Barreal Fernandez, Amaro; Karrila, Alex; Karpuk, David A.; Hollanti, Camilla

Information bounds and flatness factor approximation for fading wiretap MIMO channels

Published in:
26th International Telecommunication Networks and Applications Conference, ITNAC 2016

DOI:
10.1109/ATNAC.2016.7878822

Published: 14/03/2017

Document Version
Peer reviewed version

Please cite the original version:
Barreal Fernandez, A., Karrila, A., Karpuk, D. A., & Hollanti, C. (2017). Information bounds and flatness factor approximation for fading wiretap MIMO channels. In *26th International Telecommunication Networks and Applications Conference, ITNAC 2016* (pp. 277-282). [7878822] IEEE.
https://doi.org/10.1109/ATNAC.2016.7878822

Information Bounds and Flatness Factor Approximation for Fading Wiretap MIMO Channels

Amaro Barreal*, Alex Karrila*, David Karpuk and Camilla Hollanti
Department of Mathematics and Systems Analysis, School of Science, Aalto University, Finland.
email: firstname.lastname@aalto.fi

***Abstract*—In this article, the design of secure lattice coset codes for general wireless channels with fading and Gaussian noise is studied. Recalling the eavesdropper's probability and information bounds, a variant of the latter is given from which it is explicitly seen that both quantities are upper bounded by (increasing functions of) the expected flatness factor of the faded lattice related to the eavesdropper.**

By making use of a recently developed approximation of the theta series of a lattice, it is further shown how the average flatness factor can be approximated numerically. In particular, based on the numerical computations, the average flatness factor not only bounds but also orders correctly the performance of different lattices.

I. Introduction

In the wireless wiretap scheme two legitimate communication parties, Alice and Bob, exchange information in the presence of an *eavesdropper*, Eve. In this setting, the communication parties rely on physical layer security rather than cryptographic protocols. Hence, Eve is assumed to have no computational limitations and know the cryptographic key, if any, but to have a worse signal quality than Bob.

The objective of code design in a wiretap channel is to maximize the data rate and Bob's correct decoding probability while minimizing Eve's information. It was shown in the seminal paper of Wyner [1] that the legitimate parties can design codes with asymptotically non-zero rate, zero error probability and zero information leakage. Today, this setup is particularly interesting in wireless channels that are open in nature but vulnerable to distortions

As a practical construction of a wiretap code, [2] introduced the general technique of *coset coding*, where random bits are added to the message to confuse the eavesdropper. In the specific case of a wireless channel, where lattice codes are suitable, the code lattice Λ_b is endowed with a sublattice $\Lambda_e \subset \Lambda_b$ which carries the random bits [3].

A. Related Work and Contributions

The security of lattice coset codes can be quantized either by Eve's correct decision probability, or alternatively by the mutual information of the message and Eve's received signal. For the *additive white Gaussian noise* (AWGN) channel, upper bounds are known for both approaches [3], [4] and, more importantly, both are increasing functions of the *flatness factor* of the lattice Λ_e. Sequences of lattice coset codes achieving security and reliability are also constructed in [4].

For fading channel models, probability and information bounds were derived in [5], [6], and [7], [8], respectively. Codes achieving security and reliability in the *multiple-input multiple-output* (MIMO) channel were given in [8]. In this paper, we recall the strategy of [5], [6] and give a variant of the information bounds. With this streamlined introductory computation, we obtain bounds which are increasing functions of the expected flatness factor of the faded lattice related to the eavesdropper. The agreement of the probability and information approaches is hence explicit, and a natural and explicit generalization of the AWGN case. The computations hold in any channel model with linear fading and Gaussian noise. We remark that, to the best of our knowledge, the steps towards practical code designs in low dimensions, such as [9], are based on the probability bounds.

Motivated by this, we show how to use an approximation of the theta series of a lattice, recently derived in [10], to efficiently compute the average flatness factor of a given lattice. Hence, we do not need to rely on further approximations for the information/probability bounds, *e.g.*, the common approach using the *inverse norm sum* in SISO channels [5]. We exemplify this in the Rayleigh fast fading channel, and see an agreement between our numerical computations and the geometric design heuristics and simulations results in [9]. In particular, this agreement supports the expectation that the flatness factor not only *bounds* but also *orders correctly* the performance of different lattices, hence serving as a design criterion for practical low-dimensional constructions, as first suggested in [5], [6]. Comparing to [9], where average flatness factor heuristics are tested with simulations, our results suggest that it is indeed approachable for design heuristics as well as numerical computations.

II. Lattices, Theta Series and the Flatness Factor

A *lattice* $\Lambda \subset \mathbb{R}^n$ is a discrete subgroup of $\mathbb{R}^n$ with the property that there exist $s \leq n$ linearly independent vectors $(\mathbf{b}_1, \ldots, \mathbf{b}_s)$ of $\mathbb{R}^n$ such that

$$\Lambda = \bigoplus_{i=1}^{s} \mathbf{b}_i \mathbb{Z}.$$

We say that $(\mathbf{b}_1, \ldots, \mathbf{b}_s)$ is a $\mathbb{Z}$-basis of Λ, and call $s \leq n$ the *rank*, and n the *dimension* of Λ. The lattice is *full* if $s = n$.

A lattice $\Lambda' \subset \mathbb{R}^n$ such that $\Lambda' \subset \Lambda$ is called a *sublattice* of Λ, and Λ is referred to as a *superlattice* for Λ'.

* Equal contribution.

