SECRET KEY GENERATION FROM MULTIPLE COOPERATIVE HELPERS BY RATE UNLIMITED PUBLIC COMMUNICATION

Bin Yang, Wenjie Wang, Qinye Yin

Ministry of Education Key Lab for
Intelligent Networks and Network Security
Xi'an Jiaotong University, 710049, P. R. China

ABSTRACT

This paper investigates the secret key generation from multiple cooperative sources. We assume that the public communication between the legitimate users has unlimited capacity. In this case, the secret key rate is the upper bound of a practical system with limited public communication. The random signals from the multiple helpers can help to create correlated Gaussian sources for the legitimate users to generate secret key, and only local channel state information (CSI) is needed by the helpers to cooperate with others. The power allocation between the helpers is studied with a total transmit power constraint, which results in a non-convex optimization problem. A high efficiency sub-optimal solution is proposed. And even when the legitimate users have no idea about the eavesdropper, the scheme still can help the system to achieve fairly good performance.

Index Terms— Secret key generation, cooperative helpers, physical layer security

1. INTRODUCTION

Secret key generation from common randomness was first studied by [1][2], which showed that correlated observations of random sources could be used to distill secret keys by discussing over a public channel. The result is extended to wireless communication system recently [3]-[10]. It is shown in [4][5][6] that there is trade-off between the secret key rate and the public communication rate of the key agreement protocol from random sources. In a practical system, it is difficult to find a correlated source between the legitimate users which is independent of the eavesdropper. The wireless channel state information (CSI) is one of the options [11]: due to wireless channel reciprocity, the legitimate users can observe a pair of dependent sources of randomness while keeping the eavesdropper ignorant of them, which is suitable for time-variance fading wireless channels.

This work is partially supported by National Natural Science Foundation of China (No.60772095 & No.60971113), and Foundation for Innovative Research Groups of National Natural Science Foundation of China (No.60921003).

In this paper, we propose a novel method to create correlated random sources for the partners to generate secret keys. This method does not depend on the time-variance fading channel reciprocity, even when the channel CSI changes slowly, high secret key rate can still be achievable. To simplify the analysis, we do not consider the effect from the public channel rate, that is, we assume that the public channel is unlimited, then the achievable secret key rate is the upper bound of a practical system.

The proposed scheme includes three stages. During the first stage, the multiple helpers synchronously send random symbols to the users independently, then user A can get the mixed signals as the random source. During the second stage, the helpers repeat the random symbols in the first stage with compensating factors to user B, these factors, which is designed from the CSI of channels between the helpers and the legitimate users, can help to create correlated signals at user B. During the third stage, the legitimate users exchange information over the public channel to agree with a common secret key from the symbols received in the first two stages. For example, in the first stage, user A gets signals as $\sum_{i=1}^{N} h_i x_i$ without considering the receiver's noise, and in the second stage, user B gets the signals as $\sum_{i=1}^{N} g_i w_i x_i$. If we set w_i as h_i/g_i, then the signals of user B are also $\sum_{i=1}^{N} h_i x_i$. Then the legitimate users can easily distill a same secret key of each other. At the same time, the eavesdropper can only get $\sum_{i=1}^{N} e_i x_i$ and $\sum_{i=1}^{N} e_i \frac{h_i}{g_i} x_i$, which are both different from the signals of the legitimate users. Then the eavesdropper can hardly get the secret key of the legitimate users. In the following sections we will show details of the scheme and also a sub-optimal solution of it. More important, even when the legitimate users have no idea of the eavesdropper, the scheme can also work, which is approved by theoretical analysis and simulation results.

2. SYSTEM MODEL

In this paper, we denote vector and matrix with bold font, the Hermitian operator by H, $*$ denotes conjugation of a complex number or vector, and we measure information in bits (i.e., $\log_2(\cdot)$ is used to calculate entropy).

The system model is shown in Fig.1. There are two le-

gitimate users A and B want to generate a common secret key. They can communicate with each other through a public channel. Several helpers attempt to help the two ones to achieve the goal. A passive eavesdropper lies at some where trying to steal the secret key, he can access the public channel and signals from the helpers. All the nodes in the system are equipped with single antenna.

