda} \succeq \boldsymbol{0}$ means that $\boldsymbol{\lambda}$ is elementwise non-negative), and $\boldsymbol{s}^2$ denotes the elementwise square of $\boldsymbol{s}$. The Lagrangian dual problem of the ML problem (7) is, by definition, given by

$$
\max_{\boldsymbol{\lambda}} \quad d(\boldsymbol{\lambda})
$$
$$
\text{s.t.} \quad \boldsymbol{\lambda} \succeq \boldsymbol{0}, \tag{9}
$$

where $d(\boldsymbol{\lambda}) = \min_{\boldsymbol{s} \in 2\mathbb{Z}^N + \boldsymbol{1}} \mathcal{L}(\boldsymbol{s}, \boldsymbol{\lambda})$ is the dual function associated with (8), which can be expressed as

$$
d(\boldsymbol{\lambda}) = \varphi(\boldsymbol{\lambda}) - u^2 \boldsymbol{\lambda}^T \boldsymbol{1}, \tag{10}
$$

$$
(\Phi_{\boldsymbol{\lambda}}) \quad \varphi(\boldsymbol{\lambda}) = \min_{\boldsymbol{s} \in 2\mathbb{Z}^N + \boldsymbol{1}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T \boldsymbol{D}(\boldsymbol{\lambda})\boldsymbol{s}, \tag{11}
$$

with $\boldsymbol{D}(\boldsymbol{\lambda})$ denoting a diagonal matrix whose ith diagonal is λ_i. For convenience, we will call problem (9) the *LDR problem*.

At this point, we see several interesting observations. First, problem $(\Phi_{\boldsymbol{\lambda}})$, as a constituent component of the dual function, is a diagonally regularized lattice decoding problem. Hence, the above formulated LDR problem exhibits relation to regularized lattice decoding. More connections between LDR and regularized lattice decoding will be revealed later. Second, the naive and MMSE lattice decoders can be seen as particular instances of problem $(\Phi_{\boldsymbol{\lambda}})$. Specifically, NLD chooses $\boldsymbol{\lambda} = \boldsymbol{0}$, while MMSE lattice decoding $\boldsymbol{\lambda} = \sigma_\nu^2/\sigma_s^2 \boldsymbol{1}$. Third, the LDR problem (9) can be regarded as that of finding the best regularization vector $\boldsymbol{\lambda}$ among all the diagonally regularized lattice decoding instances to approximate the ML problem.

B. Optimality and Duality Gap Analysis

In this subsection, we analyze the optimality conditions of the LDR formulation in (9)-(11), that is, conditions under which the LDR problem exactly solves the ML problem. Moreover, we also study the approximation quality of LDR by analyzing the duality gap $f^\star - d^\star$, where

$$f^\star = \min_{\boldsymbol{s} \in 2\mathbb{Z}^N+\boldsymbol{1},\ \boldsymbol{s}^2 \preceq u^2\boldsymbol{1}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 \tag{12}$$

$$d^\star = \max_{\boldsymbol{\lambda} \succeq \boldsymbol{0}} d(\boldsymbol{\lambda}) \tag{13}$$

denote the optimal objective values of the ML problem (7) and the LDR problem (9), respectively. In particular, a smaller $f^\star - d^\star$ would indicate better approximation accuracy, and zero $f^\star - d^\star$ means ML being achieved. Note that $f^\star - d^\star \geq 0$ by weak duality (see, e.g., [31]).

We first present a simple result based on a connection between NLD and LDR.

Fact 1. *Let*

$$\hat{\boldsymbol{s}}_{\text{NLD}} \in \arg\min_{\boldsymbol{s} \in 2\mathbb{Z}^N+\boldsymbol{1}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2$$

be an optimal solution of the NLD problem (3). Consider instances where $[\hat{\boldsymbol{s}}_{\text{NLD}}]_i^2 \leq u^2$ for all $i = 1, \ldots, N$. Then, the following statements hold:

1. *$\hat{\boldsymbol{s}}_{\text{NLD}}$ is an optimal solution of the ML problem (7).*

2. *Strong duality, or $f^\star - d^\star = 0$, holds for the LDR problem. Also, $\boldsymbol{\lambda} = \boldsymbol{0}$ is an optimal solution of the LDR problem (9).*

3. *For any optimal solution $\boldsymbol{\lambda}^\star$ of the LDR problem (9), $\hat{\boldsymbol{s}}_{\text{NLD}}$ is an optimal solution of problem $(\Phi_{\boldsymbol{\lambda}})$ for $\boldsymbol{\lambda} = \boldsymbol{\lambda}^\star$.*

The proof of Fact 1 is given in Appendix A. The idea behind is to exploit the fact that NLD is a special case of LDR. Fact 1 implies that for instances where NLD is ML-optimal, LDR is also ML-optimal. Hence, we should expect that LDR would perform better than NLD—this will be shown to be true by simulations later.

Next, we consider another optimality result. The following lemma will be needed.

Lemma 1. *Let $\hat{s}_{\boldsymbol{\lambda}}$ be an optimal solution of problem $(\Phi_{\boldsymbol{\lambda}})$. Suppose that $\boldsymbol{\lambda}$ satisfies*

$$\lambda_i > \gamma_m \tag{14}$$

for some $i \in \{1, \ldots, N\}$, where m is an odd positive integer,

$$\gamma_m = \frac{c(\boldsymbol{y}, \boldsymbol{H})}{(m+2)^2 - 1}, \tag{15}$$

$$c(\boldsymbol{y}, \boldsymbol{H}) = \min_{\boldsymbol{s} \in \{\pm 1\}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 - \min_{\boldsymbol{s} \in 2\mathbb{Z}^N + 1} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2. \tag{16}$$

Then it must hold true that

$$[\hat{s}_{\boldsymbol{\lambda}}]_i^2 \leq m^2.$$

The proof of Lemma 1 is shown in Appendix B. Lemma 1 is not only useful in establishing an optimality condition of LDR, as we will see, but it is also of independent interest as discussed in the following remark.

Remark 1: Intuitively, the idea of incorporating regularization in lattice decoding is based on the belief that regularization can pull the regularized lattice decoding solution within the symbol bounds. Lemma 1 provides a theoretical justification that this intuitive belief is indeed true. It quantifies, in a sufficient manner, how much regularization is needed to achieve a desired symbol bound constraint level in the regularized lattice decoding problem $(\Phi_{\boldsymbol{\lambda}})$.

Let us now turn our attention back to the optimality analysis. From Lemma 1, we have the following observation: For any $\boldsymbol{\lambda} \succ \gamma_u \mathbf{1}$, an optimal solution $\hat{s}_{\boldsymbol{\lambda}}$ of problem $(\Phi_{\boldsymbol{\lambda}})$ always satisfies $\hat{s}_{\boldsymbol{\lambda}}^2 \preceq u^2 \mathbf{1}$, which means that $\hat{s}_{\boldsymbol{\lambda}}$ is a feasible solution of the ML problem. Hence, we would hope that such a $\boldsymbol{\lambda}$ and the corresponding $\hat{s}_{\boldsymbol{\lambda}}$ are optimal to the LDR problem and the ML problem, respectively. Remarkably, we show that this is indeed true for the case of $u = 1$.

Theorem 1. *Consider $u = 1$, or the 2-PAM case. In this case, strong duality $f^\star - d^\star = 0$ always holds. In particular, any $\boldsymbol{\lambda} \succ \gamma_1 \mathbf{1}$ is an optimal solution of the LDR problem* (9)*, and an optimal solution $\hat{s}_{\boldsymbol{\lambda}}$ of problem $(\Phi_{\boldsymbol{\lambda}})$ for any $\boldsymbol{\lambda} \succ \gamma_1 \mathbf{1}$ is also an optimal solution of the ML problem* (7)*.*

Readers are referred to Appendix C for the proof of Theorem 1. Theorem 1 indicates that in the 2-PAM constellation case, the LDR problem is ML-optimal for *any* given realization $(\boldsymbol{y}, \boldsymbol{H})$. This further implies that in principle, one can use regularized lattice decoding, which has no explicit symbol bound constraints, to solve the ML problem for the 2-PAM case.

An interesting question is whether the proof we use in Theorem 1 can be extended to the more general case of $u \geq 3$. We found that such extension is possible. Unfortunately, the result obtained becomes a loose bound on the duality gap, rather than strong duality. It also fails to cover the optimality condition in Fact 1. Herein, we give another analysis result that links to another method, *bound-constrained semidefinite relaxation* (BC-SDR) [22]. BC-SDR is a representative method in the class of convex relaxation techniques for the ML problem. It approximates the ML problem by considering the following convex optimization problem

$$g^\star = \min_{\boldsymbol{S} \in \mathbb{S}^N, \boldsymbol{s} \in \mathbb{R}^N} \quad \text{tr}(\boldsymbol{H}^T \boldsymbol{H} \boldsymbol{S}) - 2\boldsymbol{s}^T \boldsymbol{H}^T \boldsymbol{y} + \|\boldsymbol{y}\|^2$$

$$\text{s.t. } \boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T, \ \boldsymbol{1} \preceq \boldsymbol{d}(\boldsymbol{S}) \preceq u^2 \boldsymbol{1},$$

$$(17)$$

where $\mathbb{S}^N$ denotes the set of all $N \times N$ real symmetric matrices, the notation $\boldsymbol{A} \succeq \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite (PSD), the operator $\boldsymbol{d}(\boldsymbol{S}) = [\ S_{1,1}, S_{2,2}, \ldots, S_{N,N}\ ]^T \in \mathbb{R}^N$ takes the diagonal elements of $\boldsymbol{S}$ to form a vector, and $g^\star$ denotes the optimal objective value of BC-SDR. We concisely review the idea of BC-SDR as follows. Let $\boldsymbol{S} = \boldsymbol{s}\boldsymbol{s}^T$. The condition $\boldsymbol{S} = \boldsymbol{s}\boldsymbol{s}^T$ is equivalent to $\boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T$, rank$(\boldsymbol{S}) = 1$. Also, we have $S_{i,i} \in \{1, 3^2, \ldots, u^2\}$ for all i. BC-SDR relaxes $S_{i,i} \in \{1, 3^2, \ldots, u^2\}$ to $1 \leq S_{i,i} \leq u^2$ and allows $\boldsymbol{S}$ to take any rank, thereby obtaining the convex relaxation in (17). A key reason why BC-SDR is representative is that BC-SDR is shown to be equivalent to two other relaxations, namely, the polynomial-inspired SDR (PI-SDR) [28] and virtually-antipodal SDR [29], which employ different ideas to relax and offer different insights in ML approximation; see [30] for details. For example, there are theoretically proven results on the approximation accuracy of VA-SDR [32], [33], and those results apply to BC-SDR (and also PI-SDR) by using the equivalence of the three SDRs. Returning to the problem here, one can observe that LDR and BC-SDR are quite different from one another; the former and latter use discrete and continuous problem domains, respectively. In the following theorem, we provide a connection between LDR and BC-SDR.