For a convenient presentation, we define a *generator matrix* $M_\Lambda := [\mathbf{b}_1 \ \cdots \ \mathbf{b}_s] \in \mathrm{Mat}(n \times s, \mathbb{R})$, and equivalently write

$$\Lambda = \{ \lambda = M_\Lambda \mathbf{z} | \, \mathbf{z} \in \mathbb{Z}^s \}.$$

Definition 1. *Let $\Lambda \subset \mathbb{R}^n$ be a full lattice, and let $\lambda_{\min} := \min_{\lambda \in \Lambda \setminus \{\mathbf{0}\}} ||\lambda||^2$ be its* minimal norm. *The lattice Λ is called* well-rounded *if the set $\left\{ \lambda \in \Lambda \mid ||\lambda||^2 = \lambda_{\min} \right\}$ contains n linearly independent vectors.*

The *volume* of Λ is defined to be $\nu_\Lambda = |\det(M_\Lambda)|$, and is independent of the choice of basis. If Λ is not full, then $\nu_\Lambda = \det(M_\Lambda^t M_\Lambda)^{1/2}$. The dual lattice Λ^* of a full lattice Λ is the lattice generated by $M_{\Lambda^*} := (M_\Lambda^{-1})^t = (M_\Lambda^t)^{-1}$. We can easily compute the volume of a sublattice[1] $\Lambda' \subset \Lambda$ and of the dual lattice Λ^*, as

$$\nu_{\Lambda'} = \nu_\Lambda \, |\Lambda / \Lambda'| ; \quad \nu_{\Lambda^*} = 1/\nu_\Lambda,$$

where $|\Lambda/\Lambda'|$ is the group *index* of Λ' in Λ. Further, the *Voronoi cell* associated with a lattice point $\lambda \in \Lambda$ is the set

$$\mathcal{V}_\Lambda(\lambda) := \left\{ \mathbf{x} \in \mathbb{R}^n | \, ||\mathbf{x} - \lambda||^2 \leq ||\mathbf{x} - \lambda'||^2, \lambda' \in \Lambda \setminus \{\lambda\} \right\},$$

and $\mathcal{V}(\Lambda) := \mathcal{V}_\Lambda(\mathbf{0})$ denotes the *basic Voronoi cell* of Λ.

Definition 2. *Let Λ be a lattice. The* theta series *(or theta function) of Λ is the generating function*

$$\Theta_\Lambda(q) := \sum_{\lambda \in \Lambda} q^{||\lambda||^2}.$$

The theta series $\Theta_\Lambda(q)$ converges absolutely if $0 \leq q < 1$. In this article, we will need to compute the theta series of lattices which have been affected by random fading, *i.e.*, with *random* generator matrices. Unfortunately, there is no known way of computing the theta series of a random lattice in closed form. The following result is thus crucial for our purposes.

Proposition 1. *[10, Prop. 1] Let $\Lambda \subset \mathbb{R}^n$ be a full lattice with fundamental volume ν and minimal norm $\lambda_{\min}$. The theta series $\Theta_\Lambda\left(e^{-\pi\tau}\right)$, as a function of τ, can be expressed as*

$$\Theta_\Lambda\left(e^{-\pi\tau}\right) = 1 + \frac{(\pi\lambda_{\min})^{\frac{n}{2}+1}\tau}{\Gamma\left(\frac{n}{2}+1\right)\nu} \int_1^\infty t^{\frac{n}{2}} e^{-\pi\tau\lambda_{\min} t} dt + \Xi_n,$$

where $\Xi_n = \Xi_n(\tau, \Lambda, L)$ denotes the error term.

We refer to [10] for a more detailed version of this result. In Figure 1 we illustrate the accuracy of this approximation for various famous lattices for which the theta series is known in closed form and thus can be computed explicitly.

A. Lattice Sums and Flatness Factor

Let us denote the *probability density function* (PDF) of the n-dimensional spherical Gaussian as

$$g_n(\mathbf{x}; \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{||\mathbf{x}||^2}{2\sigma^2} \right),$$

[1]The index $|\Lambda/\Lambda'|$ is finite provided that $\dim(\Lambda) = \dim(\Lambda')$.

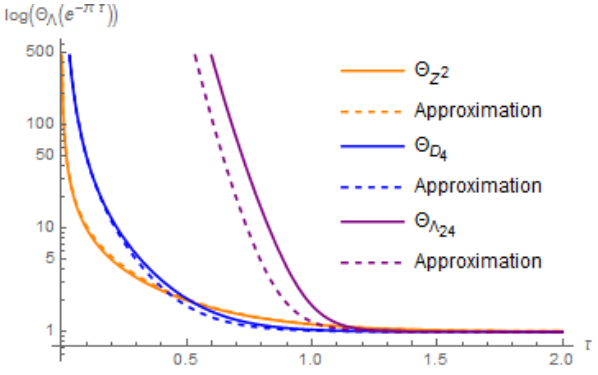

Fig. 1. Approximation of the theta series of the lattices $\mathbb{Z}^2$, D_4 and the Leech lattice Λ_{24}, in dimensions $n = 2, 4$ and 24.

and its sums over a (possibly shifted) lattice Λ as

$$g_n(\Lambda + \mathbf{x}; \sigma) := \sum_{\lambda \in \Lambda} g_n(\lambda + \mathbf{x}; \sigma).$$

We then have the identity

$$g_n(\Lambda; \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \Theta_\Lambda\left(e^{-\frac{1}{2\sigma^2}} \right),$$

from which we see that $g_n(\Lambda; \sigma)$ only differs by constants from the standard theta series. It is easy to see that $g_n(\Lambda + \mathbf{x}; \sigma)$ is Λ-periodic, and, for full lattices, it defines a PDF on the Voronoi cell $\mathcal{V}(\Lambda)$, called the *lattice Gaussian PDF*.

The following definition is crucial for all subsequent results.

