Almost universal codes for fading wiretap channels

Laura Luzzi
Laboratoire ETIS
(ENSEA - UCP - CNRS)
Cergy-Pontoise, France
laura.luzzi@ensea.fr

Cong Ling
Department of Electrical
and Electronic Engineering
Imperial College London, U.K.
c.ling@imperial.ac.uk

Roope Vehkalahti
Department of Mathematics
and Statistics,
University of Turku, Finland
roiive@utu.fi

***Abstract*—We consider a fading wiretap channel model where the transmitter has only statistical channel state information, and the legitimate receiver and eavesdropper have perfect channel state information. We propose a sequence of non-random lattice codes which achieve strong secrecy and semantic security over ergodic fading channels. The construction is almost universal in the sense that it achieves the same constant gap to secrecy capacity over Gaussian and ergodic fading models.**

I. Introduction

The wiretap channel model was introduced by Wyner [21], who showed that secure and reliable communication can be achieved simultaneously over noisy channels even without the use of secret keys. In the information theory community, the most widely accepted secrecy metric is Csiszár's *strong secrecy*: the mutual information $\mathbb{I}(\mathsf{M};\mathsf{Z}^n)$ between the confidential message M and the channel output Z^n should vanish when the code length n tends to infinity.

While in the information theory community confidential messages are often assumed to be uniformly distributed, this assumption is not accepted in cryptography. A cryptographic treatment of the wiretap channel was proposed in [3] to combine the requirements of the two communities, establishing that achieving *semantic security* in the cryptographic sense is equivalent to achieving strong secrecy for all distributions of the message. This equivalence holds also for continuous channels [10].

In the case of Gaussian wiretap channels, [10] considered the problem of designing lattice codes which achieve strong secrecy and semantic security. Following an approach by Csiszár [5, 4], strong secrecy is guaranteed if the output distributions of the eavesdropper's channel corresponding to two different messages are indistinguishable in the sense of variational distance. Moreover, the *flatness factor* of a lattice was proposed in [10] as a fundamental criterion which implies that conditional outputs are indistinguishable. Using random coding arguments, it was shown that there exist families of lattice codes which are "good for secrecy", meaning that their flatness factor is vanishing, and achieve semantic security for rates up to $1/2$ nat from the secrecy capacity.

In this paper, we consider a fading wiretap channel model where the transmitter has only access to statistical channel state information (CSI), while the legitimate receiver and the eavesdropper both have perfect knowledge of their own channels. We extend the criterion based on the flatness factor to the case of fading channels and propose a family of non-random lattice codes from algebraic number fields satisfying this criterion. We note that ideal lattices from number fields were already considered for secrecy under an error probability criterion for Gaussian and fading channels in [1, 2, 8, 16].

In this work, we consider a particular sequence of algebraic number fields with constant root discriminant. In [20, 11], it was shown that these lattice codes are "almost universal" in the sense that they achieve a constant gap to channel capacity over *any* ergodic stationary fading channel. The underlying multiplicative structure and constant root discriminant property guarantee that the received lattice after fading has a good minimum distance when the channel is not in outage.

The sequences of number fields that we consider are also used in the crypto literature for worst-case to average-case reductions of hard lattice problems [18].

In this paper, we show that these lattices also achieve strong secrecy and semantic security. The key feature is that the *dual* of the faded lattice has good minimum distance, so that the average flatness factor of the faded lattice vanishes.

In particular, for the Gaussian case this suggests a simple design criterion where the packing density of the lattice and its dual should be maximized simultaneously. We note that the dual code also plays a role in the design of LDPC codes for binary erasure wiretap channels [19].

We also improve the rate of almost universal codes by replacing spherical shaping with a discrete Gaussian distribution over the infinite lattice as in [10]. As a consequence, our nested lattice schemes achieve the same constant gap to secrecy capacity over all static and ergodic fading models.

The proposed lattice codes can be generalized in a straightforward manner to the multi-antenna case using the multiblock matrix lattices from division algebras in [11]. This generalization will be presented in an upcoming journal version.

II. Preliminaries

A. Flatness factor and discrete Gaussian distribution

In this section, we define some fundamental lattice parameters that will be used in the rest of the paper. For more background about the smoothing parameter and the flatness factor in information theory and cryptography, we refer the reader to [15, 10, 17].

Consider $\mathbb{C}^k$ as a $2k$-dimensional real vector space with a real inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \Re(\mathbf{x}^\dagger \mathbf{y})$. This inner product naturally

defines a metric on $\mathbb{C}^k$ by setting $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. [1]
Given a complex lattice $\Lambda \subset \mathbb{C}^k$, we define the dual lattice as

$$\Lambda^* = \{\mathbf{x} \in \mathbb{C}^k \mid \forall \mathbf{y} \in \Lambda, \ \ \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}.$$

Let $f_{\sqrt{\Sigma},\mathbf{c}}(\mathbf{z})$ denote the k-dimensional complex normal distribution with mean $\mathbf{c}$ and covariance matrix Σ:

$$f_{\sqrt{\Sigma},\mathbf{c}}(\mathbf{z}) = \frac{1}{\pi^k \det(\Sigma)} e^{-(\mathbf{z}-\mathbf{c})^\dagger \Sigma^{-1} (\mathbf{z}-\mathbf{c})} \quad \forall \mathbf{z} \in \mathbb{C}^k.$$

We use the notation $f_{\sigma,\mathbf{c}}(\mathbf{z})$ for $f_{\sigma I,\mathbf{c}}(\mathbf{z})$ and $f_{\sqrt{\Sigma}}$ for $f_{\sqrt{\Sigma},0}$.