Theorem 2. *For any realization $(\boldsymbol{y}, \boldsymbol{H})$ and any $u \geq 1$, the duality gap of LDR is better than or equal to that of BC-SDR:*

$$f^\star - d^\star \leq f^\star - g^\star. \tag{18}$$

The proof of Theorem 2 is relegated to Appendix D. Theorem 2 is meaningful in establishing a relationship of the approximation qualities of LDR and BC-SDR. Specifically, it implies that LDR performs no worse than BC-SDR in terms of relaxation tightness. In addition, it is possible for LDR to yield strictly better duality gap than BC-SDR. For example, for the case of $u = 1$, we have shown in Theorem 1 that strong duality always holds for LDR. However, BC-SDR does not guarantee strong duality even for $u = 1$, as indicated in previous work [34], [32]. Also, owing to the equivalence of BC-SDR, PI-SDR and VA-SDR, the same conclusion applies to PI-SDR and VA-SDR.

C. Practical Realization of LDR via Projected Subgradient

We turn our attention to the realization of LDR. By Lagrangian duality theory [31], the LDR problem (9) is a convex optimization problem. In particular, its objective function $d(\boldsymbol{\lambda})$ is concave. However, this does not mean that $d(\boldsymbol{\lambda})$ is easy to maximize. From (10), we see that $d(\boldsymbol{\lambda})$ involves a minimization problem, namely, problem $(\Phi_{\boldsymbol{\lambda}})$ in (11). Thus, $d(\boldsymbol{\lambda})$ is in general a nondifferentiable function. Our optimization strategy is to employ the projected subgradient (PS) method [35], which is a convenient approach for solving nondifferentiable convex optimization problems.

The PS method for the LDR problem (9) is described as follows. Let $\boldsymbol{\lambda}^{(k)}$ denote the iterate generated by the PS method at the kth iteration. Given an initialization $\boldsymbol{\lambda}^{(1)}$, the iterates are recursively generated via

$$\boldsymbol{\lambda}^{(k+1)} = \left[\boldsymbol{\lambda}^{(k)} + \alpha_k \boldsymbol{g}^{(k)} \right]^+, \quad k = 1, 2, \ldots \tag{19}$$

where $\boldsymbol{g}^{(k)}$ denotes a subgradient of $d(\boldsymbol{\lambda})$ at $\boldsymbol{\lambda}^{(k)}$, $\{\alpha_k\}$ is a step-size sequence which is predetermined, and $[\boldsymbol{\lambda}]^+$ denotes the projection of its input $\boldsymbol{\lambda} \in \mathbb{R}^N$ onto the set of N-dimensional nonnegative vectors $\mathbb{R}^N_+$. The projection operator $[\boldsymbol{\lambda}]^+$ has a closed form; specifically, if we let $\boldsymbol{\mu} = [\boldsymbol{\lambda}]^+$, then $\mu_i = \max\{0, \lambda_i\}$ for all i. Using basic subgradient calculus results [35], the subgradient $\boldsymbol{g}^{(k)}$ is shown to be

$$\boldsymbol{g}^{(k)} = (\boldsymbol{s}^{(k)})^2 - u^2 \mathbf{1}, \tag{20}$$

where $\boldsymbol{s}^{(k)}$ is a solution of the regularized lattice decoding problem $(\Phi_{\boldsymbol{\lambda}})$ for $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{(k)}$; i.e.,

$$\boldsymbol{s}^{(k)} = \arg \min_{\boldsymbol{s} \in 2\mathbb{Z}^N + 1} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T \boldsymbol{D}(\boldsymbol{\lambda}^{(k)})\boldsymbol{s}. \tag{21}$$

We should discuss the convergence of the PS method. It is known in the optimization literature [24], [35] that under a few fairly mild assumptions, the PS method is guaranteed to converge to the optimal objective value, which is $d^\star$ here. For the LDR problem here, we can even pin down a simplified sufficient assumption for convergence—the PS method can achieve convergence to the optimal dual value $d^\star$ for

any full column rank channel matrix $\boldsymbol{H}$. A complete description for the PS convergence results mentioned above will be provided in Remark 3. Also, practical convergence issues will be discussed in Remark 4.

To summarize, we can solve the LDR problem by using the iterative PS procedure in (19)-(21). In particular, at each iteration, we need to solve the regularized lattice decoding problem in (21), and then use its solution $\boldsymbol{s}^{(k)}$ to update the regularization vector at the next iteration, $\boldsymbol{\lambda}^{(k+1)}$. At first, this may sound computationally more expensive than a one-shot regularized lattice decoding method such as MMSE LD. However, we will illustrate by simulations that with a careful initialization and implementation, a PS-based LDR detector can be computationally comparable to the MMSE LD detector. Also, as a practical alternative, we can consider efficient approximation schemes where we use inexact sphere decoding or LRA-DF detection in place of an exact solver for problem (21). The PS procedure provides interesting insight as described in the following remark.

Remark 2: Physically, the PS procedure above can be interpreted as a recursively regularized lattice decoding method wherein some form of adaptive symbol bound control is performed. To explain this, let $\lambda_i^{(k)}$ and $s_i^{(k)}$ denote the ith elements of $\boldsymbol{\lambda}^{(k)}$ and $\boldsymbol{s}^{(k)}$, respectively. Then, we can see from (19) and (20) that if $(s_i^{(k)})^2 > u^2$, then $\lambda_i^{(k+1)}$ will be increased. Likewise, if $(s_i^{(k)})^2 < u^2$, then $\lambda_i^{(k+1)}$ will be decreased. This means that if some symbols violate the symbol bound constraints, then the PS method at the next iteration will increase the regularization variables for those symbols. Subsequently, those larger regularization variables will tend to pull the associated symbols toward the origin at the next iteration, or set a more stringent upper bound on those symbols, as suggested in Lemma 1. Similarly, for symbols lying strictly within the symbol bounds, the corresponding regularization variables will be reduced at the next iteration.

We should also give some discussions on the convergence of the PS LDR procedure.

Remark 3: Theoretically, the optimal convergence of the PS LDR procedure is usually guaranteed. To discuss this in more precise terms, we need to describe one available convergence result [24], [35]: Let $d_{\text{best}}^{(k)} = \max\{d_{\text{best}}^{(k-1)}, d(\boldsymbol{\lambda}^{(k)})\}$. For certain types of stepsize rules, such as the diminishing stepsize rule [2], we have $\lim_{k\to\infty} d_{\text{best}}^{(k)} = d^\star$ if every subgradient $\boldsymbol{g}^{(k)}$ is bounded; i.e., there exists a finite G such that

$$\|\boldsymbol{g}^{(k)}\| \leq G, \quad \text{for } k = 1, 2, \ldots$$

In our LDR problem, the boundedness assumption above is the same as requiring every regularized lattice decoding solution $\boldsymbol{s}^{(k)}$ in (21) to be bounded. Intuitively, one would argue that problem (21), or $(\Phi_{\boldsymbol{\lambda}})$,

---

[2]A typical example of the diminishing stepsize rule is $\alpha_k = \eta/\sqrt{k}$ for some fixed positive constant η.

should yield bounded solutions, except for pathological cases. In fact, this can be confirmed under a mild assumption:

Lemma 2. *Given a full column rank channel matrix $\boldsymbol{H}$, any optimal solution $\hat{\boldsymbol{s}}_{\boldsymbol{\lambda}}$ of problem $(\Phi_{\boldsymbol{\lambda}})$ for any $\boldsymbol{\lambda} \succeq \boldsymbol{0}$ is bounded.*

The proof of Lemma 2 is given in Appendix E. As a corollary of Lemma 2, the PS LDR procedure is theoretically guaranteed to converge to the optimal value for full column rank (and thus overdetermined) channels.

Remark 4: While the PS method provides an effective strategy to cope with certain difficult nondifferentiable convex optimization problems, such as the LDR problem here, it is also known to yield slow convergence in some applications. Despite this setback, the PS convergence speed may be improved if a good initialization can be found. Fortunately, such initialization seems to be available for LDR, as our extensive simulations have found. In Section IV-C, we will propose an initialization scheme that requires solving another (convex) optimization problem, but can significantly improve the convergence speed.

While the PS LDR realization procedure described above looks quite straightforward, there are fine details on how we can implement the method in a numerically efficient manner. This will be the focus of the next section.

IV. PRACTICAL IMPLEMENTATION

In this section, we elaborate on how we implement the PS LDR method.

A. Lattice Decoding for Problem (21)

The core issue with implementing the PS LDR method lies in finding the solution of the regularized lattice decoding problem (21) at each iteration. One can directly handle this issue by seeing each problem (21) as a stand-alone regularized lattice decoding problem, and applying a lattice decoder to obtain its solution. However, we can make the implementation more efficient by utilizing the regularized lattice decoding solution in the previous iteration to improve the solution search process in the present iteration. To facilitate our description, we divide our development into three steps.

1) Step 1. Integer Least Squares Reformulation: Let

$$\boldsymbol{V}_k^T \boldsymbol{V}_k = \boldsymbol{H}^T \boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(k)}) \tag{22}$$

denote a square-root decomposition of $\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(k)})$, where $\boldsymbol{V}_k \in \mathbb{R}^{N \times N}$ is a corresponding square-root factor. By the transformation

$$\boldsymbol{s} = 2\boldsymbol{z} + \boldsymbol{1}, \qquad \boldsymbol{z} \in \mathbb{Z}^N, \tag{23}$$

we can rewrite problem (21) as an integer least squares (LS) problem

$$\boldsymbol{z}^{(k)} = \arg \min_{\boldsymbol{z} \in \mathbb{Z}^N} \|\boldsymbol{f}_k - \boldsymbol{V}_k\boldsymbol{z}\|^2, \tag{24}$$

where $\boldsymbol{f}_k = \frac{1}{2}\boldsymbol{V}_k^{-T}\left(\boldsymbol{H}^T\boldsymbol{y} - \left(\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(k)})\right)\boldsymbol{1}\right)$. Note the relation $\boldsymbol{s}^{(k)} = 2\boldsymbol{z}^{(k)} + \boldsymbol{1}$.