Definition 3. *Let $\Lambda \subset \mathbb{R}^n$ be a lattice generated by $M_\Lambda \in \mathrm{Mat}(n, \mathbb{R})$, with fundamental volume ν_Λ. The* flatness factor *$\varepsilon_\Lambda(\sigma)$ of Λ is defined as*

$$\varepsilon_\Lambda(\sigma) = \varepsilon_{M_\Lambda}(\sigma) := \max_{\mathbf{x} \in \mathbb{R}^n} \left| \frac{g_n(\Lambda + M_\Lambda \mathbf{x}; \sigma)}{1/\nu_\Lambda} - 1 \right|.$$

The flatness factor was introduced in [4] as a wiretap information tool and measures the deviation of the lattice Gaussian PDF from the uniform distribution on $\mathcal{V}(\Lambda)$. As the maximum of $g_n(\Lambda + \mathbf{x}; \sigma)$ is attained for $\mathbf{x} \in \Lambda$ [4], the flatness factor of a full lattice can be expressed in terms of theta series,

$$\varepsilon_\Lambda(\sigma) = \nu_\Lambda g_n(\Lambda; \sigma) - 1 \tag{1}$$

$$= \Theta_{\Lambda^*}\left(e^{-2\pi\sigma^2} \right) - 1. \tag{2}$$

Note that the additive group structure of a lattice is actually crucial in this work: these important formulas are based on the Poisson summation formula.

III. System Model and Coset Codes

We consider a wireless fading channel with noise. Perfect channel state information is assumed at both receivers (CSIR), Bob and Eve; the transmitter is only assumed to know the channel statistics. As we are only interested in the eavesdropper's performance, we henceforth only consider the channel between Alice and Eve, and consequently forgo subscripts in the related quantities. Throughout this paper, random variables

are denoted by capital letters and their realizations with lowercase letters. Denote Alice's transmitted vector by $\mathbf{x} \in \mathbb{R}^n$, so that the channel equation is given by

$$\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{n},$$

where $\mathbf{h} \in \mathrm{Mat}(m \times n, \mathbb{R})$ is the realization of the fading, and the noise vector $\mathbf{N} \in \mathbb{R}^m$ is composed of i.i.d. components $N_i \sim \mathcal{N}(0, \sigma^2)$. We assume that $\mathbf{H}$ has full rank almost surely, but need not be a square matrix.

Remark 1. *Typically a complex fading channel model is considered, together with complex lattice codes. Such codes, however, can be be reduced to the real case with double the dimension. We consider here a real channel model since the class of real lattices is wider than that of complex lattices, most importantly including complex $\mathbb{Z}[e^{2\pi i/3}]$ -lattices, and since the average theta functions, to which both probability and information bounds reduce, are already computed explicitly [5], [6].*

To confuse Eve, Alice uses *lattice coset coding*. An original information vector thus corresponds not only to a lattice point, but to an entire coset, and the transmitted vector is then chosen randomly within the coset according to a distribution discussed below. More specifically, Alice is equipped with two nested lattices[2] $\Lambda_e \subset \Lambda_b \subset \mathbb{R}^n$, as well as an injective map from her message space $\mathcal{M}$ of cardinality $|\mathcal{M}| = |\Lambda_b/\Lambda_e|$, into the set of unique coset representatives of Λ_b/Λ_e,

$$\mathcal{E} : \mathcal{M} \to \Lambda_b \cap \mathcal{V}(\Lambda_e), \quad M \mapsto \lambda_M.$$

Alice then chooses a representative of the coset class corresponding to M at random, *i.e.*, picks $\lambda_e \in \Lambda_e$ and transmits $\mathbf{x} = \lambda_M + \lambda_e \in (\lambda_M + \Lambda_e) \in \Lambda_b/\Lambda_e$. The transmitted vector now represents the original information bits, and in addition, contains random bits which are encoded by Λ_e.

If $\mathbf{x} = M_\Lambda \mathbf{z}$ for $\mathbf{z} \in \mathbb{Z}^n$, then $\mathbf{h}\mathbf{x} = \mathbf{h}M_\Lambda \mathbf{z}$, and we can think of a lattice code under fading with CSIR as a Gaussian-channel lattice code where the code lattice realizes a random lattice with generator matrix $\mathbf{h}M_\Lambda$. We will henceforth denote the faded lattices Λ_b and Λ_e by $\Lambda_{b,\mathbf{h}}$ and $\Lambda_{e,\mathbf{h}}$, respectively.

We denote the mutual information of two random variables X, Y by $\mathrm{I}[X;Y]$, and for double conditions given a third random variable Z, we write $\mathrm{I}[X;(Y,Z)]$. The metric of interest is the information that Eve is able to extract from her observations, *i.e.*, $\mathrm{I}[M;(\mathbf{Y},\mathbf{H})]$. The fading $\mathbf{H}$, noise $\mathbf{N}$ and Alice's message M are assumed to be mutually independent. The settings considered in this article are:

1) Alice chooses coset class representatives uniformly at random, and the channel is a $\bmod \Lambda_s$ channel.
2) Alice uses Gaussian coset coding.

In the simplest setup, Alice chooses uniform representatives from a finite transmission region, and Eve's information is affected by the boundaries of that region. Naively speaking, if Eve has knowledge about the transmission region, she can often guess the transmitted vector correctly if the received vector lies outside of the transmission region. Our choice of setups can be roughly seen as removing such boundary effects by a modulo operation, and smoothing the boundary of the transmission region, respectively. The AWGN information bounds have been derived for these setups in [4], and the information-theoretic results for the fading channels in [7], [8] are for the Gaussian coset coding setup.