Definition 1: Given a complex lattice $\Lambda \subset \mathbb{C}^k$, the *flatness factor* $\epsilon_\Lambda(\sqrt{\Sigma})$ is defined as the maximum deviation of the Gaussian distribution over Λ from the uniform distribution over a fundamental region $\mathcal{R}(\Lambda)$ of Λ, with volume $V(\Lambda)$:

$$\epsilon_\Lambda(\sqrt{\Sigma}) = \max_{\mathbf{z} \in \mathcal{R}(\Lambda)} \left| V(\Lambda) \sum_{\boldsymbol{\lambda} \in \Lambda} f_{\sqrt{\Sigma},\boldsymbol{\lambda}}(\mathbf{z}) - 1 \right|.$$

Compared to [10], in this paper we use an extended version of the flatness factor for correlated Gaussians, related to the extended notion of the smoothing parameter in [17].
Note that correlations can be absorbed by the lattice in the sense that $\epsilon_\Lambda(\sqrt{\Sigma}) = \epsilon_{\sqrt{\Sigma}^{-1}\Lambda}(I)$, and that $\epsilon_\Lambda(\sqrt{\Sigma_1}) \leq \epsilon_\Lambda(\sqrt{\Sigma_2})$ if Σ_1 and Σ_2 are positive definite with $\Sigma_1 \succeq \Sigma_2$.

Definition 2: Given a lattice Λ and $\epsilon > 0$, the *smoothing parameter*[2] $\eta_\epsilon(\Lambda)$ is the smallest $s = \sqrt{2\pi}\sigma > 0$ such that $\sum_{\boldsymbol{\lambda}^* \in \Lambda^* \setminus \{\mathbf{0}\}} e^{-\pi^2 \sigma^2 \|\boldsymbol{\lambda}^*\|^2} \leq \epsilon$, where Λ^* is the dual lattice.
To extend the definition to matrices we can say that

$$\sqrt{2\pi\Sigma} \succeq \eta_\epsilon(\Lambda) \quad \text{if} \quad \epsilon_\Lambda(\Sigma) \leq \epsilon. \tag{1}$$

The smoothing parameter is upper bounded by the minimum distance of the dual lattice [15]:

$$\eta_\epsilon(\Lambda) \leq \frac{2\sqrt{k}}{\lambda_1(\Lambda^*)}. \tag{2}$$

Finally, given $\mathbf{c} \in \mathbb{C}^k$ and $\sigma > 0$, we define the *discrete Gaussian distribution* over the (shifted) lattice $\Lambda - \mathbf{c} \subset \mathbb{C}^k$ as the following discrete distribution taking values in $\Lambda - \mathbf{c}$:

$$D_{\Lambda-\mathbf{c},\sigma}(\boldsymbol{\lambda} - \mathbf{c}) = \frac{f_\sigma(\boldsymbol{\lambda} - \mathbf{c})}{f_{\sigma,\mathbf{c}}(\Lambda)}.$$

The following result is a consequence of [17, Theorem 3.1] and extends Lemma 8 in [10]:

Lemma 1: Let X_1 be sampled according to the discrete Gaussian distribution $D_{\Lambda+\mathbf{c},\sqrt{\Sigma_1}}$ and X_2 be sampled according to the continuous Gaussian $f_{\sqrt{\Sigma_2}}$. Let $\Sigma_0 = \Sigma_1 + \Sigma_2$ and $\Sigma^{-1} = \Sigma_1^{-1} + \Sigma_2^{-1}$. If

$$\epsilon_\Lambda(\sqrt{\Sigma}) \leq \epsilon \leq \frac{1}{2}, \tag{3}$$

then the distribution g of $\mathsf{X} = \mathsf{X}_1 + \mathsf{X}_2$ is close to $f_{\sqrt{\Sigma_0}}$:

$$\mathbb{V}(g, f_{\sqrt{\Sigma_0}}) \leq 4\epsilon,$$

where $\mathbb{V}(\,,\,)$ is the L^1 distance.

B. Ideal lattices from number fields with constant root discriminant

Let F be a number field of degree $[F : \mathbb{Q}] = n$, with ring of integers $\mathcal{O}_F$. We denote by d_F the discriminant of the number field. We define the *codifferent* of F as

$$\mathcal{O}_F^\vee = \{x \in K : \mathrm{Tr}_{F/\mathbb{Q}}(x\mathcal{O}_F) \subseteq \mathbb{Z}\}.$$

The codifferent is a fractional ideal, that is, there exists some integer a such that $a\mathcal{O}_F^\vee$ is a proper ideal of $\mathcal{O}_F$, and its algebraic norm is the inverse of the discriminant:

$$N(\mathcal{O}_F^\vee) = 1/d_F. \tag{4}$$

We focus on the case of totally complex extensions $F/\mathbb{Q}$ of degree $n = 2k$. The *relative canonical embedding* of F into $\mathbb{C}^k$ is given by

$$\psi(x) = (\sigma_1(x), \ldots, \sigma_k(x)),$$

where $\{\sigma_1, \ldots, \sigma_k\}$ is a set of $\mathbb{Q}$-embeddings $F \to \mathbb{C}$ such that we have chosen one from each complex conjugate pair. Then $\Lambda = \psi(\mathcal{O}_F)$ is a lattice in $\mathbb{C}^k$. The codifferent embeds as the complex conjugate of the dual lattice:

$$\Lambda^* = 2\overline{\psi(\mathcal{O}_F^\vee)}. \tag{5}$$

Using (2), we obtain

$$\eta_\epsilon(\Lambda) \leq \frac{\sqrt{k}}{\lambda_1(\overline{\psi(\mathcal{O}_F^\vee)})}. \tag{6}$$