There are three stages in the scheme. In the first stages, all the helpers send random symbols to partner A synchronously. Then the symbols partner A received are

$$x = \sum_{i=1}^{N} h_i s_i + n_1, \tag{1}$$

where $h_i,\ i = 1, \ldots, N$ denotes the complex channel gain between the helpers and user A, $s_i,\ i = 1, \ldots, N$ denotes complex zero-mean Gaussian random symbols sent by the helpers which is independent of each other, and n_1 is the noise of partner A. The sending power of the helpers is

$$P_i = E(|s_i|^2),\ i = 1, \ldots, N. \tag{2}$$

During the second stage, all the helpers repeat the symbols sent in the first stage multiplied with weight factors $w_i,\ i = 1, \ldots, N$. Then the signals user B received are

$$y = \sum_{i=1}^{N} g_i w_i s_i + n_2, \tag{3}$$

where $g_i,\ i = 1, \ldots, N$ denotes the complex channel gain between the helpers and user B, and n_2 is the noise of partner B. To create correlated sources for the partners, we set

$$w_i = \sqrt{\rho}\frac{h_i}{g_i},\ i = 1, \ldots, N, \tag{4}$$

where ρ is power factor to adjust the total transmit power in the second stage. Then the received symbols of user B are

$$y = \sqrt{\rho}\sum_{i=1}^{N} h_i s_i + n_2. \tag{5}$$

During these two stages, the eavesdropper gets the signals

$$v_1 = \sum_{i=1}^{N} e_i s_i + n_{E1},\ v_2 = \sqrt{\rho}\sum_{i=1}^{N} e_i \frac{h_i}{g_i} s_i + n_{E2}, \tag{6}$$

where $e_i,\ i = 1, \ldots, N$ denotes the complex channel gain between the helpers and the eavesdropper, n_{E1} and n_{E2} are the noise of the eavesdropper in the two stages respectively. We can see that v_1 and v_2 are different from x and y, which help the legitimate users get a prior position to the eavesdropper.

During the third stage, user A and user B use public communication to distill a common secret key. In this paper, we concentrate on the class of key agreement protocols in which only user A sends a message to user B over the public channel. User A computes the key S_n from x. And then user A computes the message C_n from x and sends the message to user B over the public channel. User B then computes the key S'_n from y and C_n. The error probability of the protocol is defined by

$$\varepsilon_n = \Pr\{S_n \neq S'_n\}. \tag{7}$$

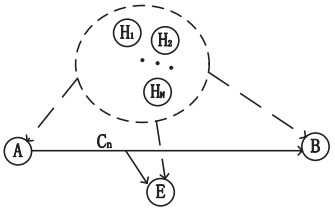

Fig. 1. System model.

The security of the protocol is measured by

$$\nu_n = \log|\mathcal{S}_n| - H(S_n|C_n, y_{E1}, y_{E2}), \tag{8}$$

where $\mathcal{S}_n$ is the range of the key S_n. Secret key rate R_k is achievable if there exists a sequence of protocols satisfying

$$\lim_{n\to\infty} \varepsilon_n = 0 \qquad \text{(Reliability Condition)}, \tag{9}$$

$$\lim_{n\to\infty} \nu_n = 0 \qquad \text{(Secrecy Condition)}. \tag{10}$$

It is shown in [5][6] that the secret key rate is limited by the public channel capacity for Gaussian sources. The upper bound can be achieved only when the communication rate is unlimited. In this paper, we are more concerned with the common randomness generation from multiple helpers than the key agreement protocol. We want to illustrate how to design the Gaussian signals from multiple helpers to help the system get higher secret key rate. Then we assume that the public channel capacity is infinite, then we can only consider the upper bound of secret key capacity.

Note that all the noise terms n_1, n_2, n_{E1} and n_{E2} are zero-mean white independent complex Gaussian random variables with the variance σ_n^2.