IV. Information and Probability Bounds

Let Λ_b and Λ_e be of dimension n, and assume that Eve simply decodes the received signal to the closest lattice point in $\Lambda_{e,\mathbf{h}}$. Then, probability of Eve correctly decoding the message is upper bounded by [5], [6]

$$\begin{aligned}
&\mathbb{P}\left[\text{Decode correctly in } \Lambda_e : \Lambda_b \text{ coset code}\right] \\
=&\mathbb{E}_{\mathbf{H}}\left[\mathbb{P}\left[\text{Decode correctly in } \Lambda_{e,\mathbf{h}} : \Lambda_{b,\mathbf{h}} \text{ AWGN coset code}\right]\right] \\
\leq&\mathbb{E}_{\mathbf{H}}\left[\nu_{\Lambda_{b,\mathbf{h}}} g_n(\Lambda_{e,\mathbf{h}};\sigma)\right] \\
=&\left[\Lambda_b : \Lambda_e\right]^{-1}\left(\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right] + 1\right). \qquad (3)
\end{aligned}$$

The first step is Fubini's theorem and the independence of $\mathbf{N}$, $\mathbf{H}$ and M, while the second step follows from the probability bound in the AWGN setup [3]. The bound is given for Rayleigh fading *single-input single-output* (SISO) and MIMO channels in [5] and [6], respectively. We recall that the flatness factor is defined through (1) for non-full lattices as well.

We will give a variant of the information bounds [7], [8] from which it is explicit that the bounds agree (up to constants) with this upper bound on the probability that Eve correctly decodes the message. Hence, to minimize this probability, it will be necessary to minimize the upper bound on the mutual information, and vice versa.

Lemma 1. *[4, Lemma 2] Let $\mathbf{Y}$ be an $\mathbb{R}^n$-valued random variable with PDF $\rho_{\mathbf{Y}}$, and consider a message space $\mathcal{M}$ such that $|\mathcal{M}| \geq 4$. Suppose that there exists some PDF $\tilde{\rho}$ in $\mathbb{R}^n$ such that for all $m \in \mathcal{M}$, the variational distance*

$$V(\rho_{\{\mathbf{Y}|M=m\}}, \tilde{\rho}) := \int_{\mathbf{y}\in\mathbb{R}^n} |\rho_{\{\mathbf{Y}|M=m\}}(\mathbf{y}) - \tilde{\rho}(\mathbf{y})| d^n\mathbf{y}$$

is upper bounded[3] by $\delta \leq e^{-1}/2$. Then,

$$I\left[M;\mathbf{Y}\right] \leq 2\delta \log(|\mathcal{M}|) - 2\delta \log(2\delta).$$

A. The $\bmod \Lambda_s$ Wiretap Channel

We first consider a strategy where the boundary of the shaping region set to be the boundary of $\mathcal{V}(\Lambda_s)$, where Λ_s is a third lattice $\Lambda_s \subset \Lambda_e \subset \Lambda_b \subset \mathbb{R}^n$, called the *shaping lattice*. The random shifting vector λ_e is chosen uniformly at random from the set of representatives of Λ_e/Λ_s in $\mathcal{V}(\Lambda_s)$. Then, slightly artificially, we assume that Eve only receives knowledge of the equivalence class $\mathbf{y}/\Lambda_{s,\mathbf{h}}$.

[2]The notation Λ_b, Λ_e is chosen to indicate that the message intended for Bob is taken from Λ_b, while Λ_e is the lattice that is chosen to confuse Eve.

[3]This assumption is implicit in [4] but necessary, seen by taking $\delta \to \infty$. This yields a small difference between our and earlier information bounds.

Theorem 1. *In the* $\bmod \Lambda_s$ *channel setup, let the message* M *have any distribution on the message space* $\mathcal{M}$ *of cardinality* $|\mathcal{M}| \geq 4$*, and assume that* $\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right] \leq e^{-1}/2$*. Then,*

$$\begin{aligned}\mathrm{I}\left[M;(\mathbf{Y}/\Lambda_{s,\mathbf{h}},\mathbf{H})\right] \leq 2(e+1)\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right]\log(|\mathcal{M}|) \\ - 2\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right]\log(2\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right]).\end{aligned}$$

Proof. By the independence assumptions, we have the identity

$$\mathrm{I}\left[M;(\mathbf{Y}/\Lambda_{s,\mathbf{h}},\mathbf{H})\right] = \mathbb{E}_{\mathbf{H}}\left[\mathrm{I}\left[M;(\mathbf{Y}/\Lambda_{s,\mathbf{h}}|\mathbf{H}=\mathbf{h})\right]\right].$$

We divide $\mathbf{Y}$ into components $\mathbf{Y}_\perp$ and $\mathbf{Y}_\parallel$, perpendicular and parallel to the nested lattices $\Lambda_{*,\mathbf{h}}$. By the independence assumptions, given $\mathbf{H} = \mathbf{h}$, $\mathbf{Y}_\perp$ is independent of the transmitted lattice point and hence of the message M. Thus, $\mathrm{I}\left[M;(\mathbf{Y}/\Lambda_{s,\mathbf{h}}|\mathbf{H}=\mathbf{h})\right] = \mathrm{I}\left[M;(\mathbf{Y}_\parallel/\Lambda_{s,\mathbf{h}}|\mathbf{H}=\mathbf{h})\right]$. For a fixed channel realization $\mathbf{h}$, the channel is just a Gaussian $\bmod \Lambda_{s,\mathbf{h}}$ channel. Then, given the message m, the projection modulo Eve's lattice, $\mathbf{Y}_\parallel/\Lambda_{e,\mathbf{h}}$, has the lattice $\Lambda_{e,\mathbf{h}}$ Gaussian distribution, of which the PDF of $\mathbf{Y}_\parallel/\Lambda_{s,\mathbf{h}}$ is simply a scaling. Denoting $\varepsilon = \varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)$, Lemma 1 and the trivial upper bound $\log(|\mathcal{M}|)$ on the mutual information yield