From the AM-GM inequality we have that for any fractional ideal $\mathcal{I}$ of $\mathcal{O}_F$,

$$\lambda_1(\psi(\mathcal{I})) \geq \sqrt{k}(N(\mathcal{I}))^{\frac{1}{2k}}.$$

In particular, from (4) we get

$$\lambda_1(\overline{\psi(\mathcal{O}_F^\vee)}) = \lambda_1(\psi(\mathcal{O}_F^\vee)) \geq \frac{\sqrt{k}}{|d_F|^{\frac{1}{2k}}}. \tag{7}$$

Combining equations (6) and (7), we find that the smoothing parameter of Λ is upper bounded by the root discriminant [18, Lemma 6.5]: given $\epsilon = 2^{-2k}$,

$$\eta_\epsilon(\Lambda) \leq |d_F|^{\frac{1}{2k}}. \tag{8}$$

The following theorem by Martinet [13] proves the existence of infinite towers of totally complex number fields with constant root discriminant:

Theorem 2.1: There exists an infinite tower of totally complex number fields $\{F_k\}$ of degree $2k = 5 \cdot 2^t$, such that

$$|d_{F_k}|^{\frac{1}{2k}} = G, \tag{9}$$

for $G \approx 92.368$.

[1] This inner product corresponds to identifying $\mathbb{C}^k$ with $\mathbb{R}^{2k}$ with the canonical real inner product, through the isometry $\phi(z_1, \ldots, z_k) = (\Re(z_1), \ldots, \Re(z_k), \Im(z_k), \ldots, \Im(z_k))$. Note also that if $\Sigma = \Sigma^\dagger$, then $\langle \mathbf{z}, \Sigma\mathbf{z} \rangle = \Re(\mathbf{z}^\dagger \Sigma \mathbf{z}) = \mathbf{z}^\dagger \Sigma \mathbf{z} = \phi(\mathbf{z})^T \Sigma_\mathbb{R} \phi(\mathbf{z})$, where $\Sigma_\mathbb{R} = \begin{pmatrix} \Re(\Sigma) & -\Im(\Sigma) \\ \Im(\Sigma) & \Re(\Sigma) \end{pmatrix}$. In particular, the properties of real Gaussian distributions carry over to circularly symmetric complex Gaussian distributions.

[2] Note that we define the smoothing parameter per complex dimension, which differs by a factor $\sqrt{2}$ from the definition in [15]. We have adjusted the bounds on $\eta_\epsilon(\Lambda)$ accordingly.

We now focus on the corresponding lattice sequence $\Lambda^{(k)} \subset \mathbb{C}^k$. Their volume is a function of the discriminant:

$$\mathrm{Vol}(\Lambda^{(k)}) = 2^{-k}\sqrt{|d_F|} = 2^{-k}G^k \tag{10}$$

Let $\epsilon = 2^{-2k}$. From Theorem 2.1 and equation (8),

$$\eta_\epsilon(\Lambda^{(k)}) \leq |d_F|^{\frac{1}{2k}} = G.$$

Since the flatness factor is a decreasing function of σ,

$$\forall \sigma > \frac{G}{\sqrt{2\pi}}, \quad \varepsilon_{\Lambda^{(k)}}(\sigma) \leq 2^{-2k}. \tag{11}$$

III. Fading wiretap channel

We consider an ergodic fading channel model where the outputs Y^k and Z^k at Bob and Eve's end are given by

$$\begin{cases} \mathsf{Y}_i = \mathsf{H}_{b,i}\mathsf{X}_i + \mathsf{W}_{b,i}, \\ \mathsf{Z}_i = \mathsf{H}_{e,i}\mathsf{X}_i + \mathsf{W}_{e,i}, \end{cases} \qquad i = 1, \dots, k \tag{12}$$

where $\mathsf{W}_{b,i}$, $\mathsf{W}_{e,i}$ are i.i.d. complex Gaussian vectors with zero mean and variance σ_b^2, σ_e^2 per complex dimension. The input X^k satisfies the average power constraint

$$\frac{1}{k}\sum_{i=1}^{k} |\mathsf{X}_i|^2 \leq P. \tag{13}$$

We suppose that $\mathsf{H}_{b,i}$, $\mathsf{H}_{e,i}$ are isotropically invariant channels such that the channel capacities C_b and C_e are well-defined and the weak law of large numbers holds: $\forall \delta > 0$,

$$\lim_{k\to\infty} \mathbb{P}\left\{ \left| \frac{1}{k}\sum_{i=1}^{k} \ln\left(1 + \frac{P}{\sigma_b^2}|h_{b,i}|^2\right) - C_b \right| > \delta \right\} = 0, \tag{14}$$

$$\lim_{k\to\infty} \mathbb{P}\left\{ \left| \frac{1}{k}\sum_{i=1}^{k} \ln\left(1 + \frac{P}{\sigma_e^2}|h_{e,i}|^2\right) - C_e \right| > \delta \right\} = 0. \tag{15}$$

All rates are expressed in nats per complex channel use.