2) Step 2. Lattice Reduction: We reformulate problem (24) to a lattice-reduced form. The popularized LLL algorithm is chosen for lattice reduction, and some of its operational details are concisely described as follows. Let $\boldsymbol{G} \in \mathbb{R}^{N \times N}$ be a matrix to be LLL-reduced. Denote its QR decomposition by $\boldsymbol{G} = \bar{\boldsymbol{Q}}\bar{\boldsymbol{R}}$ where $\bar{\boldsymbol{Q}} \in \mathbb{R}^{N \times N}$ is unitary and $\bar{\boldsymbol{R}} \in \mathbb{R}^{N \times N}$ upper triangular. Given $(\bar{\boldsymbol{Q}}, \bar{\boldsymbol{R}})$, an LLL algorithm finds a 3-tuple $(\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{U})$ such that

$$\boldsymbol{Q}\boldsymbol{R} = \bar{\boldsymbol{Q}}\bar{\boldsymbol{R}}\boldsymbol{U}, \tag{25}$$

$\boldsymbol{Q}\boldsymbol{R}$ is a lattice-reduced matrix of $\bar{\boldsymbol{Q}}\bar{\boldsymbol{R}} = \boldsymbol{G}$, and $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{U} \in \mathbb{R}^{N \times N}$ are unitary, upper triangular and unimodular, respectively. For convenience, we will use the following notation

$$(\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{U}) = \mathsf{LLL}(\bar{\boldsymbol{Q}}, \bar{\boldsymbol{R}}) \tag{26}$$

to represent the LLL process. Returning to our problem, our task is to LLL-reduce the basis matrix $\boldsymbol{V}_k$ in problem (24). This can be done as follows:

$$(\bar{\boldsymbol{Q}}_k, \bar{\boldsymbol{R}}_k) = \mathsf{QR}(\boldsymbol{V}_k), \tag{27a}$$

$$(\boldsymbol{Q}_k, \boldsymbol{R}_k, \boldsymbol{U}_k) = \mathsf{LLL}(\bar{\boldsymbol{Q}}_k, \bar{\boldsymbol{R}}_k), \tag{27b}$$

where QR is a shorthand notation for the QR decomposition process. We can see that $\boldsymbol{Q}_k\boldsymbol{R}_k$ forms an LLL-reduced basis matrix of $\boldsymbol{V}_k$, with $\boldsymbol{U}_k$ being the associated unimodular transformation matrix. By substituting $\boldsymbol{Q}_k\boldsymbol{R}_k = \boldsymbol{V}_k\boldsymbol{U}_k$, and introducing the transformation $\tilde{\boldsymbol{z}} = \boldsymbol{U}_k^{-1}\boldsymbol{z}$, we equivalently turn problem (24) to

$$\tilde{\boldsymbol{z}}^{(k)} = \arg \min_{\tilde{\boldsymbol{z}} \in \mathbb{Z}^N} \|\tilde{\boldsymbol{f}}_k - \boldsymbol{R}_k\tilde{\boldsymbol{z}}\|^2, \tag{28}$$

where $\tilde{\boldsymbol{f}}_k = \boldsymbol{Q}_k^T\boldsymbol{f}_k$.

3) An Alternative Option to Step 2 by Successive LLL Update: We offer an alternative option to Step 2 that can provide computational savings with the LLL reduction process. Essentially, if a given basis matrix $\boldsymbol{G}$ is almost LLL-reduced, then it would generally take fewer number of iterations for the LLL algorithm to complete the process. Now, rather than LLL-reducing $\boldsymbol{V}_k$ as in the previous Step 2, *we consider the LLL reduction of $\boldsymbol{V}_k\boldsymbol{U}_{k-1}$, where $\boldsymbol{U}_{k-1}$ is the LLL unimodular matrix of $\boldsymbol{V}_{k-1}$.* The idea is that if $\boldsymbol{V}_k$ and $\boldsymbol{V}_{k-1}$ are not too different, which is possible when there are only small changes with the PS update (cf. (19)), then $\boldsymbol{V}_k\boldsymbol{U}_{k-1}$ may already be well conditioned in the LLL sense. Note that a similar idea has been considered in a different context, namely, LRA detection under time-correlated fading channels [36]. The proposed successive LLL update process is described as follows. We replace (27) by

$$(\bar{\boldsymbol{Q}}_k, \bar{\boldsymbol{R}}_k) = \mathsf{QR}(\boldsymbol{V}_k\boldsymbol{U}_{k-1}), \tag{29a}$$

$$(\boldsymbol{Q}_k, \boldsymbol{R}_k, \tilde{\boldsymbol{U}}_k) = \mathsf{LLL}(\bar{\boldsymbol{Q}}_k, \bar{\boldsymbol{R}}_k), \tag{29b}$$

$$\boldsymbol{U}_k = \boldsymbol{U}_{k-1}\tilde{\boldsymbol{U}}_k. \tag{29c}$$

It can be verified that $\boldsymbol{Q}_k\boldsymbol{R}_k$ is a lattice-reduced matrix of $\boldsymbol{V}_k$, and that $\boldsymbol{U}_k$ is the corresponding unimodular transformation matrix. Thus, the equivalent formulation in (28) applies. We should note that there is a more efficient way to compute (29). Let $\mathsf{chol}(\cdot)$ denote the Cholesky decomposition operation; i.e., $\boldsymbol{W} = \mathsf{chol}(\boldsymbol{A}) \iff \boldsymbol{W}^T\boldsymbol{W} = \boldsymbol{A}, \boldsymbol{W}$ upper triangular. It can be shown that (29a)-(29b) can be equivalently implemented by

$$\boldsymbol{W}_k = \mathsf{chol}(\boldsymbol{U}_{k-1}^T(\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(k)}))\boldsymbol{U}_{k-1}), \tag{30a}$$

$$(\boldsymbol{Q}_k, \boldsymbol{R}_k, \tilde{\boldsymbol{U}}_k) = \mathsf{LLL}(\boldsymbol{I}, \bar{\boldsymbol{W}}_k). \tag{30b}$$

In particular, by using (30), we do not need to compute $\boldsymbol{V}_k$, which requires a square-root decomposition operation (cf. (22)).

4) Step 3. Applying a Sphere Decoder: As the last step, we solve problem (28) by a sphere decoding algorithm. Note that problem (28), which has its equivalent channel matrix $\boldsymbol{R}_k$ being upper triangular, already takes a standard problem form for a sphere decoding algorithm to run. Sphere decoding considers a within-sphere point search process; roughly speaking, the latter may be described by

$$\begin{aligned} \min_{\tilde{\boldsymbol{z}}\in\mathbb{Z}^N} \quad & \|\tilde{\boldsymbol{f}}_k - \boldsymbol{R}_k\tilde{\boldsymbol{z}}\|^2 \\ \text{s.t.} \quad & \|\tilde{\boldsymbol{f}}_k - \boldsymbol{R}_k\tilde{\boldsymbol{z}}\|^2 \leq C^{(k)} \end{aligned} \tag{31}$$

for some given squared sphere radius $C^{(k)}$. We refer readers to the literature, e.g., [4], [7], for more complete descriptions of the within-sphere point processes used in various sphere decoding algorithms. In practice, it is generally found that a well chosen sphere radius may significantly narrow down the within-sphere search space, or the feasible set of (31), thereby making the optimal point search more efficient [7], [10]. Here, we make use of the result in the previous LDR iteration to set the sphere radius, namely, by

$$C^{(k)} = \| \boldsymbol{f}_k - \boldsymbol{V}_k \boldsymbol{z}^{(k-1)} \|^2. \tag{32}$$

The rationale is that if the previous iterate is LDR-optimal, or near LDR-optimal, then the sphere radius choice above may eliminate a substantial number of points that are not necessary to visit.

B. Putting Together the Algorithm

By plugging the above lattice decoding steps into the PS procedure in (19)-(21), we construct a complete LDR algorithm. A pseudo-code form description of the algorithm is given in Algorithm 1. In the algorithm, $\mathcal{Q}$ denotes the elementwise quantization function to the symbol constellation set $\{\pm 1, \pm 3, \ldots \pm u\}$, and $\lfloor \cdot \rceil$ denotes the elementwise integer rounding function. Also, note that we adopt the successive LLL update method (Steps 5-8). In the sequel, we will call the resultant detector the *LDR lattice decoding (LD) detector*.

As previously mentioned, we can also consider variations of the LDR LD detector where efficient approximation schemes are used in place of the exact sphere decoder. One alternative is to force the sphere decoding algorithm to terminate when its number of nodes visited exceeds a prescribed limit—this leads to an inexact runtime-limited LDR LD detector. Another alternative is to apply the LRA-DF method [15], [37], or equivalently, the Babai's nearest plane algorithm [26], to problem (28). To be specific, we generate a point, denoted by $\tilde{\boldsymbol{z}}_{\mathsf{DF}}^{(k)} \in 2\mathbb{Z}^N + \mathbf{1}$, via

$$\tilde{z}_{\mathsf{DF},i}^{(k)} = \left\lfloor [\tilde{\boldsymbol{f}}_k]_i - \sum_{j=i+1}^{N} [\boldsymbol{R}_k]_{ij} \tilde{z}_{\mathsf{DF},j}^{(k)} \right\rceil, \quad \text{for } i = N, N-1, \ldots, 1. \tag{33}$$

We will name the subsequent detector the *LDR LRA-DF detector*.