$$\begin{aligned}&\mathrm{I}\left[M;(\mathbf{Y}/\Lambda_{s,\mathbf{h}},\mathbf{H})\right] \\ &\leq \mathbb{E}_{\mathbf{H}}\left[\mathbb{1}_{\{\varepsilon\leq e^{-1}/2\}}(2\varepsilon\log(|\mathcal{M}|)-2\varepsilon\log(2\varepsilon))\right] \\ &\quad + \mathbb{E}_{\mathbf{H}}\left[\mathbb{1}_{\{\varepsilon>e^{-1}/2\}}\log(|\mathcal{M}|)\right] \\ &= \mathbb{P}_{\mathbf{H}}\left[\varepsilon\leq e^{-1}/2\right]\mathbb{E}_{\{\mathbf{H}|\varepsilon\leq e^{-1}/2\}}\left[2\varepsilon\log(|\mathcal{M}|)-2\varepsilon\log(2\varepsilon)\right] \\ &\quad + \mathbb{E}_{\mathbf{H}}\left[\mathbb{1}_{\{\varepsilon>e^{-1}/2\}}\right]\log(|\mathcal{M}|).\end{aligned}$$

We upper bound each of the summands separately. For the first one, we apply Jensen's inequality to the convex function $x\log x$ and bound $\mathbb{P}_{\mathbf{H}}\left[\varepsilon\leq e^{-1}/2\right]\leq 1$, which yields

$$\begin{aligned}&\mathbb{P}_{\mathbf{H}}\left[\varepsilon\leq e^{-1}/2\right]\mathbb{E}_{\{\mathbf{H}|\varepsilon\leq e^{-1}/2\}}\left[2\varepsilon\log(|\mathcal{M}|)-2\varepsilon\log(2\varepsilon)\right] \\ &\leq 2\mathbb{E}_{\{\mathbf{H}|\varepsilon\leq e^{-1}/2\}}\left[\varepsilon\right]\left(\log(|\mathcal{M}|)-\log(2\mathbb{E}_{\{\mathbf{H}|\varepsilon\leq e^{-1}/2\}}\left[\varepsilon\right])\right) \\ &\leq 2\mathbb{E}_{\mathbf{H}}\left[\varepsilon\right]\left(\log(|\mathcal{M}|)-\log(2\mathbb{E}_{\mathbf{H}}\left[\varepsilon\right])\right).\end{aligned}$$

The second inequality holds since $0 \leq \mathbb{E}_{\{\mathbf{H}|\varepsilon\leq e^{-1}/2\}}\left[\varepsilon\right] \leq \mathbb{E}_{\mathbf{H}}\left[\varepsilon\right] \leq e^{-1}/2$, and $x\log x$ is decreasing in this interval.

For the second term, we can use Markov's inequality to get

$$\mathbb{E}_{\mathbf{H}}\left[\mathbb{1}_{\{\varepsilon>e^{-1}/2\}}\right] \leq \frac{\mathbb{E}_{\mathbf{H}}\left[\varepsilon\right]}{e^{-1}/2} = 2e\mathbb{E}_{\mathbf{H}}\left[\varepsilon\right].$$

The result follows. □

B. Discrete Gaussian Coset Coding

While an insightful scenario, the restrictions imposed in the $\bmod \Lambda_s$ channel setting are not necessarily realistic. To be more general, we consider a second approach, where the boundary is smoothed instead of removed. Here, the vectors λ_e corresponding to the random bits are chosen so that the message $\mathbf{X} = \lambda_M + \lambda_e$ follows the Gaussian distribution centered on the shifted lattice $\Lambda_e + \lambda_M$, that is, for all $\mathbf{x} \in \Lambda_e + \lambda_M$,

$$P\left[\mathbf{X}=\mathbf{x}\right] = \frac{g_n(\mathbf{x};\sigma_s)}{g_n(\Lambda_e+\lambda_M;\sigma_s)} =: D_{\Lambda_e,\lambda_M}(\mathbf{x};\sigma_s).$$

The parameter σ_s^2 is called the *shaping variance*.

The following result, which can be regarded as a special case of [8, Lemma 1] in our notation, is needed for the generalization of the information bound to this second setting.

Lemma 2. *Fix* $\mathbf{h} \in \mathrm{Mat}(m\times n,\mathbb{R})$ *and let* $\mathbf{X}$ *have the centered discrete Gaussian distribution* $D_{\Lambda_e,\lambda_M}(\mathbf{x};\sigma_s)$*, where* Λ_e *is full. Let* $\mathbf{N}\sim\mathcal{N}(\mathbf{0},\sigma^2\mathbf{I}_m)$ *be a spherical (continuous) Gaussian vector independent of* $\mathbf{X}$*. Assume furthermore that* $\varepsilon_{\sqrt{\sigma^2/\sigma_s^2 I_n+\mathbf{h}^t\mathbf{h}}\Lambda_e}(\sigma)\leq 1/2$*, where* $\sqrt{\cdot}$ *denotes matrix square root. Then, the PDF* $\rho(\mathbf{y})$ *of* $\mathbf{Y}=\mathbf{h}\mathbf{X}+\mathbf{N}$ *and the PDF* $\tilde{\rho}(\mathbf{y})$ *of* $\mathcal{N}(\mathbf{0},(\sigma^2 I_m+\sigma_s^2\mathbf{h}\mathbf{h}^t))$*,*

$$\tilde{\rho}(\mathbf{y}) = \frac{\exp\left(-\frac{1}{2}\mathbf{y}^t(\sigma^2 I_m+\sigma_s^2\mathbf{h}\mathbf{h}^t)^{-1}\mathbf{y}\right)}{(\sqrt{2\pi})^m\sqrt{\det(\sigma^2 I_m+\sigma_s^2\mathbf{h}\mathbf{h}^t)}},$$

have variational distance at most $4\varepsilon_{\sqrt{\sigma^2/\sigma_s^2 I_n+\mathbf{h}^t\mathbf{h}}\Lambda_e}(\sigma)$*.*