We suppose that Alice has no instantaneous CSI (apart from knowledge of channel statistics), and Bob and Eve have perfect CSI of their own channels. A confidential message M and an auxiliary message M' with rate R and R' respectively are encoded into X^k. We denote by $\hat{\mathsf{M}}$ the estimate of the confidential message at Bob's end. With slight abuse of notation, we define $\mathsf{H}_e = \mathrm{diag}(\mathsf{H}_{e,1}, \dots, \mathsf{H}_{e,k})$, $\mathsf{H}_b = \mathrm{diag}(\mathsf{H}_{b,1}, \dots, \mathsf{H}_{b,k})$.

Definition 3: A coding scheme achieves *strong secrecy* if

$$\lim_{k\to\infty} \mathbb{P}\{\hat{\mathsf{M}} \neq \mathsf{M}\} = 0, \qquad \text{(reliability condition)}$$

$$\lim_{k\to\infty} \mathbb{I}(\mathsf{M}; \mathsf{Z}^k, \mathsf{H}_e) = 0. \qquad \text{(secrecy condition)}$$

The secrecy capacity for this wiretap model is given by [9]

$$C_s = C_b - C_e. \tag{16}$$

Let $\Lambda^{(k)} \subset \mathbb{C}^k$ be the lattice sequence defined in the previous section. We consider scaled versions $\Lambda_b = \alpha_b\Lambda^{(k)}$, $\Lambda_e = \alpha_e\Lambda^{(k)}$ such that $\Lambda_e \subset \Lambda_b$ and $|\Lambda_b/\Lambda_e| = e^{kR}$.

We consider the secrecy scheme in [10], where each confidential message $m \in \mathcal{M} = \{1, \dots, e^{kR}\}$ is associated to a coset leader $\boldsymbol{\lambda}_m \in \Lambda_b \cap \mathcal{R}(\Lambda_e)$ for a fundamental region $\mathcal{R}(\Lambda_e)$. To transmit the message m, Alice samples X^k from the discrete Gaussian $D_{\Lambda_e + \boldsymbol{\lambda}_m, \sigma_s}$ with $\sigma_s^2 = P$. It follows from [10, Lemma 6] that as $k \to \infty$, the variance per complex dimension of X^k tends to P provided that

$$\lim_{k\to\infty} \epsilon_{\Lambda_e}(\sqrt{P}) = 0. \tag{17}$$

From [10, Lemma 7], the information rate R' of the auxiliary message (corresponding to the choice of a point in Λ_e) is

$$R' \approx \ln(\pi e P) - \frac{1}{k}\ln V(\Lambda_e) = \ln(\pi e P) - \frac{1}{k}\ln(\alpha_e^{2k} 2^{-k} G^k).$$

Therefore, we have

$$\alpha_e^2 \approx \frac{2\pi e P}{G e^{R'}}. \tag{18}$$

From (11), $\epsilon_{\Lambda_e}(\sqrt{P}) = \epsilon_{\alpha_e\Lambda}(\sqrt{P}) = \epsilon_\Lambda\left(\sqrt{P}/\alpha_e\right) \to 0$ provided that $\frac{\sqrt{P}}{\alpha_e} > \frac{G}{\sqrt{2\pi}}$, and (17) holds for

$$R' > \ln(eG) = \ln(G) + 1. \tag{19}$$

We now state the main result of the paper which will be proven in the following sections:

Proposition 1: The proposed wiretap coding scheme with $\sigma_s^2 = P$ achieves strong secrecy for any message distribution p_{M} (and thus semantic security) for any secrecy rate

$$R < C_b - C_e - \ln\left(2G^2/\pi\right).$$

A. Secrecy

The received lattice at Eve's end is $\mathsf{H}_e\Lambda$. Since M and H_e are independent, the leakage can be expressed as follows:

$$\mathbb{I}(\mathsf{M}; \mathsf{Z}^k, \mathsf{H}_e) = \mathbb{I}(\mathsf{M}; \mathsf{H}_e) + \mathbb{I}(\mathsf{M}; \mathsf{Z}^k|\mathsf{H}_e) = \mathbb{I}(\mathsf{M}; \mathsf{Z}^k|\mathsf{H}_e) =$$
$$= \mathbb{E}_{\mathsf{H}_e}\left[\mathbb{I}(p_{\mathsf{M}|\mathsf{H}_e}; p_{\mathsf{Z}^k|\mathsf{H}_e})\right] = \mathbb{E}_{\mathsf{H}_e}\left[\mathbb{I}(p_{\mathsf{M}}; p_{\mathsf{Z}^k|\mathsf{H}_e})\right]$$