C. Box Relaxation as an Initialization

Like many other iterative optimization methods, the PS LDR method may exhibit fast convergence if a good starting point $\boldsymbol{\lambda}^{(1)}$ is given. For the initialization aspect, a simple and logical choice is to employ the MMSE regularization; i.e., $\boldsymbol{\lambda}^{(1)} = \sigma_\nu^2 / \sigma_s^2 \mathbf{1}$. By our empirical experience, the MMSE initialization

---

Algorithm 1 The LDR LD Detector

Input a problem instance $(\boldsymbol{y}, \boldsymbol{H})$, a starting point $\boldsymbol{\lambda}^{(1)} \succeq \boldsymbol{0}$, a step-size sequence $\{\alpha_k\}$, and stopping

    parameters K_{max}, ϵ;

1: $k = 1$

2: $\boldsymbol{U}^{(0)} = \boldsymbol{I}$

3: $\tilde{\boldsymbol{z}}^{(0)} = \left\lfloor \frac{1}{2}((\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(1)}))^{-1}\boldsymbol{H}^T\boldsymbol{y} + \mathbf{1}) \right\rceil$

4: **repeat**

5:     $\boldsymbol{W}_k = \mathsf{chol}(\boldsymbol{U}_{k-1}{}^T(\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(k)}))\boldsymbol{U}_{k-1})$

6:     $\boldsymbol{f}_k = \frac{1}{2}\boldsymbol{W}_k{}^{-T}\boldsymbol{U}_{k-1}{}^T(\boldsymbol{H}^T\boldsymbol{y} - (\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda}^{(k)}))\mathbf{1})$

7:     $(\boldsymbol{Q}_k, \boldsymbol{R}_k, \tilde{\boldsymbol{U}}_k) = \mathsf{LLL}(\boldsymbol{I}, \boldsymbol{W}_k)$

8:     $\boldsymbol{U}_k = \boldsymbol{U}_{k-1}\tilde{\boldsymbol{U}}_k$

9:     $\tilde{\boldsymbol{f}}_k = \boldsymbol{Q}_k{}^T\boldsymbol{f}_k$

10:     $C^{(k)} = \|\boldsymbol{f}_k - \boldsymbol{W}_k\tilde{\boldsymbol{z}}^{(k-1)}\|^2$

11:     run a sphere decoding algorithm, with $C^{(k)}$ as the squared sphere radius, to solve for

$$\tilde{\boldsymbol{z}}^{(k)} = \arg\min_{\tilde{\boldsymbol{z}} \in \mathbb{Z}^N} \|\tilde{\boldsymbol{f}}_k - \boldsymbol{R}_k\tilde{\boldsymbol{z}}\|^2$$

12:     $\boldsymbol{z}^{(k)} = \boldsymbol{U}_k\tilde{\boldsymbol{z}}^{(k)}$

13:     $\boldsymbol{s}^{(k)} = 2\boldsymbol{z}^{(k)} + \mathbf{1}$

14:     $\boldsymbol{g}^{(k)} = \left(\boldsymbol{s}^{(k)}\right)^2 - u^2\mathbf{1}$

15:     $\boldsymbol{\lambda}^{(k+1)} = \left[\boldsymbol{\lambda}^{(k)} + \alpha_k\boldsymbol{g}^{(k)}\right]^+$

16:     $k = k + 1$

17: **until** $k > K_{\mathsf{max}}$ or $\|\boldsymbol{\lambda}^{(k)} - \boldsymbol{\lambda}^{(k-1)}\| \leq \epsilon$

18: Find $\ell = \arg\min_{i=1,2,\ldots,k-1} \|\boldsymbol{y} - \boldsymbol{H}\mathcal{Q}(\boldsymbol{s}^{(i)})\|$

Output $\hat{\boldsymbol{s}} = \mathcal{Q}(\boldsymbol{s}^{(\ell)})$.

---

scheme can indeed lead to reasonable improvement in PS convergence speed. But we also found a better

initialization scheme based on our extensive numerical study. The idea is to consider a further relaxation

of the LDR problem (9)

$$d^\star \geq \max_{\boldsymbol{\lambda} \succeq \boldsymbol{0}} \left\{ \min_{\boldsymbol{s} \in \mathbb{R}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T(\boldsymbol{s}^2 - u^2\mathbf{1}) \right\}, \tag{34}$$

where we relax the original problem domain of $\boldsymbol{s}$ from $2\mathbb{Z}^N + \mathbf{1}$ to $\mathbb{R}^N$. To be specific, we aim at using

the solution of the outer part of problem (34) to initialize the PS LDR method. It is interesting to note

that problem (34) is related to a conventional MIMO detection method. Consider the following problem

$$\min_{\boldsymbol{s} \in \mathbb{R}^N} \ \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2$$
$$\text{s.t. } s_i^2 \leq u^2, \qquad i = 1, \ldots, N,$$

(35)

which is a continuous *box relaxation* (BR) of the ML problem (7) and has previously been studied in the context of multiuser detection [38], [39]. It can be shown that problem (34) is equivalent to problem (35)—the equivalence is in the sense that the dual of the BR problem (35) takes exactly the same form as problem (34), and that strong duality holds owing to the fact that the BR problem (35) is convex and satisfies Slater's condition [31]. Hence, the initialization scheme based on (34) may be seen as using the BR method to warm-start the PS LDR method.

The BR initialization scheme proposed above does not have a closed form in general. In order to implement the BR initialization scheme efficiently, we need to build a low-complexity solver for problem (34). Our solution approach consists of two steps: First, we solve the BR problem (35), which has a simple structure and for which it is relatively easy to find an efficient optimization algorithm. Second, by utilizing the strong duality relationship of problems (34) and (35), we construct a solution to problem (34) from the BR solution. For the first step, we custom-build an optimization algorithm for the BR problem (35); the algorithm is mainly based on the active set method in [40], with a modification to accelerate the LS procedure inside by the one-column updating method [41]. The details are heavily numerical, and are skipped due to limit of space. For the second step, we consider the Karush-Kuhn-Tucker (KKT) conditions of the BR problem, which is shown to be

$$\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{s} + \boldsymbol{D}(\boldsymbol{\lambda})\boldsymbol{s} - \boldsymbol{H}^T\boldsymbol{y} = \boldsymbol{0},$$ (36a)

$$\lambda_i(s_i^2 - u^2) = 0, \quad i = 1, \ldots, N,$$ (36b)

$$\lambda_i \geq 0, s_i^2 - u^2 \leq 0, \quad i = 1, \ldots, N.$$ (36c)

Since (36a)-(36c) is the necessary and sufficient conditions for $(\boldsymbol{s}, \boldsymbol{\lambda})$ to be optimal to problems (34) and (35), we plug the BR solution $\boldsymbol{s}$ obtained in the first step into (36a)-(36b) to find the corresponding optimal $\boldsymbol{\lambda}$. The resultant solution is shown to be

$$\lambda_i = \begin{cases} -\dfrac{[\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{s} - \boldsymbol{H}^T\boldsymbol{y}]_i}{s_i}, & \text{if } s_i \neq 0 \\ 0, & \text{if } s_i = 0 \end{cases}$$ (37)

for $i = 1, \ldots, N$. To summarize, the BR initialization scheme works by running a custom-built active set algorithm to find a solution $\boldsymbol{s}$ of the BR problem (35), substituting the obtained $\boldsymbol{s}$ into (37) to construct a solution $\boldsymbol{\lambda}$ of problem (34), and using the obtained $\boldsymbol{\lambda}$ as a starting point of the PS LDR method.

V. SIMULATIONS

Simulations were conducted to study the symbol error rate (SER) and complexity performance of the proposed LDR MIMO detection approach. We consider a standard complex-valued QAM MIMO scenario

$$\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{\nu}_C, \tag{38}$$

where the channel matrix $\boldsymbol{H}_C \in \mathbb{C}^{M_C \times N_C}$ follows an elementwise i.i.d. complex circular Gaussian distribution with zero mean and unit variance, the symbol vector $\boldsymbol{s}_C \in \mathbb{C}^{N_C}$ is elementwise i.i.d. uniformly distributed with each element $[\boldsymbol{s}_C]_i$ drawn from the standard $(u+1)^2$-QAM constellation set, $\sigma_{\nu_C}^2 \in \mathbb{C}^{M_C}$ is additive white complex circular Gaussian noise with zero mean and variance $\sigma_{\nu_C}^2$, and M_C and N_C denotes the output and input problem sizes, respectively. The model (38) can be equivalently represented by the real-valued model (1) (see, e.g., [30, Eqn. (2)]) with problem size $(M, N) = (2M_C, 2N_C)$ and PAM size $u + 1$. The SNR is defined as $\mathrm{SNR} = \mathrm{E}[\|\boldsymbol{H}_C \boldsymbol{s}_C\|^2]/\mathrm{E}[\|\boldsymbol{\nu}_C\|^2] = N_C \sigma_{s_C}^2 / \sigma_{\nu_C}^2$, where $\sigma_{s_C}^2$ is the variance of the elements of $\boldsymbol{s}_C$.

The benchmarked algorithms are the ML sphere decoding (SD) detector, the MMSE LD detector, the NLD detector, runtime-limited inexact implementations of the aforementioned SD and LD detectors, and the BC-SDR detector (cf. problem (17) and [22]). The ML SD detector is implemented by the Schnorr-Euchner enumeration-based SD algorithm in [4, Algorithm 2]. The MMSE LD detector is also implemented by the same SD algorithm, and its lattice reduction process is LLL reduction. The LLL reduction algorithm we employed follows the pseudo-code description in [37]. The inexact MMSE LD detector is a variation of the MMSE LD detector where we force the SD algorithm to terminate when the number of nodes visited exceeds a prescribed worst-case limit, denoted by N_{node} here. We set $N_{\mathrm{node}} = \min\{N^3, N \times \mathrm{SNR}\}$ (SNR is in linear scale). The same applies to inexact ML SD and inexact NLD. For the BC-SDR detector, Gaussian randomization rounding [28] is employed and the number of randomizations is 64.