Theorem 2. *Consider the Rayleigh fading channel with discrete Gaussian coset coding. Let the message* M *have any distribution on the message space* $\mathcal{M}$ *of cardinality* $|\mathcal{M}|\geq 4$*. Assume that* $E := \mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\sqrt{\sigma^2/\sigma_s^2 I_n+\mathbf{h}^t\mathbf{h}}\Lambda_e}(\sigma)\right] \leq \frac{1}{8e}$*. Then,*

$$\mathrm{I}\left[M;(\mathbf{Y},\mathbf{H})\right] \leq 8(1+e)E\log(|\mathcal{M}|) - 8E\log\left(8E\right).$$

Proof. The proof closely follows the steps of that of Theorem 1. We start by writing

$$\mathrm{I}\left[M;(\mathbf{Y},\mathbf{H})\right] = \mathbb{E}_{\mathbf{H}}\left[\mathrm{I}\left[M;(\mathbf{Y}|\mathbf{H}=\mathbf{h})\right]\right].$$

For a fixed channel realization $\mathbf{h}$, by Lemma 2 the distribution of the received vector $\mathbf{Y}$ is close to a fixed Gaussian distribution $\tilde{\rho}$ for all messages M, with variational distance

$$V\left(\rho_{\{\mathbf{Y}|M=m\}},\tilde{\rho}\right)\leq 4\varepsilon(\sigma),$$

for $\varepsilon\leq 1/2$. For the values $\mathbf{h}$ such that $\varepsilon\leq\frac{1}{8e}$ we get an information bound using Lemma 1. Otherwise, we have the trivial upper bound $\log(|\mathcal{M}|)$, yielding

$$\begin{aligned}\mathrm{I}\left[M;(\mathbf{Y},\mathbf{H})\right] \leq \mathbb{E}_{\mathbf{H}}\left[\mathbb{1}_{\left\{\varepsilon\leq\frac{1}{8e}\right\}}\left[8\varepsilon\left(\log(|\mathcal{M}|)-\log\left(8\varepsilon\right)\right)\right]\right] \\ + \mathbb{E}_{\mathbf{H}}\left[\mathbb{1}_{\left\{\varepsilon>\frac{1}{8e}\right\}}\log(|\mathcal{M}|)\right]. \qquad (4)\end{aligned}$$

The rest of the proof is identical to Theorem 1, using Jensen's inequality for the first, and Markov's inequality for the second summand to obtain termwise bounds which we can substitute back to (4) to conclude the proof. □

C. Observations

The information bounds tend to zero with the respective average flatness factor. By the dual formula (2) for the flatness factor, $\varepsilon_\Lambda(\sigma)$ decreases monotonously to zero as $\sigma\to\infty$ for any Λ. As it only depends on ratios of the different parameters σ,σ_s,σ_h, it is easy to deduce that the respective average flatness factors also decrease monotonously to zero at poor signal quality. Hence, the bounds prove information-theoretic security for poor eavesdropper's channel quality.

The information bound of Theorem 2 involves the quantity

$$E := \mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\sqrt{\sigma^2/\sigma_s^2 I_n+\mathbf{h}^t\mathbf{h}}\Lambda_e}(\sigma)\right].$$

In reasonable scenarios, we have $\sigma^2/\sigma_s^2 \ll 1$, and it is easily deduced that $E \underset{\sigma_s \to \infty}{\to} \mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\sqrt{\mathbf{h}^t\mathbf{h}}\Lambda_e}(\sigma)\right] = \mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right]$. Hence, independent of the considered setup, the goal is to design a lattice Λ_e so that $\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right]$ is minimized. We remark that this quantity only depends on Λ_e and σ_h/σ, not for example on Λ_b. Using (1), we can also compute

$$\mathbb{E}_{\mathbf{H}}\left[\varepsilon_{\Lambda_{e,\mathbf{h}}}(\sigma)\right] = |\Lambda_b/\Lambda_e|\mathbb{E}_{\mathbf{H}}\left[\nu_{\Lambda_{b,\mathbf{h}}} g_n(\Lambda_{e,\mathbf{h}};\sigma)\right] - 1,$$

which coincides up to constants with the probability bound (3). This is an important agreement of the probability and information approaches.

We also point out that the agreeing probability bound (3) has been analyzed with error terms in [5] and is asymptotic at poor signal quality. We thus expect the average flatness factor to predict and correctly order the performance of different lattices at the interesting low-SNR regime.

V. Simulation Results

In the following, we make use of Proposition 1 to approximate the average flatness factor of different lattice coset codes numerically. The aim is to compare the predictions of our theta approximation to lattice design heuristics and channel simulations given *e.g.*, in [9]. We focus on the SISO Rayleigh fading channel to allow for a comparison with the results obtained in [9], and hence restrict the choice of channel matrices $\mathbf{h}$ to diagonal matrices.