We want to show that the *average* leakage with respect to the fading is small. In order to do so, we will show that the output distributions $p_{\mathsf{Z}^k|\mathsf{H}_e}$ are close to a Gaussian distribution with high probability. For a fixed realization $H_e = \mathrm{diag}(h_{e,1}, \dots, h_{e,k})$, $H_e\mathsf{X}^k \sim D_{H_e\Lambda_e + H_e\boldsymbol{\lambda}_m, \sqrt{H_e H_e^\dagger}\sqrt{P}}$. Using Lemma 1 with $\Sigma_1 = H_e H_e^\dagger P$, $\Sigma_2 = \sigma_b^2 I$, we have

$$\mathbb{V}(p_{\mathsf{Z}^k|H_e}, f_{\Sigma_0}) \leq \epsilon \tag{20}$$

provided that

$$\varepsilon_{H_e\Lambda_e}(\sqrt{\Sigma}) = \varepsilon_{\sqrt{\Sigma}^{-1}H_e\Lambda_e}(1) \leq \epsilon \leq \frac{1}{2}, \tag{21}$$

where we define $\Sigma_0 = H_e H_e^\dagger P + \sigma_b^2 I$, $\Sigma = \frac{(H_e H_e^\dagger)^{-1}}{P} + \frac{I}{\sigma_b^2}$. If (20) holds, then it follows from [10, Lemma 2] that

$$\mathbb{I}(p_{\mathsf{M}}; p_{\mathsf{Z}^k|H_e}) \leq 8k\epsilon R - 8\epsilon\log 8\epsilon. \tag{22}$$

Recalling the upper bound (2), we have

$$\eta_\epsilon(\sqrt{\Sigma^{-1}}H_e\Lambda) \leq \frac{2\sqrt{k}}{\lambda_1(\sqrt{\Sigma}(H_e^\dagger)^{-1}\Lambda^*)}. \tag{23}$$

Using (5) and the arithmetic mean - geometric mean inequality,

$$\lambda_1(\sqrt{\Sigma}(H_e^\dagger)^{-1}\Lambda^*) = 2\lambda_1(\sqrt{\Sigma}(H_e^\dagger)^{-1}\overline{\psi(\mathcal{O}_F^\vee)}) =$$

$$= 2 \min_{x \in \mathcal{O}_F^\vee \setminus \{0\}} \left\| \sqrt{\Sigma} (H_e^\dagger)^{-1} \overline{\psi(x)} \right\| \geq$$

$$\geq 2 \min_{x \in \mathcal{O}_F^\vee \setminus \{0\}} \sqrt{k} \prod_{i=1}^k \left(\frac{P\sigma_e^2}{\sigma_e^2 + P|h_{e,i}|^2} \right)^{\frac{1}{2k}} \prod_{i=1}^k |\sigma_i(x)|^{\frac{1}{k}} =$$

$$= \frac{2\sqrt{k}\sqrt{P}\sigma_e}{G \prod_{i=1}^k (\sigma_e^2 + P|h_{e,i}|^2)^{\frac{1}{2k}}}.$$

The last equality follows from the fact that

$$\min_{x \in \mathcal{O}_F^\vee \setminus \{0\}} \prod_{i=1}^k |\sigma_i(x)|^{\frac{1}{k}} = \min_{a \in \mathcal{O}_F^\vee \setminus \{0\}} \left| N_{K/\mathbb{Q}}(a) \right|^{\frac{1}{2k}} =$$

$$= N(\mathcal{O}_F^\vee)^{\frac{1}{2k}} = \frac{1}{|d_F|^{1/2k}} = \frac{1}{G}. \quad (24)$$

Replacing in (23), we find that for $\epsilon = 2^{-2k}$,

$$\eta_\epsilon(\sqrt{\Sigma^{-1}} H_e \Lambda) \leq G \prod_{i=1}^k (\sigma_e^2 + P|h_{e,i}|^2)^{\frac{1}{2k}} / \sqrt{P}\sigma_e.$$

Equivalently, in terms of flatness factor we have

$$\varepsilon_{\sqrt{\Sigma^{-1}} H_e \Lambda} \left(\frac{G \prod_{i=1}^k (\sigma_e^2 + P|h_{e,i}|^2)^{\frac{1}{2k}}}{\sqrt{2\pi P}\sigma_e} \right) \leq 2^{-2k}$$

for fixed fading H_e. Given $\delta > 0$, the law of large numbers (15) implies that $\mathbb{P}\left\{ \prod_{i=1}^k \left(1 + \frac{P}{\sigma_e^2}|h_{e,i}|^2\right)^{\frac{1}{k}} > e^{C_e + \delta} \right\} \to 0$. Now suppose that

$$\alpha_e G e^{\frac{C_e+\delta}{2}} / \sqrt{2\pi P} \leq 1. \quad (25)$$

We can bound the leakage as follows:

$$\mathbb{E}_{\mathsf{H}_e} \left[\mathbb{I}(p_{\mathsf{M}}; p_{\mathsf{Z}^k|\mathsf{H}_e}) \right] \leq$$

$$\leq \mathbb{P}\Big\{ \prod_{i=1}^k \left(1 + \frac{P|h_{e,i}|^2}{\sigma_e^2}\right)^{\frac{1}{k}} > e^{C_e+\delta} \Big\} (kR) +$$

$$+ \mathbb{E}_{\mathsf{H}_e} \left[\mathbb{I}(p_{\mathsf{M}}; p_{\mathsf{Z}^k|\mathsf{H}_e}) \;\Big|\; \prod_{i=1}^k \left(1 + \frac{P|h_{e,i}|^2}{\sigma_e^2}\right)^{\frac{1}{k}} \leq e^{C_e+\delta} \right] \quad (26)$$