The settings of the proposed LDR LD detector, inexact LDR LD detector and LDR LRA-DF detector are as follows. Unless specified, the BR initialization scheme in Section IV-C is used to generate the starting point. The stopping parameters of the PS iterations are set as $K_{\mathsf{max}} = 5$ and $\epsilon = 10^{-9}$; cf. Algorithm 1, line 17. We use the same SD implementations as in MMSE LD, both exact and inexact, to process each regularized lattice decoding problem in LDR LD.
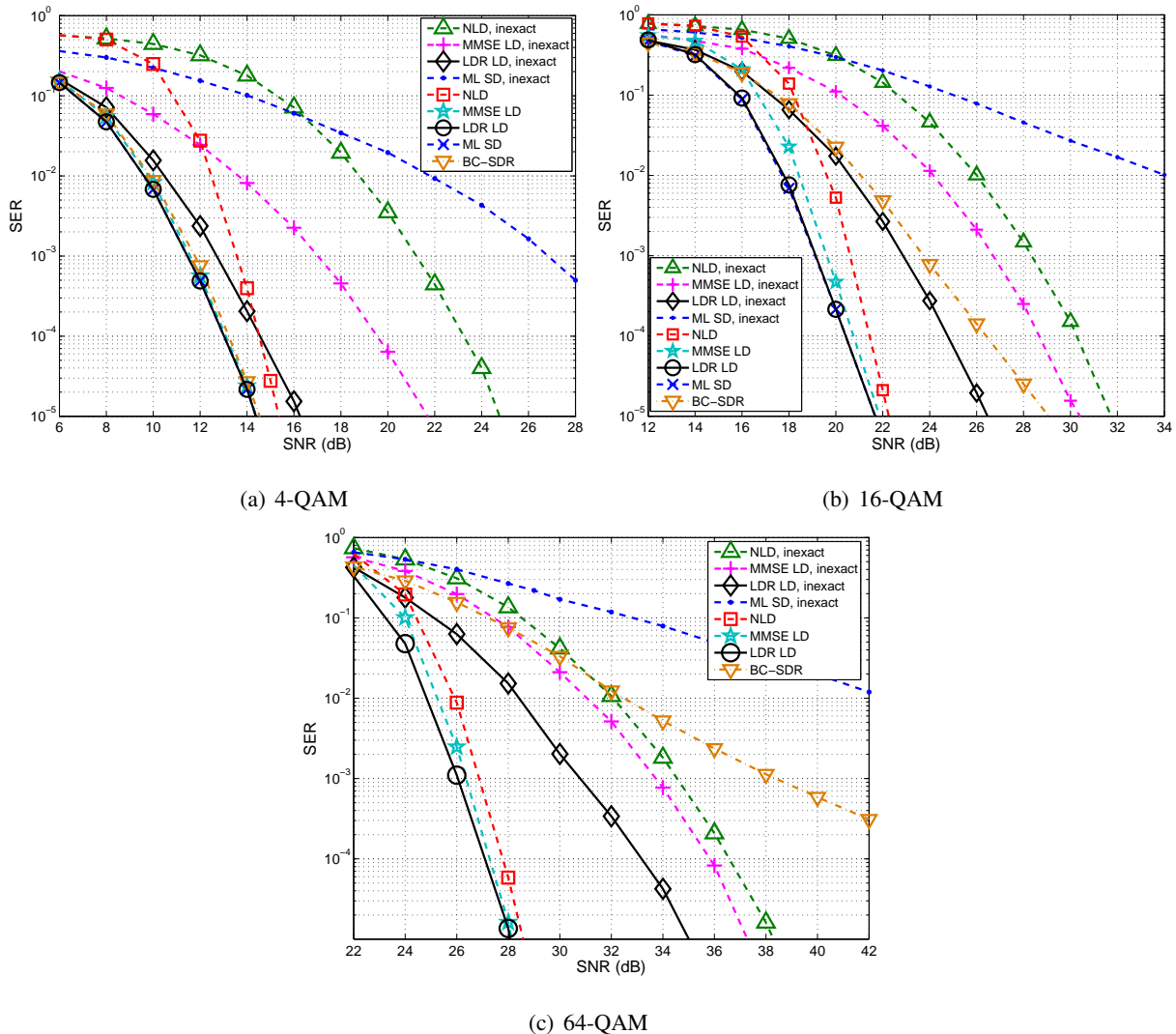
(a) 4-QAM

(b) 16-QAM

(c) 64-QAM

Fig. 1. Symbol error rate comparison for the SD-based detectors. $(M_C, N_C) = (16, 16)$.

A. Symbol Error Rate Performance of the LDR LD Detectors

In this subsection, we illustrate the SER performance of the LDR LD detector. The problem size is chosen as $(M_C, N_C) = (16, 16)$. Fig. 1 plots the SERs of the various detectors versus SNR under 4-QAM, 16-QAM and 64-QAM constellations. We have several key observations. First, the LDR LD detector achieves the same SER performance as the ML SD detector in the 4-QAM and 16-QAM cases. We are unable to verify whether the same desirable result holds for 64-QAM, since the ML SD detector is too slow to run in the 64-QAM case. Note that the identical performance of the ML SD and LDR LD detectors for the 4-QAM case is expected, since Theorem 1 shows that LDR is theoretically ML-optimal

for 4-QAM. Second, the LDR LD detector gives SER performance no worse than the BC-SDR detector. In fact, the SER gaps between the LDR LD and BC-SDR detectors are significant for the 16-QAM and 64-QAM cases. This numerical observation is consistent with the duality gap theorem in Theorem 2, which suggests that LDR should provide approximation accuracies at least no worse than BC-SDR. Third, the LDR LD detector generally yields better SER performance than the NLD and MMSE LD detectors. Further comparing the LD detectors, we observe that NLD may suffer from more than 2dB performance loss relative to LDR LD, especially for 4-QAM and 16-QAM. The gap nevertheless reduces for 64-QAM. Another observation is that MMSE LD is quite close to LDR LD. This suggests that MMSE regularization is a good regularization under the exact SD implementation.

However, when we consider the runtime-limited inexact implementation, the SER gaps between the LDR method and the MMSE regularization-based method widen significantly. The inexact implementation results are also included in Fig. 1. We first notice that an inexact detector exhibits performance loss compared to its original exact counterpart, which is expected as a compromise for computational efficiency. The inexact LDR LD detector is seen to lose less, and perform best among all the inexact detectors. We observe that at SER= 10^{-5}, the SNR gains of the inexact LDR LD detector over the inexact MMSE LD detector (the second best) are 6dB, 4dB and 2dB for 4-QAM, 16-QAM and 64-QAM, respectively. The simulation results also show that the inexact ML SD detector is highly ineffective.

It is further observed that the inexact LDR LD detector has more performance advantages on small QAM sizes than large QAM sizes. This observation may be intuitively explained as follows: For small QAM sizes, constellation points on the symbol bounds constitute a larger portion of all constellation points, rendering out-of-bound symbol events more likely to occur. The LDR detector, which focuses on mitigating out-of-bound symbol effects, has more chances to kick in and improve the performance.

It is also interesting to examine the performance-complexity tradeoffs of the inexact LDR LD detector. Fig. 2 illustrates the SER performance of the inexact LDR LD detector and the inexact MMSE LD detector under various complexity limits. In the figure, the number c represents the power of the complexity limit. Specifically, for a given c, we set the worst-case complexity limit of the two inexact LD detectors to $N_{\text{node}} = \min\{N^c, N^{c-1} \times \text{SNR}\}$. We observe that the SER performance of the two detectors improves as c increases, and that the inexact LDR LD detector for $c = 4$ or above achieves a reasonably good performance. Also, for each given c, the inexact LDR LD detector is seen to outperform the inexact MMSE LD detector.
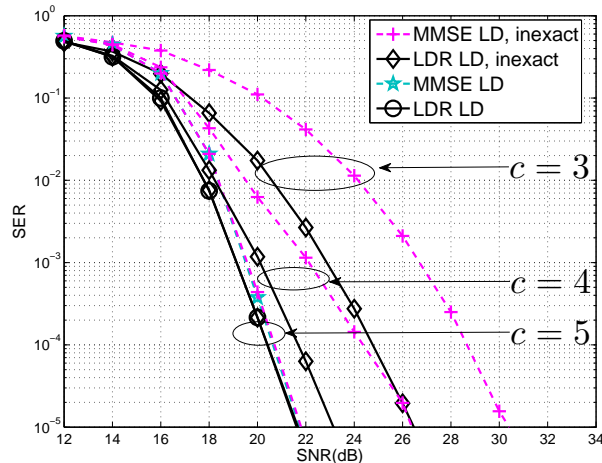
Fig. 2. Symbol error rate comparison of the inexact LD detectors under various complexity limits. $(M_C, N_C) = (16, 16)$, 16-QAM.

From all the observations above, we conclude that the LDR LD detector can yield near-ML performance, although the MMSE LD detector is also close. The inexact LDR LD detector has considerable SNR gains over inexact MMSE LD as well as the other inexact detectors; the gains are particularly significant for smaller QAM sizes.

B. Convergence of the LDR LD Detector

This subsection aims at studying the convergence behaviors of the PS iterations of the LDR LD detector. We start with considering the objective value convergence for a single problem realization. Fig. 3 shows the result, where the dual objective values $d_{\text{best}}^{(k)}$ of the LDR LD detector are plotted against the PS iteration number k (see Remark 3 for the precise definition of $d_{\text{best}}^{(k)}$). We tested the three initialization schemes, namely, random initialization, MMSE initialization and BR initialization. The problem size is $(M_C, N_C) = (16, 16)$, and the QAM size is 16. For reference, we also plot the optimal ML objective value, obtained via the ML SD detector. From Fig. 3, we observe that the three differently initialized LDR LD detectors converge to the same objective value. Also, for this problem realization, they converge to the optimal ML objective value. However, the convergence speed of the three initialization schemes has significant differences. It is seen that random initialization is the worse, taking more than 35 iterations to converge. MMSE initialization is better than random initialization, but the best is BR initialization: BR initialization takes only one iteration to converge at SNR=22dB, and three iterations at SNR=16dB.

We further illustrate the convergence speed of the LDR LD detectors by examining the average numbers

of PS iterations to terminate. The stopping parameters of the PS iterations is set to be $K_{\max} = 50$ and $\epsilon = 10^{-9}$. A number of $10,000$ independent problem realizations was run. Again, we consider 16-QAM and $M_C = N_C$. Table I shows the average numbers of iterations with respect to the problem size N_C; the SNR is fixed at SNR= 22dB. It is observed that the BR-initialized LDR LD detector takes about one iteration on average to complete the task, while the MMSE-initialized LDR LD detector takes $2 - 4$ iterations. Table II shows the average number of iterations with respect to the SNR, with $N_C = 16$. We can see that the average numbers of iterations of the LDR LD detectors tend to increase when the SNR decreases. Nevertheless, the BR-initialized LDR LD detector is very efficient in terms of the number of iterations used.
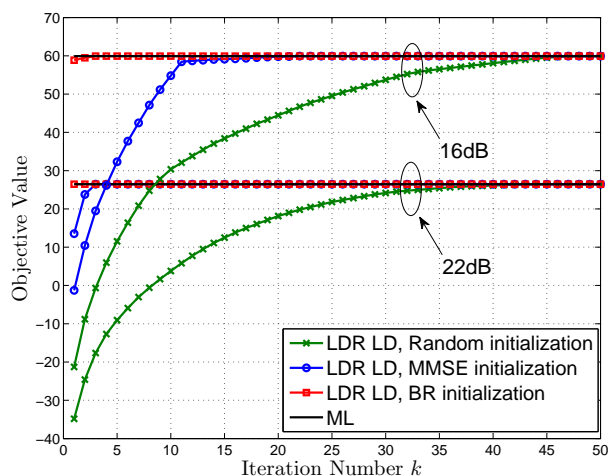
Fig. 3. Convergence of the LDR LD detector in one realization. 16-QAM, $(M_C, N_C) = (16, 16)$.

	Problem size N_C					
	6	10	14	18	22	26
BR init.	1.083	1.016	1.009	1.010	1.008	1.006
MMSE init.	2.040	2.000	3.000	3.000	4.000	4.000

TABLE I

AVERAGE NUMBER OF PS ITERATIONS OF THE LDR LD DETECTOR. 16-QAM, SNR=22dB, $M_C = N_C$, $K_{\max} = 50$, $\epsilon = 10^{-9}$.