3. SECRET KEY RATE ANALYSIS

From the result in [2][5][6], the secret key rate from correlated Gaussian sources by unlimited public communication is

$$\begin{aligned} R_k &= I(x; y|\mathbf{v}) \\ &= H(x|\mathbf{v}) - H(x|y\mathbf{v}) \end{aligned} \tag{11}$$

where $I(\cdot;\cdot)$ is mutual information, and$\mathbf{v} \triangleq [v_1, v_2]^T$. Then we have

$$H(x|\mathbf{v}) = \log|Q_{x\mathbf{v}}| - \log|Q_{\mathbf{v}}|, \tag{12}$$

and

$$H(x|y\mathbf{v}) = \log|Q_{xy\mathbf{v}}| - \log|Q_{y\mathbf{v}}|, \tag{13}$$

where

$$\mathbf{Q}_{x\mathbf{v}} \triangleq E\left(\begin{bmatrix} x \\ \mathbf{v} \end{bmatrix}[x^*, \mathbf{v}^H]\right), \tag{14}$$

$$\mathbf{Q}_{\mathbf{v}} \triangleq E(\mathbf{v}\mathbf{v}^H), \tag{15}$$

$$\mathbf{Q}_{xy\mathbf{v}} \triangleq E\left(\begin{bmatrix} x \\ y \\ \mathbf{v} \end{bmatrix}[x^*, y^*, \mathbf{v}^H]\right), \tag{16}$$

$$\mathbf{Q}_{y\mathbf{v}} \triangleq E\left(\begin{bmatrix} y \\ \mathbf{v} \end{bmatrix}[y^*, \mathbf{v}^H]\right). \tag{17}$$

And then we have

$$|\mathbf{Q}_{x\mathbf{v}}| = \sum_{i>j>k} \alpha_{ijk} P_i P_j P_k + \tag{18}$$

$$\sigma_n^2 \left(\sum_{i>j} \beta_{ij} P_i P_j + \sigma_n^2 \left(\sum_i \gamma_i P_i + \sigma_n^2 \right) \right) \tag{19}$$

$$|\mathbf{Q}_{\mathbf{v}}| = \sum_{i>j} \mu_{ij} P_i P_j + \sigma_n^2 \left(\sum_{i=1}^{N} \eta_i P_i + \sigma_n^2 \right), \tag{20}$$

$$|\mathbf{Q}_{y\mathbf{v}}| = \rho \, |Q_{x\mathbf{v}}| + (1-\rho)\sigma_n^2 \, |Q_{\mathbf{v}}| \,, \tag{21}$$

$$|\mathbf{Q}_{xy\mathbf{v}}| = (1+1/\rho)\sigma_n^2 \, |Q_{y\mathbf{v}}| - \sigma_n^4/\rho \, |Q_{\mathbf{v}}| \,, \tag{22}$$

where

$$\alpha_{ijk} = \rho \left| e_j h_i h_k \left(\frac{e_i}{g_i} - \frac{e_k}{g_k} \right) - e_i h_j h_k \left(\frac{e_j}{g_j} - \frac{e_k}{g_k} \right) - e_k h_i h_j \left(\frac{e_i}{g_i} - \frac{e_j}{g_j} \right) \right|^2,$$

$$\beta_{ij} = |h_i e_j - h_j e_i|^2 + \rho \left| \frac{h_i h_j e_j}{g_j} - \frac{h_j h_i e_i}{g_i} \right|^2 + \rho \left| \frac{e_i h_j e_j}{g_j} - \frac{e_j h_i e_i}{g_i} \right|^2,$$

$$\gamma_i = |h_i|^2 + |e_i|^2 + \rho \left| \frac{e_i h_i}{g_i} \right|^2,$$

$$\mu_{ij} = \rho |e_i e_j|^2 \left| \frac{h_i}{g_i} - \frac{h_j}{g_j} \right|^2, \quad \eta_i = |e_i|^2 \left(1 + \rho \left| \frac{h_i}{g_i} \right|^2 \right),$$

where $Re\{\cdot\}$ denotes real part of a complex number.

When the total power of the helpers is limited to P_0, namely, $\sum_{i=1}^{N} P_i \le P_0$ and $\rho \sum_{i=1}^{N} \left| \frac{h_i}{g_i} \right|^2 P_i \le P_0$, which means

$$\rho \le \frac{P_0}{\sum_{i=1}^{N} \left| \frac{h_i}{g_i} \right|^2 P_i}. \tag{23}$$

Then the problem is

$$\begin{aligned} \max_{\mathbf{P}} \quad & R_k \\ s.t. \quad & \mathbf{P}^T \mathbf{1} \le P_0, \\ & \mathbf{P}^T |\mathbf{w}|^2 \le P_0, \\ & P_i \ge 0, i = 1 \dots N, \end{aligned} \tag{24}$$

where $\mathbf{P} = [P_1, P_2, \dots, P_N]^T$, $\mathbf{1} = [1, 1, \dots, 1]^T$ and $|\mathbf{w}|^2 = [|w_1|^2, |w_2|^2, \dots, |w_N|^2]^T$. The problem is not convex, then the solution is not tractable. Here we propose a sub-optimal solution of this problem.