The first term vanishes when $k \to \infty$.
Now consider the second term. Under the hypothesis that $\prod_{i=1}^k \left(1 + \frac{P}{\sigma_e^2}|h_{e,i}|^2\right)^{\frac{1}{k}} \leq e^{C_e+\delta}$, we have

$$\varepsilon_{\sqrt{\Sigma^{-1}} H_e \Lambda_e}(1) = \varepsilon_{\alpha_e \sqrt{\Sigma^{-1}} H_e \Lambda}(1) \leq \varepsilon_{\sqrt{\Sigma^{-1}} H_e \Lambda} \left(\frac{G e^{\frac{C_e+\delta}{2}}}{\sqrt{2\pi P}} \right) \leq$$

$$\leq \varepsilon_{\sqrt{\Sigma^{-1}} H_e \Lambda} \left(\frac{G \prod_{i=1}^k (\sigma_e^2 + P|h_{e,i}|^2)^{\frac{1}{2k}}}{\sqrt{2\pi P}\sigma_e} \right) \leq 2^{-2k}.$$

Using (22), the second term is also vanishing and the lattice coding scheme achieves strong secrecy over Eve's channel.
From the conditions (25) and (18), we find that in order to have strong secrecy we need $eGe^{C_e+\delta} \leq e^{R'}$, or equivalently $R' \geq C_e + \delta + 1 + \ln(G)$. Since this is true for any $\delta > 0$, we find that a rate

$$R' \geq C_e + 1 + \ln(G). \quad (27)$$

is required for strong secrecy.

Remark 1: Although we focused on ergodic fading, the same scheme achieves strong secrecy over the Gaussian and static fading wiretap channels. In fact, for these models the first term in (26) is zero, and the second term still vanishes.

B. Reliability

Let $\mathbf{y} = H_b\mathbf{x} + \mathbf{w}_b$ be the received signal at Bob. We suppose that Bob performs MMSE-GDFE preprocessing as in [6]: let $\rho_b = \frac{P}{\sigma_b^2}$, and consider the QR decomposition

$$\widetilde{H}_b = \begin{pmatrix} H_b \\ \frac{1}{\rho_b} I \end{pmatrix} = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} R.$$

Observe that $\|\mathbf{y} - H_b\mathbf{x}\|^2 + \frac{1}{\rho_b}\|\mathbf{x}\|^2 = \left\| Q_1^\dagger \mathbf{y} - R\mathbf{x} \right\|^2 + C$, where C is some constant which does not depend on $\mathbf{x}$.
Since the distribution of $\mathbf{x}$ is not uniform, MAP decoding is not equivalent to ML. However, similarly to [10, Theorem 5], for fixed H_b which is known at the receiver, the result of MAP decoding can be written as

$$\hat{\mathbf{x}}_{\text{MAP}} = \operatorname*{argmax}_{\mathbf{x} \in \Lambda_b} p(\mathbf{x}|\mathbf{y}) = \operatorname*{argmax}_{\mathbf{x} \in \Lambda_b} p(\mathbf{x}) p(\mathbf{y}|\mathbf{x}) =$$

$$= \operatorname*{argmax}_{\mathbf{x} \in \Lambda_b} e^{-\frac{\|\mathbf{x}\|^2}{2P}} e^{-\frac{\|\mathbf{y} - H_b\mathbf{x}\|^2}{2\sigma_b^2}} =$$

$$= \operatorname*{argmin}_{\mathbf{x} \in \Lambda_b} \left(\frac{1}{\rho_b}\|\mathbf{x}\|^2 + \|\mathbf{y} - H_b\mathbf{x}\|^2 \right) = \operatorname*{argmin}_{\mathbf{x} \in \Lambda_b} \left\| Q_1^\dagger \mathbf{y} - R\mathbf{x} \right\|^2$$

Thus, Bob can compute

$$\mathbf{y}' = Q_1^\dagger \mathbf{y} = R\mathbf{x} + \mathbf{v},$$

where $\mathbf{v} = Q_1^\dagger \mathbf{w}_b - \frac{1}{\rho_b}(R^{-1})^\dagger \mathbf{x}$ [6]. The noise $\mathbf{v}$ is the sum of a discrete Gaussian with distribution $D_{\Lambda', \sqrt{\Sigma_1}}$, where $\Lambda' = \frac{1}{\rho_b}(R^{-1})^\dagger \Lambda_b$, $\Sigma_1 = \frac{\sigma_b^2}{\rho_b}(RR^\dagger)^{-1}$, and of a continuous Gaussian random variable $f_{\sqrt{\Sigma_2}}$, where $\Sigma_2 = \sigma_b^2 Q_1 Q_1^\dagger$.

For any message $m \in \mathcal{M}$, $P_e(m) \leq \mathbb{P}\{\mathbf{v} \notin \mathcal{V}(R\Lambda_b)\}$ and consequently the same upper bound holds for the the average:

$$P_e = \sum_{m \in \mathcal{M}} P_e(m) p(m) \leq \mathbb{P}\{\mathbf{v} \notin \mathcal{V}(R\Lambda_b)\}.$$

Although $\mathbf{v}$ is not Gaussian, we will show that its tails behave similarly to a Gaussian random variable.
A random vector $\mathbf{z}$ taking values in $\mathbb{C}^k$ is *δ-subgaussian* with parameter σ if $\forall \mathbf{t} \in \mathbb{C}^k$, $\mathbb{E}[e^{\Re(\mathbf{t}^\dagger \mathbf{z})}] \leq e^\delta e^{\frac{\sigma^2}{2}\|\mathbf{t}\|^2}$. Note that for a complex Gaussian vector $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}}(0, \Sigma)$, $\mathbb{E}[e^{\Re(\mathbf{t}^\dagger \mathbf{z})}] = e^{\frac{1}{2}\mathbf{t}^\dagger \Sigma \mathbf{t}}$.
Suppose that a fixed message m has been transmitted, so that $\mathsf{X}^k \sim D_{\Lambda_e + \boldsymbol{\lambda}_m, \sqrt{P}}$. The following result holds (see also [14, Lemma 2.8]:).