	SNR (dB)				
	16	17.5	19	20.5	22
BR init.	5.746	2.026	1.098	1.025	1.009
MMSE init.	23.76	9.402	5.115	4.001	3.000

TABLE II

AVERAGE NUMBER OF PS ITERATIONS OF THE LDR LD DETECTOR. 16-QAM, $(M_C, N_C) = (16, 16)$, $K_{\max} = 50$, $\epsilon = 10^{-9}$.

C. Complexity Performance of the LDR LD Detectors

In this subsection, we examine the complexities of the proposed LDR detectors. Fig. 4 presents the average number of floating point operations (FLOPs) of the various detectors with respect to the problem sizes, where we set SNR=22dB, $M_C = N_C$ and the QAM size to be 16. Note that the LDR LD detector is initialized by BR, and the overhead of computing the initialization, i.e., BR optimization, has been included in evaluating the FLOPs of the LDR detectors. For problem sizes $N_C \leq 6$, the complexity of the ML SD detector is very low, and is even faster than the suboptimal detectors. But its complexity becomes unacceptably large as the problem size increases, rendering the ML SD detector impractical for large problem sizes. All the other detectors have much lower complexities than the ML SD detector for problem sizes $N_C \geq 10$. We can see that the complexities of LDR LD and MMSE LD increase at a rate much lower than that of the ML SD detector, though they still exhibit exponential complexity behaviors eventually. Moreover, as expected, the complexities of the inexact detectors are much lower than those of their exact counterparts when the problem size is large.

The complexity comparison in Fig. 4 also reveals that for problem sizes $N_C \geq 18$, the LDR LD detector can be more efficient than the MMSE LD. This result seems counter-intuitive, since the LDR LD detector is an iterative LD method, rather than one-shot LD as in MMSE LD. The reason for this actually lies in the chosen initialization scheme for LDR LD, i.e., BR initialization. As illustrated previously, the BR-initialized LDR LD detector takes very few number of iterations to converge. For the problem setting here, the average number of iterations is about one (cf. Table I). Moreover, by empirical experience, we found that regularization via BR is helpful in improving the SD computational speed. To explain, in Table III we give a breakdown of the complexities of MMSE LD and LDR LD. In the table, "SD" represents the FLOPs consumed by the SD algorithm, and "others" the FLOPs of other operations, which include LLL reduction, BR optimization (for LDR LD only), and other matrix operations. We can see that LDR LD is always more expensive than MMSE LD on "others"; this makes sense because LDR

LD requires solving the BR problem for initialization. However, LDR LD is much cheaper than MMSE LD on "SD" when $N_C \geq 14$. In fact, the SD complexities of both LDR LD and MMSE LD dominates the total complexities for $N_C \geq 14$. As a result, the LDR LD detector can be faster than MMSE LD for moderate to large problem sizes. As a future direction, it will be interesting to further investigate why BR regularization can accelerate SD so significantly.
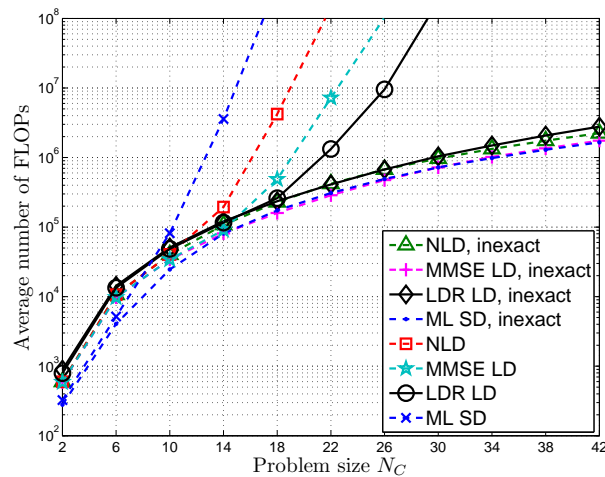

Fig. 4. Average number of FLOPs of various detectors. 16-QAM, SNR=22dB, $M_C = N_C$.

		Problem Size N_C			
		6	14	22	26
SD	MMSE LD	**4.0e2**	1.9e4	7.0e6	1.0e8
	LDR LD	4.1e2	**5.6e3**	**9.7e5**	**9.0e6**
Others	MMSE LD	**9.4e3**	**7.2e4**	**2.1e5**	**3.1e5**
	LDR LD	1.1e4	8.0e4	3.5e5	5.4e5
Total	MMSE LD	**9.8e3**	9.1e4	7.2e6	1.0e8
	LDR LD	1.1e4	**8.5e4**	**1.3e6**	**9.5e6**

TABLE III

AVERAGE NUMBER OF FLOPS OF THE OPERATIONS OF THE LDR LD AND MMSE LD DETECTORS. 16-QAM, SNR=22 dB, $M_C = N_C$.

D. Symbol Error Rate Performance of the LDR LRA-DF Detector

In this subsection, we test the SER performance of the LDR LRA-DF detector. The 16-QAM case is considered. Fig. 5 shows the SERs of the LDR LRA-DF detector and the MMSE LRA-DF detector under

various problem sizes. It can be observed that the LDR LRA-DF detector exhibits dramatic performance gains compared to MMSE LRA-DF as the problem size increases. While LDR LRA-DF needs about 27dB to achieve SER=10^{-5} for the three problem sizes tested, MMSE LRA-DF requires a much higher SNR to achieve the same SER level—and this is particularly true for larger problem sizes. This demonstrates that the LDR method is also very effective in boosting the performance of the LRA-DF receiver approach.
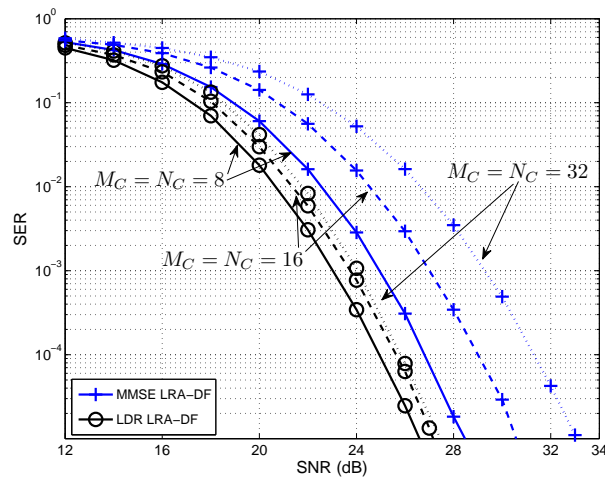

Fig. 5. Symbol error rates of the LDR and MMSE LRA-DF detectors. 16-QAM.

VI. Conclusion and Discussion

To conclude, this paper addressed a regularization optimization problem in lattice decoding, by considering the LDR of the ML detection problem. It was found that the LDR problem can be regarded as that of finding the best diagonally regularized lattice decoding to approximate the ML detector, and that the well-known NLD and MMSE LD detectors can be seen as particular instances of LDR. We proved that in the 2-PAM constellation case, lattice decoding with a proper regularization is optimal. We also established a connection between LDR and a previously developed semidefinite relaxation-based method, showing that the former yields relaxation tightness at least no worse than the latter. The projected subgradient method was derived to solve the LDR problem, thereby obtaining the best regularization. Based on the idea of projected subgradient, we developed the LDR LD detector, and its approximations using conventional suboptimal lattice decoding methods. Simulation results showed that the LDR LD approach yields promising symbol error probability and complexity performance.

The present research endeavor also demonstrated that the BR initialization scheme developed for the LDR LD approach plays a non-negligible role. While the LDR LD detector can converge to the optimum

for any given feasible initialization, as suggested by convergence results in optimization theory, our extensive simulation results have shown that BR initialization can improve both the convergence speeds and sphere decoding complexities quite significantly. Thus, an interesting future direction would be to further understand the merits of BR initialization via analysis.

This work focuses on hard decision-based MIMO detection. Since soft decision-based MIMO detection is important to practical systems (see the literature, such as [42] and the references therein), it would be meaningful to further study how the proposed LDR solutions may be used to provide soft decisions. Such a future direction seems feasible, since some existing soft decision generation tricks for LRA methods, such as bit flipping [43] and K-best SD [44], are also applicable to the LDR approach. It would also be interesting to investigate whether the structures of LDR may be exploited to provide efficient methods for soft decision generations.

APPENDIX A

PROOF OF FACT 1

Statement 1 is straightforward: Since the NLD problem (3) is a relaxation of the ML problem (7), NLD is automatically ML-optimal under instances where the NLD solution $\hat{s}_{\text{NLD}}$ is a feasible point of the ML problem. For Statement 2, we first note that $\|y - H\hat{s}_{\text{NLD}}\|^2 = \varphi(0) = d(0) \leq d^\star \leq f^\star$. Since we have $f^\star = \|y - H\hat{s}_{\text{NLD}}\|^2$ by Statement 1, we obtain $f^\star = d^\star$. This subsequently implies that $\lambda = 0$ is an optimal solution of the LDR problem. For Statement 3, we prove the statement using contradiction. Suppose that $\hat{s}_{\text{NLD}}$ is not an optimal solution of problem (Φ_λ) for $\lambda = \lambda^\star$. This is equivalent to saying

$$\varphi(\lambda^\star) < \|y - H\hat{s}_{\text{NLD}}\|^2 + \hat{s}_{\text{NLD}}^T D(\lambda^\star)\hat{s}_{\text{NLD}}. \tag{39}$$

Applying (39) to the strong duality result in Statement 2, we have

$$0 = d(\lambda^\star) - f^\star \tag{40a}$$

$$< \|y - H\hat{s}_{\text{NLD}}\|^2 + \hat{s}_{\text{NLD}}^T D(\lambda^\star)\hat{s}_{\text{NLD}} - u^2 \mathbf{1}^T \lambda^\star - f^\star \tag{40b}$$

$$= \left(\hat{s}_{\text{NLD}}^2 - u^2 \mathbf{1}\right)^T \lambda^\star \tag{40c}$$

where (40c) is owing to $f^\star = \|y - H\hat{s}_{\text{NLD}}\|^2$. However, since $\lambda^\star \succeq 0$ and $[\hat{s}_{\text{NLD}}]_i^2 \leq u^2$ for all $i = 1, \ldots, N$, we always obtain $\left(\hat{s}_{\text{NLD}}^2 - u^2 \mathbf{1}\right)^T \lambda^\star \leq 0$, which is a contradiction to (40c).