Even if Eve's information only depends on Λ_e, to ensure a meaningful comparison between codes we fix a superlattice Λ_b (*i.e.*, fix Bob's vector decoding error probability) and the index[4] $|\Lambda_b/\Lambda_e|$ to match the information rate. For convenience, we define the variable $\tau = 1/(2\pi\sigma^2)$, and consider ε_Λ as a function of τ. In the limit of large and small values of τ, our quantity of interest of two different lattices of equal index converge to the same value. Hence, we will center our attention to a range of τ where a difference is visible, which in our plots corresponds to a very large range of values for σ^2.

Remark 2. *It has been recently shown in [9] that in order to minimize the expression* (3) *for the SISO channel, the property of the considered lattice being well-rounded is favorable. As will be visible from the following simulation results, this criterion is also key for minimizing the average flatness factor. When comparing* orthogonal *lattices of fixed volume, this reduces to taking a (close to) square lattice,* i.e.*, $d\mathbb{Z}^n$ is expected to perform best among lattices with diagonal generator matrix and fixed volume $\nu = d^n$.*

We start by comparing two 2-dimensional lattices, Λ_1 and Λ_2 in Figure 2, with generator matrices

$$M_{\Lambda_1} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}, M_{\Lambda_2} = \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix},$$

which we interpret as index-16 sublattices of $\Lambda = \mathbb{Z}^2$. Some characteristics of these lattices are summarized in Table I.

[4]As we are fixing a common superlattice fixing the index is equivalent to comparing sublattices of the same volume.

$\mathbf{n=2}$	$\lambda_{\min}$	$\#\{\lambda \in \Lambda \mid \|\lambda\|^2 = \lambda_{\min}\}$	WR	Index
Λ_1	16	4	Yes	16
Λ_2	1	2	No	16

TABLE I
CHARACTERISTICS OF THE LATTICES IN DIMENSION $n = 2$.

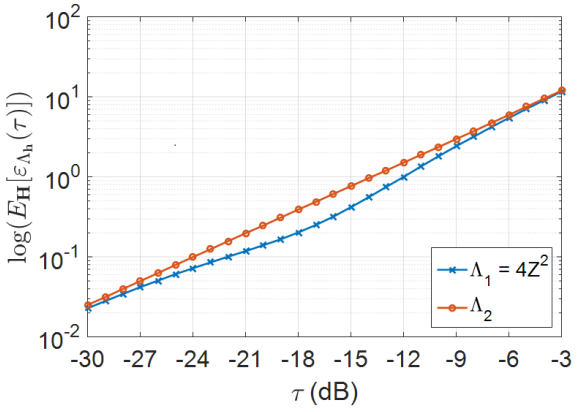

Fig. 2. Average flatness factor of Λ_1 and Λ_2.

The choice of M_{Λ_2} is deliberately bad, hence from the observation in Remark 2 it is not surprising that Λ_1 exhibits a lower average flatness factor than Λ_2. The difference can be as large as 3 dB, which is remarkable without further optimization. This example allows already to hint that both the property of being well-rounded as well as having a shortest vector which is as long as possible is advantageous for minimizing the average flatness factor. This statement will become clearer from the subsequent, more interesting examples.

$\mathbf{n=4}$	$\lambda_{\min}$	$\#\{\mathbf{x} \in \Lambda \mid \|\mathbf{x}\|^2 = \lambda_{\min}\}$	WR	Index
Π_1	8	24	Yes	32
Π_2	1	2	No	32
Ω_1	4	4	No	256
Ω_2	16	8	Yes	256
Ω_3	20	12	Yes	256
Γ_1	4	2	No	302
Γ_2	22	12	Yes	302

TABLE II
CHARACTERISTICS OF THE LATTICES IN DIMENSION $n = 4$.

We consider $\Pi_1 = 2D_4$, a scaled version of the checkerboard lattice, as an index-32 sublattice of $\mathbb{Z}^4$, and compare its average flatness factor to another index-32 sublattice Π_2 of $\mathbb{Z}^4$. The generator matrices are given by

$$M_{\Pi_1} = 2 \cdot \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}, M_{\Pi_2} = \begin{bmatrix} 2 & 0 & -2 & -2 \\ 3 & -5 & -2 & 3 \\ -4 & 8 & 4 & -4 \\ -6 & 5 & 0 & -7 \end{bmatrix}.$$

In a recent article [9], the authors analyze the performance of three different index-256 sublattices of $\mathbb{Z}^4$ with respect to Eve's decoding error probability. We denote the lattices by Ω_1, $\Omega_2 = 4\mathbb{Z}^4$, and Ω_3, with generator matrices given by

$$M_{\Omega_1} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, M_{\Omega_2} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, M_{\Omega_3} = \begin{bmatrix} -2 & -3 & 4 & -1 \\ 0 & -1 & 0 & 3 \\ 0 & -3 & -2 & -3 \\ -4 & -1 & 0 & -1 \end{bmatrix}.$$

In Figure 3 we compare all five presented lattices with respect to their average flatness factor. Note that as remarked above, only a comparison between lattices of the same index is meaningful. The comparison between Π_1 and Π_2 again agrees

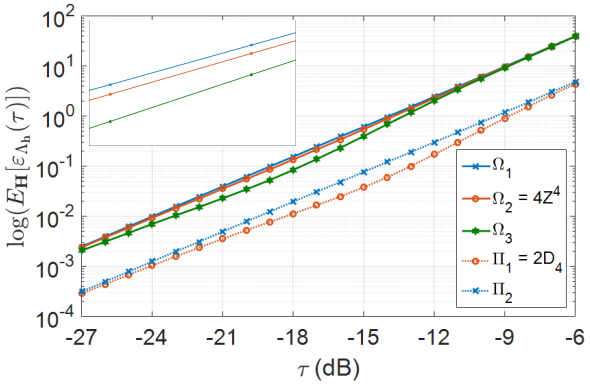

Fig. 3. Average flatness factor of Π_1, Π_2, and of Ω_1, Ω_2, Ω_3.

with what should be expected. Note that it is known that D_4 attains the maximum possible length of the shortest vector among lattices in dimension 4, and is moreover well-rounded.