Lemma 2: Let $\mathsf{X}^k \sim D_{\Lambda + \mathbf{c}, \sigma}$ be a k-dimensional discrete complex Gaussian random variable, and let $A \in M_k(\mathbb{C})$. Suppose that $\epsilon_\Lambda(\sigma) < 1$. Then $\forall \mathbf{t} \in \mathbb{C}^k$,

$$\mathbb{E}[e^{\Re(\mathbf{t}^\dagger A \mathbf{x})}] \leq \left(\frac{1 + \epsilon_\Lambda(\sigma)}{1 - \epsilon_\Lambda(\sigma)} \right) e^{\frac{\sigma^2}{2}\|A^\dagger \mathbf{t}\|^2}.$$

It follows that X^k is δ-subgaussian with parameter $\sqrt{P}$ for $\delta = \ln\left(\frac{1+\epsilon}{1-\epsilon}\right)$ provided that $\epsilon = \epsilon_{\Lambda_e}(\sqrt{P}) < 1$, which is

guaranteed by (19). This is weaker than the condition (27) we have already imposed for secrecy, so it doesn't affect the achievable rate. Consequently, for the equivalent noise $\mathbf{v}$,

$$\mathbb{E}[e^{\Re(\mathbf{t}^\dagger \mathbf{v})}] = \mathbb{E}\left[e^{\Re(\mathbf{t}^\dagger Q_1^\dagger \mathbf{w}_b)}\right] \mathbb{E}\left[e^{-\Re\left(\frac{1}{\rho_b}\mathbf{t}^\dagger (R^{-1})^\dagger \mathbf{x}\right)}\right] \leq$$

$$\leq \left(\frac{1+\epsilon}{1-\epsilon}\right) e^{\frac{\sigma_b^2}{2}\mathbf{t}^\dagger \left(Q_1^\dagger Q_1 + \frac{1}{\rho_b}(R^{-1})^\dagger R^{-1}\right)\mathbf{t}} = \left(\frac{1+\epsilon}{1-\epsilon}\right) e^{\frac{\sigma_b^2}{2}\|t\|^2}.$$

This implies that the tails of $\mathbf{v}$ vanish exponentially fast: from [7, Theorem 2.1], it follows that $\forall t > 0$,

$$\mathbb{P}\left\{\|\mathbf{v}\|^2 / k\sigma_b^2 > 1 + 2\sqrt{t/k} + 2t\right\} \leq e^\delta e^{-t}.$$

In particular, taking $\eta = \sqrt{\frac{t}{k}}$, we find that $\forall \eta > 0$,

$$\mathbb{P}\left\{\|\mathbf{v}\|^2 / k\sigma_b^2 > 1 + \eta\right\} \leq e^\delta e^{-k\eta^2}.$$

Let d_R denote the minimum distance in the received lattice:

$$d_R^2 = \min_{\boldsymbol{\lambda}\in\Lambda_b\setminus\{0\}} \sum_{i=1}^k |R_i\lambda_i|^2 = \min_{x\in\mathcal{O}_F\setminus\{0\}} \alpha_b^2 \sum_{i=1}^k |R_i\sigma_i(x)|^2 \geq$$

$$\geq \min_{x\in\mathcal{O}_F\setminus\{0\}} \alpha_b^2 k \prod_{i=1}^k \left(\frac{1}{\rho_b} + |h_{b,i}|^2\right)^{\frac{1}{k}} \prod_{i=1}^k |\sigma_i(x)|^{\frac{1}{k}} \geq$$

$$\geq \alpha_b^2 k \prod_{i=1}^k \left(\frac{1}{\rho_b} + |h_{b,i}|^2\right)^{\frac{1}{k}}. \tag{28}$$

The previous bound follows from the AM-GM inequality and the fact that the minimum non-zero norm of the code is 1. We use the same argument as in [11] to bound P_e: given $\eta > 0$,

$$P_e \leq \mathbb{P}\{\mathbf{v} \notin \mathcal{V}(R\Lambda_b)\} \leq \mathbb{P}\{\mathbf{v} \notin \mathcal{B}(d_R/2)\} \leq$$

$$\leq \mathbb{P}\left\{\frac{\|\mathbf{v}\|^2}{k\sigma_b^2} \geq 1+\eta\right\} + \mathbb{P}\left\{\frac{d_R^2}{4k\sigma_b^2} < 1+\eta\right\}. \tag{29}$$