Sub-Optimal Solution with High SNR

From (12) we know that

$$\lim_{P_0 \to \infty} H(x|\mathbf{v}) = \log(\mathbf{G}) \to \infty, \tag{25}$$

where

$$\mathbf{G}(\mathbf{P}) = \frac{\sum_{i>j>k} \alpha_{ijk} P_i P_j P_k}{\sum_{i>j} \mu_{ij} P_i P_j}, \tag{26}$$

and because $\lim_{P_0 \to \infty} |Q_{\mathbf{v}}|/|Q_{y\mathbf{v}}| \to 0$

$$\lim_{P_0 \to \infty} H(x|y\mathbf{v}) = \log\left((1+1/\rho)\sigma_n^2 \right) < \infty. \tag{27}$$

Normally when all the channels are fixed, ρ will be a limited value, unless some of the channels are deep fading, namely, there exists a index $i' \in \{1, 2, \dots, N\}$ that $|g_{i'}|^2 \to 0$, which means $H_{i'}$ almost cannot sends any signals to user B. In this case, we should keep the power of $H_{i'}$ to be zero, namely, $P_{i'} = 0$, which makes ρ still be a limited value. Then we have

$$\lim_{P_0 \to \infty} R_k = \lim_{P_0 \to \infty} H(x|\mathbf{v}) - \log\left((1+1/\rho)\sigma_n^2 \right) \to \infty. \tag{28}$$

This means when P_0 is large enough, the secret key rate can also be very large. In a practical system, we hope the system can achieve high secret key rate. So we are more interested in the high SNR area. In this area, R_k is mainly determined by the value of $\log(\mathbf{G})$. Then the optimization of R_k can be approximately solved by the optimization of $\mathbf{G}$ when the total transmit power P_0 is large enough. Then the problem can be converted to a equation constraint optimization problem

$$\begin{aligned} \max_{\mathbf{P}} \quad & \mathbf{G}(\mathbf{P}) \\ s.t. \quad & \mathbf{P}^T \mathbf{1} = P_0, \\ & P_i \ge 0, i = 1 \dots N. \end{aligned} \tag{29}$$

Unfortunately, $\mathbf{G}$ is still not convex function, it is difficult to get the optimal solution. As an option, if we can find a locally maximum value instead of the global optimal value, it will be a fairly good solution. So we can use the Newton's method with equality constraints to solve the problem (29) (for more detail, please refer to [13]). The Newton step $\Delta\mathbf{P}$ is characterized by

$$\begin{bmatrix} \nabla^2 \mathbf{G} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{P} \\ v \end{bmatrix} = \begin{bmatrix} \nabla \mathbf{G} \\ 0 \end{bmatrix}, \tag{30}$$

where v is the associated optimal dual variable for the optimization problem. And we define the Newton decrement for the problem as

$$\lambda(\mathbf{P}) = \left(\Delta\mathbf{P} \nabla^2 \mathbf{G} \Delta\mathbf{P}^T \right)^{1/2}. \tag{31}$$

Then we can get a sub-optimal solution as following:

1. random generate N_t power allocation vectors: $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_{N_t}$, and find out the highest secret key rate point with the power vector $\mathbf{P_m}$ as the start point of iteration, that is $R_k(\mathbf{P_m}) = \max(R_k(\mathbf{P}_1), R_k(\mathbf{P}_2), \dots, R_k(\mathbf{P}_{N_t}))$.

2. compute the Newton step and decrement $\Delta\mathbf{P}$, $\lambda(\mathbf{P})$ from (30) and (31).

3. choose step size t by backtracking line search.

4. update $\mathbf{P}$: $\mathbf{P} = \mathbf{P} + t\Delta\mathbf{P}$.

5. iterate to step **2** with stopping criterion: $\lambda^2/2 \le \varepsilon$ or $|\mathbf{P}|^T \mathbf{1} > P_0$.