The order of the lattices Ω_i, $1 \leq i \leq 3$, agrees with the results obtained in [9]. We note that Ω_1 and Ω_2 respect the statement from Remark 2, but their difference is almost negligible. The well-rounded lattice Ω_3, however, is up to 1-1.5 dB ahead of Ω_2.

In addition, we compare two index-302 sublattices of $\mathbb{Z}^4$, of respective generator matrices

$$M_{\Gamma_1} = \begin{bmatrix} -1 & 1 & 2 & 2 \\ -1 & 0 & 2 & -5 \\ 1 & -2 & 5 & -1 \\ -1 & -5 & 1 & 2 \end{bmatrix}, M_{\Gamma_2} = \begin{bmatrix} 1 & 1 & 3 & -2 \\ -4 & 1 & 0 & -4 \\ -1 & -2 & 3 & 1 \\ 2 & -4 & 2 & -1 \end{bmatrix}.$$

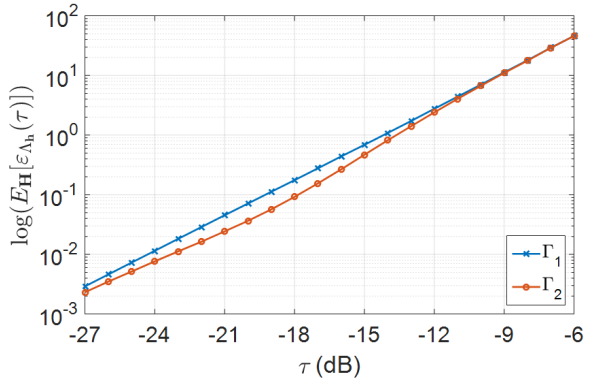

Fig. 4. Average flatness factor of Γ_1 and Γ_2.

As in the previous examples, the well-rounded lattice Γ_2 exhibits a smaller average flatness factor than Γ_1, and the difference reaches up to 2 dB. This is in agreement with the simulation results obtained in [9].

VI. Conclusions

We studied the design of secure lattice coset codes in a general wireless model with any fading and Gaussian noise, and focused on the $\mathrm{mod}\ \Lambda_s$ channel as well as on Gaussian coset coding. Recalling the eavesdropper's probability bounds [5], [6] and deriving a variant of the information bounds [7], [8] we saw that both are an increasing function of the average flatness factor, *i.e.*, the theta series of the faded eavesdroppers lattice $\Lambda_{e,\mathbf{h}}$ averaged over channel realizations $\mathbf{h}$.

We then computed numerically the average flatness factors of different faded lattices with help of an approximation of the theta series of a lattice derived in [10]. By making use of this result, computing the average flatness factor becomes computationally inexpensive, which allows us to avoid further additional approximations based *e.g.*, on the inverse norm sum. As already suggested [9], our findings show that the property of being well-rounded, is necessary in order to minimize the average flatness factor and, hence, the upper bound on the eavesdropper's information. All our numerical computations agreed with the channel simulations in [9].

Interesting further topics of investigation include the more detailed design of optimal secure lattice codes and, for SISO channels, the construction of sequences of codes where the information tends to zero. In both of these cases, well-rounded lattices studied in [9] will likely play a central role.

Acknowledgment

This work is supported by the Academy of Finland under Grants #268364, #276031, #282938 and #283262, as well as a grant from the Finnish Foundation for Technology Promotion.

We gratefully acknowledge the authors of [8] for bringing recent publications to our attention.

References

[1] A. D. Wyner, "The Wire-Tap Channel", *Bell System Technical Journal,* Vol. 54, Oct 1975.

[2] L. H. Ozarow and A. D. Wyner, "The Wire-Tap Channel II", *Bell System Technical Journal,* vol. 63, pp. 2135–2157, 1984.

[3] F. Oggier, P. Sole and J.-C. Belfiore, "Lattice Codes for the Wiretap Gaussian Channel: Construction and Analysis", *IEEE Trans. Inf. Theory*, to appear.

[4] C. Ling, L. Luzzi, J.-C. Belfiore, D. Stehle, "Semantically Secure Lattice Codes for the Gaussian Wiretap Channel", *IEEE Trans. Inf. Theory*, vol.60, no.10, pp. 6399–6416, 2014.

[5] J.-C. Belfiore and F. Oggier, "Lattice Code Design for the Rayleigh Fading Wiretap Channel", *Proc. IEEE ICC*, 2011.

[6] J.-C. Belfiore and F. Oggier, "An Error Probability Approach to MIMO Wiretap Channels", *IEEE Trans. Inf. Theory*, vol.61, no.8, pp. 3396–3403, 2013.

[7] H. Mirghasemi and J.-C. Belfiore, "Lattice Code Design Criterion For MIMO Wiretap Channels", *Proc. IEEE ITW*, 2015.

[8] L. Luzzi, C. Ling and R. Vehkalahti, "Almost Universal Codes for Fading Wiretap Channels", arXiv:1601.02391, 2016.

[9] O. Gnilke, H. Tran and A. Karrila and C. Hollanti, "Well-Rounded Lattices for Reliability and Security in Rayleigh Fading SISO Channels", arXiv:1605.0041. To appear in Proc. IEEE ITW, 2016.

[10] A. Barreal, D. Karpuk and C. Hollanti, "Decoding in Compute-and-Forward Relaying: Real Lattices and the Flatness of Lattice Sums", arXiv:1601.05596, 2016.