APPENDIX B

PROOF OF LEMMA 1

The proof is by contradiction. Assume that (14) holds, and yet $[\hat{s}_{\boldsymbol{\lambda}}]_i^2 > m^2$. Then we have

$$\varphi(\boldsymbol{\lambda}) = \|\boldsymbol{y} - \boldsymbol{H}\hat{s}_{\boldsymbol{\lambda}}\|^2 + \hat{s}_{\boldsymbol{\lambda}}^T \boldsymbol{D}(\boldsymbol{\lambda})\hat{s}_{\boldsymbol{\lambda}}$$

$$\geq \|\boldsymbol{y} - \boldsymbol{H}\hat{s}_{\boldsymbol{\lambda}}\|^2 + \lambda_i \left((m+2)^2 - 1\right) + \boldsymbol{\lambda}^T \mathbf{1}, \tag{41}$$

where we have used the fact that $[\hat{s}_{\boldsymbol{\lambda}}]_j^2 \geq 1$ for $j \neq i$ and $[\hat{s}_{\boldsymbol{\lambda}}]_i^2 \geq (m+2)^2$ (recall $\hat{s}_{\boldsymbol{\lambda}} \in 2\mathbb{Z}^N + \mathbf{1}$). By applying (14) to the second term of (41), we get

$$\varphi(\boldsymbol{\lambda}) > \|\boldsymbol{y} - \boldsymbol{H}\hat{s}_{\boldsymbol{\lambda}}\|^2 + c(\boldsymbol{y}, \boldsymbol{H}) + \boldsymbol{\lambda}^T \mathbf{1}$$

$$\geq \min_{\boldsymbol{s} \in \{\pm 1\}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T \mathbf{1}, \tag{42}$$

where (42) is due to $\|\boldsymbol{y} - \boldsymbol{H}\hat{s}_{\boldsymbol{\lambda}}\|^2 \geq \min_{\boldsymbol{s} \in 2\mathbb{Z}+\mathbf{1}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2$. From (42), we can equivalently write

$$\varphi(\boldsymbol{\lambda}) > \min_{\boldsymbol{s} \in \{\pm 1\}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T \boldsymbol{D}(\boldsymbol{\lambda})\boldsymbol{s} \geq \varphi(\boldsymbol{\lambda}), \tag{43}$$

where the second inequality is by the definition of $\varphi(\boldsymbol{\lambda})$, cf. (11). Eq. (43) is clearly a contradiction. We therefore conclude that under the condition in (14), $[\hat{s}_{\boldsymbol{\lambda}}]_i^2 \leq m^2$ must hold.

APPENDIX C

PROOF OF THEOREM 1

Our proof is based on a strong duality result for general integer programming problems [45]. The result, for the ML problem, is stated as follows.

Lemma 3. *[Strong Lagrangian duality [45]] Let* $\boldsymbol{\lambda} \succeq \mathbf{0}$, *and* $\hat{s}_{\boldsymbol{\lambda}} \in 2\mathbb{Z}^N + \mathbf{1}$ *be an optimal solution of problem* $(\Phi_{\boldsymbol{\lambda}})$. *If the following conditions are satisfied:*

$$\hat{s}_{\boldsymbol{\lambda}}^2 \preceq u^2 \mathbf{1},$$

$$\boldsymbol{\lambda}^T \left(\hat{s}_{\boldsymbol{\lambda}}^2 - u^2 \mathbf{1}\right) = 0, \tag{44}$$

then $\boldsymbol{\lambda}$ *and* $\hat{s}_{\boldsymbol{\lambda}}$ *are optimal solutions of the LDR problem* (9) *and ML problem* (7), *respectively. Moreover, strong duality* $f^{\star} - d^{\star}$ *holds.*

Suppose that $\boldsymbol{\lambda} \succ \gamma_1 \mathbf{1}$. By Lemma 1, we know that $[\hat{s}_{\boldsymbol{\lambda}}]_i^2 \leq 1$ for all i. Since $\hat{s}_{\boldsymbol{\lambda}} \in 2\mathbb{Z}^N + \mathbf{1}$, we must have $[\hat{s}_{\boldsymbol{\lambda}}]_i^2 = 1$ for all i. Thus, the conditions in Lemma 3 are satisfied, and strong duality holds. It also follows from Lemma 3 that any $\boldsymbol{\lambda} \succ \gamma_1 \mathbf{1}$ is LDR-optimal, and the corresponding $\hat{s}_{\boldsymbol{\lambda}}$ ML-optimal.

APPENDIX D

PROOF OF THEOREM 2

The statement in Theorem 2 is the same as saying that $d^\star \geq g^\star$. To prove this result, we first use the implication

$$\boldsymbol{s} \in 2\mathbb{Z}^N + \mathbf{1} \Longrightarrow \boldsymbol{s}^2 \succeq \mathbf{1}$$

to obtain a lower bound on $d(\boldsymbol{\lambda})$ in (10)-(11)

$$d(\boldsymbol{\lambda}) \geq \min_{\boldsymbol{s} \in \mathbb{R}^N, \ \boldsymbol{s}^2 \succeq \mathbf{1}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T(\boldsymbol{s}^2 - u^2\mathbf{1}). \tag{45}$$

By weak duality, the right-hand side (RHS) of (45) is lower-bounded by its Lagrangian dual, which is given by

$$d(\boldsymbol{\lambda}) \geq \max_{\boldsymbol{\mu} \succeq \mathbf{0}} \min_{\boldsymbol{s} \in \mathbb{R}^N} \ \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T(\boldsymbol{s}^2 - u^2\mathbf{1}) - \boldsymbol{\mu}^T(\boldsymbol{s}^2 - \mathbf{1}), \tag{46}$$

where $\boldsymbol{\mu} \succeq \mathbf{0}$ is the dual variable for the constraint $\boldsymbol{s}^2 \succeq \mathbf{1}$. By recalling that $d^\star = \max_{\boldsymbol{\lambda} \succeq \mathbf{0}} d(\boldsymbol{\lambda})$, we arrive at

$$d^\star \geq \max_{\substack{\boldsymbol{\lambda} \succeq \mathbf{0}, \\ \boldsymbol{\mu} \succeq \mathbf{0}}} \min_{\boldsymbol{s} \in \mathbb{R}^N} \ \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T(\boldsymbol{s}^2 - u^2\mathbf{1}) - \boldsymbol{\mu}^T(\boldsymbol{s}^2 - \mathbf{1}). \tag{47}$$

Our next step is to show that the RHS of (47) is equivalent to $g^\star$ in (17). The RHS of (47) can be rewritten as

$$\begin{aligned}
\max_{\boldsymbol{\lambda}, \boldsymbol{\mu}, r} \ & r \\
\text{s.t. } & \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{\lambda}^T(\boldsymbol{s}^2 - u^2\mathbf{1}) - \boldsymbol{\mu}^T(\boldsymbol{s}^2 - \mathbf{1}) \geq r, \ \forall \boldsymbol{s} \in \mathbb{R}^N \\
& \boldsymbol{\lambda} \succeq \mathbf{0}, \ \boldsymbol{\mu} \succeq \mathbf{0}.
\end{aligned} \tag{48}$$

By the lemma given in [46, p.163], we can equivalently turn problem (48) to

$$\begin{aligned}
\max_{\boldsymbol{\lambda}, \boldsymbol{\mu}, r} \ & r \\
\text{s.t. } & \begin{bmatrix} \boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda} - \boldsymbol{\mu}) & -\boldsymbol{H}^T\boldsymbol{y} \\ -\boldsymbol{y}^T\boldsymbol{H} & \boldsymbol{\mu}^T\mathbf{1} - u^2\boldsymbol{\lambda}^T\mathbf{1} - r + \|\boldsymbol{y}\|^2 \end{bmatrix} \succeq \mathbf{0}, \\
& \boldsymbol{\lambda} \succeq \mathbf{0}, \ \boldsymbol{\mu} \succeq \mathbf{0}.
\end{aligned} \tag{49}$$

Let us take a look at the dual of problem (49). The Lagrangian function of problem (49) is written as

$$\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}, r, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{X}) = r + \boldsymbol{p}^T\boldsymbol{\mu} + \boldsymbol{q}^T\boldsymbol{\lambda} + \text{tr}\left(\boldsymbol{X} \begin{bmatrix} \boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{D}(\boldsymbol{\lambda} - \boldsymbol{\mu}) & -\boldsymbol{H}^T\boldsymbol{y} \\ -\boldsymbol{y}^T\boldsymbol{H} & \boldsymbol{\mu}^T\mathbf{1} - u^2\boldsymbol{\lambda}^T\mathbf{1} - r + \|\boldsymbol{y}\|^2 \end{bmatrix}\right),$$

where $p \succeq 0, q \succeq 0$ and $X \succeq 0$ are the dual variables for $\mu \succeq 0$, $\lambda \succeq 0$ and the first constraint in (49), respectively. By partitioning X as

$$X = \begin{bmatrix} S & s \\ s^T & t \end{bmatrix}$$

where $S \in \mathbb{S}^N$, $s \in \mathbb{R}^N$, $t \in \mathbb{R}$, the function $\mathcal{L}$ can be reorganized as

$$\mathcal{L}(\lambda, \mu, r, p, q, X) = (1-t)r + (q + d(S) - tu^2 \mathbf{1})^T \lambda + (p - d(S) + t\mathbf{1})^T \mu$$
$$+ \mathrm{tr}(H^T H S) - 2s^T H^T y + t\|y\|^2.$$

We note that the dual function respective to the above $\mathcal{L}$ is bounded only if the first three terms above are zero; i.e., $1 - t = 0$, $q + d(S) - tu^2 \mathbf{1} = 0$ and $p - d(S) + t\mathbf{1} = 0$. As a result, the dual of problem (49) is

$$\min_{p,q,S,s} \quad \mathrm{tr}(H^T H S) - 2s^T H^T y + \|y\|^2 \tag{50a}$$

$$\text{s.t.} \quad q + d(S) - u^2 \mathbf{1} = 0, p - d(S) + \mathbf{1} = 0, \tag{50b}$$

$$p \succeq 0, \ q \succeq 0, \tag{50c}$$

$$\begin{bmatrix} S & s \\ s^T & 1 \end{bmatrix} \succeq 0. \tag{50d}$$

Since problem (50) is strictly feasible and bounded from below, by the conic duality theorem [46, Theorem 2.4.1], problems (49) and (50) attain the same optimal objectives; i.e., strong duality holds. Furthermore, by substituting (50b) into (50c) and applying Schur's complement [31] to (50d), we show that problem (50) is same as the BC-SDR problem in (17). The desired result $d^\star \geq g^\star$ is therefore concluded.