6. get sub-optimal vector $\mathbf{P}^*$, then the sub-optimal secret key rate is $R_k^* = \max(R_k(\mathbf{P_m}), R_k(\mathbf{P}^*))$.

Without CSI of the Eavesdropper

In a practical system, usually we have no idea about the eavesdropper, then we set the sending power of the helpers as the same

$$P_i = P_0/N, \quad i = 1 \dots N. \tag{32}$$

Then

$$\rho = \frac{N}{\sum_{i=1}^{N} \left| \frac{h_i}{g_i} \right|^2}. \tag{33}$$

When P_0 goes to infinite

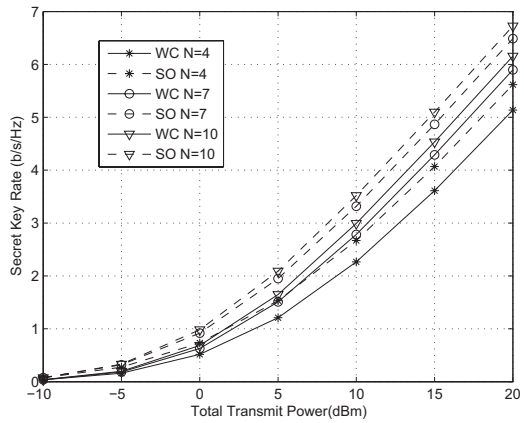

Fig. 2. Average secret key rate versus transmit power for Rayleigh fading channels. SO denotes the sub-optimal solution, and WC denotes the algorithm without CSI of the eavesdropper.

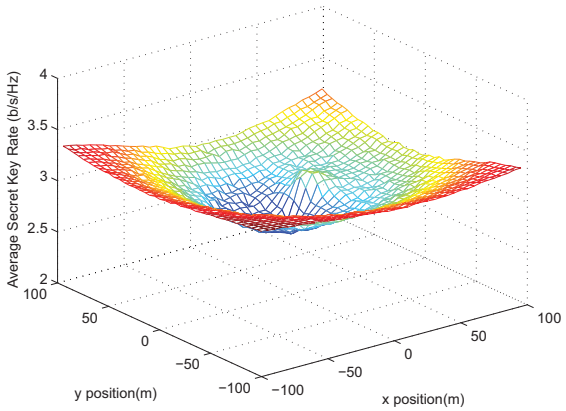

Fig. 3. Average secret key rate when the eavesdropper moving in a two-dimension plane with the locations of helpers and users being fixed.

$$\lim_{P_0\to\infty} R_k = \log\left(\frac{P_0\sum\limits_{i>j>k}\alpha_{ijk}}{N\sum\limits_{i>j}\beta_{ij}}\right) - \log\left(1+\frac{\sigma_n^2}{N}\sum_{i=1}^{N}\left|\frac{h_i}{g_i}\right|^2\right). \quad (34)$$

Actually, $\sum \alpha_{ijk}$ is impossible to be zero in a practical system. This means in high SNR area, R_k increases linearly with the total power P_0. Such a simple algorithm can help to get infinite secret key rate by increasing the total signal power without considering the CSI of the eavesdropper. This is the key benefit of our scheme.

4. SIMULATION RESULTS

In this section, we demostrate the performance of the proposed schemes numerically. We perform the simulations with two channel models: Rayleigh fading channel and line-of-sight (LOS) channel.

Firstly, we consider wireless communication system in fading environment. The channels of the users and the eavesdropper are all Rayleigh fading channels which are independent of each other, and follow the unit variance complex Gaussian distribution. The noise power σ_n^2 is set as 0dBm, and $N_t = 1000$. We compare the results with different number of the helpers and different algorithms, 1000 times of experiments are performed to get an average secret key rate, which are shown in Fig. 2. We can see that the secret key rates increase linearly with the total transmit power, and more helpers result in better performance. Even when we have no idea about the eavesdropper, the performance of the algorithm is fairly good.