Since the first term vanishes exponentially fast when $k \to \infty$, we can focus on the second term. From (28), the second term in (29) is upper bounded by

$$\mathbb{P}\left\{\frac{\alpha_b^2}{4\sigma_b^2} \prod_{i=1}^k \left(\frac{1}{\rho_b} + |h_{b,i}|^2\right)^{\frac{1}{k}} < 1+\eta\right\} =$$

$$= \mathbb{P}\left\{\frac{1}{k}\sum_{i=1}^k \ln\left(1+\rho_b|h_{b,i}|^2\right) < \ln\left(\frac{4P(1+\eta)}{\alpha_b^2}\right)\right\} =$$

$$= \mathbb{P}\left\{\frac{1}{k}\sum_{i=1}^k \ln\left(1+\rho_b|h_{b,i}|^2\right) < \ln\left(\frac{2Ge^{R_b}(1+\eta)}{\pi e}\right)\right\},$$

recalling that $\alpha_b^2 \approx \frac{2\pi e P}{Ge^{R_b}}$ from (18) and the fact that $|\Lambda_b/\Lambda_e| = e^{kR}$. Since the left hand side tends to C_b when $k \to \infty$ due to the law of large numbers (14), the last expression will vanish provided that $R_b < C_b - \ln\left(\frac{2G}{\pi e}\right) - \ln(1+\eta)$. Since η is arbitrary, any rate

$$R_b = R + R' < C_b - \ln(2G/\pi e) \tag{30}$$

is achievable for Bob. From equations (27) and (30), the proposed coding scheme achieves strong secrecy for any message distribution (and thus semantic security) for any secrecy rate

$$R < C_b - C_e - \ln\left(2G^2/\pi\right).$$

This concludes the proof of Proposition 1.

ACKNOWLEDGEMENTS

Cong Ling's work was supported in part by FP7 project PHYLAWS (EU FP7-ICT 317562). The research of R. Vehkalahti was funded by Finnish Cultural Foundation.
The authors would like to thank Jean-Claude Belfiore and Hamed Mirghasemi for useful discussions.

REFERENCES

[1] J.-C. Belfiore, F. Oggier, "Lattice code design for the Rayleigh fading wiretap channel", *IEEE International Conference on Communications* (ICC) 2011

[2] J.-C. Belfiore, F. Oggier, "An error probability approach to MIMO wiretap channels", *IEEE Trans. Commun.*, vol. 61 n. 8, 2013

[3] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel", *Advances in Cryptology*, Lecture Notes in Computer Science, vol. 7417, Springer-Verlag, 2012, pp. 294–311.

[4] M. Bloch and J. Laneman, "Strong secrecy from channel resolvability", *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.

[5] I. Csiszár, "Almost independence and secrecy capacity", *Problems of Information Transmission*, vol. 32, pp. 40–47, 1996.

[6] H. El Gamal, G. Caire, M. O. Damen, "Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels", *IEEE Trans. Inform. Theory*, vol. 50, n. 6, pp. 968–985, 2004

[7] D. Hsu, S. M. Kakade, T. Zhang, "A tail inequality for quadratic forms of subgaussian random vectors", *Electron. Commun. Probab.* 17 (2012), no. 52, 1–6.

[8] D. Karpuk, A.-M. Ernvall-Hytönen, C. Hollanti, E. Viterbo, "Probability estimates for fading and wiretap channels from Ideal Class Zeta Functions", *Advances in Mathematics of Communication* vol. 9 n. 4, pp. 391–413, 2015

[9] S.-C. Lin, "On ergodic secrecy capacity of fast fading MIMOME wiretap channel with statistical CSIT", *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2013

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

[11] L. Luzzi, R. Vehkalahti, "Almost universal codes achieving ergodic MIMO capacity within a constant gap", arxiv.org/pdf/1507.07395

[12] V. Lyubashevsky, C. Peikert, O. Regev, "On ideal lattices and learning with errors", *Journal ACM*, vol. 60, n. 6, Nov. 2013

[13] J. Martinet, "Tours de corps de classes et estimations de discriminants", *Inventiones Mathematicae* n. 44, 1978, pp. 65–73.

[14] D. Micciancio and C. Peikert, "Trapdoors for lattices: Simpler, tighter, faster, smaller", *Advances in Cryptology - EUROCRYPT 2012*, Lecture Notes in Computer Science vol 7237, pp. 700–718

[15] D. Micciancio and O. Regev, "Worst-case to average-case reductions based on Gaussian measures", in *Proc. Ann. Symp. Found. Computer Science*, Rome, Italy, Oct. 2004, pp. 372–381.

[16] S. S. Ong, F. Oggier, "Wiretap lattice codes from number fields with no small norm elements", *Designs, Codes and Cryptography*, vol 73 n.2, pp. 425–440, 2014

[17] C. Peikert, "An efficient and parallel Gaussian sampler for lattices", *Proc. CRYPTO*, vol. 6223, Springer-Verlag, 2010, pp. 80–97.

[18] C. Peikert and A. Rosen, "Lattices that admit logarithmic worst-case to average-case connection factors", *Proc. STOC*, pp. 478–487, 2007.

[19] A. Subramanian, A. Thangaraj, M. Bloch, S. W. McLaughlin, "Strong Secrecy on the Binary Erasure Wiretap Channel Using Large-Girth LDPC Codes", *IEEE Trans. Inf. Forensic Secur.*, vol.6, no.3, pp. 585–594, 2011

[20] R. Vehkalahti and L. Luzzi, "Number field lattices achieve Gaussian and Rayleigh channel capacity within a constant gap", in *IEEE Int. Symp. Inform. Theory* (ISIT), June 2015

[21] A. D. Wyner, "The wire-tap channel", *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.