APPENDIX E

PROOF OF LEMMA 2

First, we note that

$$\|y - H\hat{s}_\lambda\| \geq \|H\hat{s}_\lambda\| - \|y\| \geq \sigma_{\min}\|\hat{s}_\lambda\| - \|y\|, \tag{51}$$

where $\sigma_{\min}$ is the smallest singular value of H, which is strictly positive for a full column rank H. From (51), we obtain

$$\|\hat{s}_\lambda\| \leq \frac{1}{\sigma_{\min}} \left(\|y - H\hat{s}_\lambda\| + \|y\| \right). \tag{52}$$

Second, recall from (Φ_λ) that $\hat{s}_\lambda$ is an optimal solution of

$$\min_{s \in 2\mathbb{Z}^N + 1} \|y - Hs\|^2 + s^T D(\lambda) s. \tag{53}$$

Let

$$\mathcal{I} = \{\, i \in \{1, \ldots, N\} \mid \lambda_i > \gamma_1 \,\},$$

$$\mathcal{J} = \{\, i \in \{1, \ldots, N\} \mid \lambda_i \leq \gamma_1 \,\},$$

where γ_1 has been defined in (15). By Lemma 1, $\hat{s}_{\boldsymbol{\lambda}}$ must satisfy $[\hat{s}_{\boldsymbol{\lambda}}]_i^2 \leq 1$ for all $i \in \mathcal{I}$. Hence, problem (53) can be equivalently expressed as

$$\min_{\boldsymbol{s}} \ \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}_{\mathcal{J}}^T \boldsymbol{D}(\boldsymbol{\lambda}_{\mathcal{J}})\boldsymbol{s}_{\mathcal{J}} \tag{54}$$
$$\text{s.t. } \boldsymbol{s}_{\mathcal{I}}^2 = \boldsymbol{1}, \ \boldsymbol{s}_{\mathcal{J}} \in 2\mathbb{Z}^{|\mathcal{J}|} + \boldsymbol{1},$$

where $\boldsymbol{s}_{\mathcal{I}}$ denotes a subvector of $\boldsymbol{s}$ whose elements are $\{s_i\}_{i \in \mathcal{I}}$, and $\boldsymbol{s}_{\mathcal{J}}$ and $\boldsymbol{\lambda}_{\mathcal{J}}$ are defined in the same way. By noting that $\hat{s}_{\boldsymbol{\lambda}}$ is optimal to (54), we have

$$\|\boldsymbol{y} - \boldsymbol{H}\hat{s}_{\boldsymbol{\lambda}}\|^2 \leq \|\boldsymbol{y} - \boldsymbol{H}\hat{s}_{\boldsymbol{\lambda}}\|^2 + \hat{s}_{\boldsymbol{\lambda},\mathcal{J}}^T \boldsymbol{D}(\boldsymbol{\lambda}_{\mathcal{J}})\hat{s}_{\boldsymbol{\lambda},\mathcal{J}}$$
$$= \min_{\boldsymbol{s}_{\mathcal{I}}^2 = \boldsymbol{1}, \ \boldsymbol{s}_{\mathcal{J}} \in 2\mathbb{Z}^{|\mathcal{J}|}+\boldsymbol{1}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}_{\mathcal{J}}^T \boldsymbol{D}(\boldsymbol{\lambda}_{\mathcal{J}})\boldsymbol{s}_{\mathcal{J}}$$
$$\leq \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{1}\|^2 + \boldsymbol{1}^T \boldsymbol{D}(\boldsymbol{\lambda}_{\mathcal{J}})\boldsymbol{1}$$
$$\leq \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{1}\|^2 + N\gamma_1. \tag{55}$$

Moreover, it can be shown from (15)-(16) that

$$\gamma_1 \leq \frac{1}{8}\|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{1}\|^2. \tag{56}$$

Finally, by plugging (55)-(56) into (52), we obtain a finite upper bound on $\|\hat{s}_{\boldsymbol{\lambda}}\|$, as desired.

REFERENCES

[1] W. Su, Z. Safar, and K. Liu, "Full-rate full-diversity space-frequency codes with optimum coding advantage," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 229 –249, Jan. 2005.

[2] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Proces. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[3] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization*. Springer Verlag, 1993.

[4] M. 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. 2389–2402, Oct. 2003.

[5] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[6] H. El Gamal, G. Caire, and M. 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.

[7] A. D. Murugan, H. El 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.

[8] M. Taherzadeh and A. K. Khandani, "On the limitations of the naive lattice decoding," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4820–4826, Oct. 2010.

[9] J. Jaldén and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4765 –4780, Oct. 2010.

[10] A. Singh, P. Elia, and J. Jaldén, "Achieving a vanishing SNR gap to exact lattice decoding at a subexponential complexity," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3692–3707, Jun. 2012.

[11] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, pp. 515–534, 1982.

[12] D. Wübben, D. Seethaler, J. Jaldén, and G. Matz, "Lattice reduction," *IEEE Signal Proces. Mag.*, vol. 28, no. 3, pp. 70–91, May. 2011.

[13] D. Seethaler, J. Jaldén, C. Studer, and H. Bölcskei, "On the complexity distribution of sphere decoding," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5754 –5768, Sep. 2011.

[14] H. Yao and G. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE Global Conf. Commun.*, vol. 1, Taipei, Taiwan, Nov. 2002, pp. 424–428.

[15] C. Windpassinger and R. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction," in *Proc. Information Theory Workshop*, Munich, Germany, 31 March-4 April 2003, pp. 345–348.

[16] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," *IEEE Trans. Commun.*, vol. 56, no. 2, pp. 309–318, Feb. 2008.

[17] Y. H. Gan, C. Ling, and W. H. Mow, "Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection," *IEEE Trans. Signal Process*, vol. 57, no. 7, pp. 2701 –2710, Jul. 2009.

[18] L. Luzzi, G. Othman, and J. Belfiore, "Augmented lattice reduction for MIMO decoding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2853–2859, Sep. 2010.

[19] S. Liu, C. Ling, and D. Stehlé, "Decoding by sampling: A randomized lattice algorithm for bounded-distance decoding," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5933–5945, Sep. 2011.

[20] L. Luzzi, D. Stehlé, and C. Ling, "Decoding by embedding: Correct decoding radius and DMT optimality," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2960 – 2973, May 2013.

[21] M. Taherzadeh, A. Mobasher, and A. Khandani, "LLL reduction achieves the receive diversity in MIMO decoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4801–4805, Dec. 2007.

[22] N. Sidiropoulos and Z.-Q. Luo, "A semidefinite relaxation approach to MIMO detection for high-order QAM constellations," *IEEE Signal Process. Lett.*, vol. 13, no. 9, pp. 525–528, Sep. 2006.

[23] W.-K. Ma, T. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA," *IEEE Trans. Signal Process*, vol. 50, no. 4, pp. 912–922, Apr. 2002.

[24] S. Boyd and A. Mutapcic, "Subgradient methods," *Notes of EE364b, Standford University*, 2008.

[25] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.

[26] L. Babai, "On Lovasz' lattice reduction and the nearest lattice point problem," *Combinactorica*, vol. 6, no. 1, pp. 1–13, 1986.

[27] A. Burg, M. Borgmanr, M. Wenk, C. Studer, and H. Bölcskei, "Advanced receiver algorithms for MIMO wireless communications," in *Proc. Design Automation and Test Europe Conf. (DATE)*, vol. 1, Mar. 2006, pp. 593–598.

[28] A. Wiesel, Y. Eldar, and S. Shitz, "Semidefinite relaxation for detection of 16-QAM signaling in MIMO channels," *IEEE Signal Process. Lett.*, vol. 12, no. 9, pp. 653 – 656, Sep. 2005.

[29] Z. Mao, X. Wang, and X. Wang, "Semidefinite programming relaxation approach for multiuser detection of QAM signals," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4275 –4279, Dec. 2007.

[30] W.-K. Ma, C.-C. Su, J. Jaldén, T.-H. Chang, and C.-Y. Chi, "The equivalence of semidefinite relaxation MIMO detectors for higher-order QAM," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 1038–1052, Dec. 2009.

[31] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[32] M. Kisialiou and Z.-Q. Luo, "Probabilistic analysis of semidefinite relaxation for binary quadratic minimization," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1906–1922, Mar. 2010.

[33] A. M.-C. So, "Non-asymptotic performance analysis of the semidefinite relaxation detector in digital communications," *Preprint*, 2010. [Online]. Available: http://www1.se.cuhk.edu.hk/~manchoso/papers/mimo_sdp_mpsk.pdf

[34] J. Jaldén, C. Martin, and B. Ottersten, "Semidefinite programming for detection in linear systems-optimality conditions and space-time decoding," in *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Process. (ICASSP)*, vol. 4, 2003, pp. IV – 9–12.

[35] D. P. Bertsekas, *Convex Analysis and Optimization*. Nashua, NH, USA: Athena Scientific, 2003, with A. Nedić and A. E. Ozdaglar.

[36] H. Najafi, M. Jafari, and M. Damen, "On adaptive lattice reduction over correlated fading channels," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1224 –1227, May. 2011.

[37] D. Wübben, R. Böhnke, V. Kühn, and K.-D. Kammeyer, "Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction," in *Proc. IEEE Int. Conf. Commun.*, vol. 2, Paris, France, Jun. 2004, pp. 798–802.

[38] P. H. Tan, L. Rasmussen, and T. Lim, "Constrained maximum-likelihood detection in CDMA," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 142 –153, Jan. 2001.

[39] A. Yener, R. Yates, and S. Ulukus, "CDMA multiuser detection: a nonlinear programming approach," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1016 –1024, Jun. 2002.

[40] P. Stark and R. Parker, "Bounded-variable least-squares: an algorithm and applications," *Computational Statistics*, pp. 129–141, 1995.

[41] M. Emtiyaz, "Updating inverse of a matrix when a column is added/removed." [Online]. Available: http://www.cs.ubc.ca/~emtiyaz/Writings/OneColInv.pdf

[42] P. Fertl, J. Jaldén, and G. Matz, "Performance assessment of MIMO-BICM demodulators based on mutual information," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1366–1382, 2012.

[43] R. Wang and G. Giannakis, "Approaching MIMO channel capacity with soft detection based on hard sphere decoding," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 587–590, 2006.

[44] X.-F. Qi and K. Holt, "A lattice-reduction-aided soft demapper for high-rate coded MIMO-OFDM systems," *IEEE Signal Process. Lett.*, vol. 14, no. 5, pp. 305–308, 2007.

[45] D. Li and X. Sun, *Nonlinear Integer Programming*. New York, NY, USA: Springer, 2006.

[46] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization*. Philadelphia, PA, USA: MPS-SIAM Series on Optimization, 2001.