Secondly we do the simulation in LOS channel model. For simplicity, the source and destination are placed along a horizontal line. The source and the destination are placed at (0, 0) and (50, 0) respectively. There are four helpers placed near them with the positions being (15, 25), (25,30), (27, 10) and (20, 20) respectively. The eavesdropper moves on the plane with x axis and y axis from -100m to 100m respectively, the secret key rate of the system is calculated. Here we only perform the algorithm without CSI of the eavesdropper, which have worse performance than the sub-optimal solution shown in Fig.2. To highlight the effects of distances, channels between any two nodes are modeled by a simple line-of-sight channel model including the path loss effect and a random phase. For example, $h = d^{-c/2}e^{j\theta}$ where d is the distance between two nodes, $c = 3.5$ is the path loss exponent, θ is the random phase uniformly distributed within $[0,\ 2\pi)$. The noise power is $\sigma_n^2 = -60$dBm, and P_0 is 10dBm. We perform Monte Carlo experiments consisting of 1000 independent trials to obtain the average results. It is shown in Fig.3 that the achievable secret key rate is between 2.5bits/s/Hz to 3.5bits/s/Hz when the eavesdropper moves in both dimensions and both directions. This means wherever the eavesdropper is located certain system performance can be guaranteed by the proposed scheme.

5. CONCLUSION

In this paper, we have investigated the design of the correlated Gaussian sources with multiple cooperative helpers for secret key generation. The benefits of the proposed scheme are: 1. there is no global channel state information (CSI) to be shared between the helpers, only local CSI is enough for the helpers to perform the algorithm; 2. the secret key rate can be arbitrary high by increasing the sending power of the helpers; 3. the algorithm can achieve fairly good performance even without any information on the eavesdropper.

6. REFERENCES

[1] U.M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory,* vol. 39, no. 3, pp. 733-742, 1993.

[2] R. Ahlswede, I. Csiszár, "Common randomness in information theory and cryptography part I: Secret sharing," *IEEE Trans. Inf. Theory,* vol. 39, no. 4,pp. 1121-1132, 1993.

[3] M. Bloch, J. Barros, M.R.D. Rodrigues, S.W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory,* vol. 54, no. 6, pp. 2515-2534, Jun. 2008.

[4] T. Chou, S. Draper, A. Sayeed, "Key generation using external source excitation: Capacity, reliability, and secrecy exponent," *IEEE Trans. Inf. Theory,* vol. 58, no. 4, pp. 2455-2474, Apr. 2012.

[5] S. Watanabe, Y. Oohama, "Secret Key Agreement From Correlated Gaussian sources by rate limited public communication," *IEICE Trans. Fundamentals*, vol. E93A, no. 11, pp. 1976-1983, Nov. 2010.

[6] S. Watanabe, Y. Oohama, "Secret Key Agreement From Vector Gaussian Sources by Rate Limited Public Communication," *IEEE Trans. Inf. Forensics and Security*, vol. 6, no. 3, pp. 541-550, Sep. 2011.

[7] S. Nitinawarat, P. Narayan, "Secret key generation for correlated gaussian sources," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3373-3391, June 2012.

[8] C. Ling, L. Luzzi, M.R. Bloch, "Secret key generation from gaussian sources using lattice hashing," *ISIT, 2013 IEEE International Symposium on*, Istanbul, Turkey, pp. 2621-2625, 7-12 July 2013.

[9] M. Zafer, "Limitations of generating a secret key using wireless fading under active adversary," *Networking, IEEE/ACM Trans.,* vol. 20, no. 5, pp. 1440-1451, Oct. 2012.

[10] A. Khisti, S.N. Diggavi, G.W. Wornell, "Secret-key generation using correlated sources and channels," *IEEE Trans. Inf. Theory,* vol. 58, no. 2, pp. 652-670, Feb. 2012.

[11] A. Agrawal, Z. Rezki, A.J. Khisiti, M-S Alouini, "Noncoherent capacity of secret-key agreement with public discussion," *IEEE Trans. Inf. Forensics and Security,* vol. 6, no. 3, pp. 565-574, Sep. 2011.

[12] B. Yang, W. Wang, B. Yao, Q. Yin, "Destination assisted secret wireless communication with cooperative helpers," *IEEE Sig. Proc. Lett.,* vol. 20, no. 11, pp. 1030-1033, Nov. 2013.

[13] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, no. 6, pp. 585-596, Nov. 1999.

[14] S. Boyd, L. Vandenberghe, "Convex Optimization," *Combridge University Press*, 2004.