Achievable Secrecy Rates for the Broadcast Channel with Confidential Message and Finite Constellation Inputs

Zeina Mheich, Florence Alberge, Pierre Duhamel

Achievable Secrecy Rates for the Broadcast Channel with Confidential Message and Finite Constellation Inputs

Z. Mheich, F. Alberge and P. Duhamel
Univ. Paris-Sud, UMR8506 Orsay, F-91405; CNRS, Gif-sur-Yvette, F-91192;
Supélec, Gif-sur-Yvette, F-91192, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette cedex, France
Tel: +33 1 69851757; fax: +33 1 69851765
e-mail:{alberge, zeina.mheich, pierre.duhamel}@lss.supelec.fr

***Abstract*—This paper considers the Broadcast Channel with Confidential Message (BCCM) where the sender attempts to send altogether a common message to two receivers and a confidential message to one of them. The achievable rate regions are derived for the power-constrained Gaussian BCCM with finite input alphabet using various transmission strategies. Namely, time sharing, superposition modulation and superposition coding are used as broadcast strategies. For superposition modulation and superposition coding, the maximal achievable rate regions are obtained by maximizing over both constellation symbol positions and the joint probability distribution. The maximization of the secrecy rate for wiretap channels is also studied as a particular case of the BCCM problem. We compare the considered transmission strategies in terms of percentage gains in achievable rates. We concentrate on the impact of the finite alphabet constraint on achievable rates, and show that this constraint may change well known results obtained in the Gaussian case. We show also that the secrecy constraint can change the shape of the achievable rate region in superposition modulation used in some standards when symbols are equiprobable. On a more practical side, it is shown that a performance close to the optimum can be obtained by strategies with reduced complexity.**

***Index Terms*—Information-theoretic security, finite-alphabet input, broadcast channel with confidential message, achievable rate region.**

I. Introduction

Security is an important issue for wireless communications. Vulnerability to eavesdropping comes from the shared nature of the wireless environment. Traditionally, cryptographic techniques are used at higher layers of the protocol stack for security purpose. In these techniques, security relies on the assumption of limited computational power at the eavesdropper. Recently, the wireless communications community has devoted a considerable attention to the information theoretic security at the physical layer, which makes use of totally different concepts. Indeed, in physical layer security, secrecy is achieved by exploiting the randomness of the wireless channels and does not assume any computational restrictions at the eavesdropper.

In the wiretap channel model, introduced by Wyner in [1], a transmitter wants to send reliably confidential message to a legitimate receiver and to keep the transmitted message secure from an eavesdropper. The level of ignorance at the wiretapper with respect to the confidential message is measured by the equivocation rate. Wyner demonstrated that secure communication is possible without sharing a secret key and determined the secrecy capacity of the memoryless degraded wiretap channel. The secrecy capacity is the maximal achievable rate to communicate reliably with the destination while the wiretapper is not able to obtain any information from the incoming signal. The secrecy capacity for the Gaussian wiretap channel was given later in [2]. Csiszar and Korner studied in [3] a more general model of the wiretap channel called broadcast channel with confidential message where the channels do not obey necessarily any degradation relationship. In this model, there is a common message for two receivers in addition to the confidential message for one receiver. More recently, fading was also introduced in the secret transmission model [4], [5] and the Gaussian multiple-input–multiple-output (MIMO) and multiple-input–single-output (MISO) wiretap channel are revisited in [6] and [7] respectively.

The secrecy capacity for the Gaussian wiretap channel and the secrecy-capacity region for the Gaussian BCCM are achieved using random Gaussian codebooks. However, Gaussian alphabets are not used in real systems since they are not practical to implement, and instead finite constellations such as Pulse Amplitude Modulation (PAM) or Quadrature Amplitude Modulation (QAM) are considered, usually with equal probability. The impact of finite size constellations on the achievable secrecy rate is analyzed in [8] and [9] in the particular case of equiprobable symbols. It is shown that the secrecy rate curves for a finite constellation plotted against the Signal-to-Noise-Ratio (SNR) and for a fixed noise variance of the eavesdropper's channel have a global maximum at an internal point. This comes in contrast to what is known in the case of Gaussian codebook input where the secrecy capacity curve is a bounded, monotonically increasing function of SNR. Ref. [10] investigates the secrecy rate of the Gaussian wiretap channel with standard M-PAM inputs. The authors provide the necessary conditions for both the M-PAM input power and the M-PAM input distribution to maximize the secrecy rate which they specialize to the asymptotic low-power and high-power regimes. Ref. [11] and [12] study the effect of finite discrete-constellation on the secrecy achievable rate of

multiple-antenna wiretap channels and [13] investigates the power allocation and artificial noise design for orthogonal frequency-division multiplexing (OFDM) wiretap channels with discrete channel inputs. In [14], the authors investigated the design of optimum linear transmit precoding for the maximum secrecy rate over multiple-input–multiple-output–multiple-antenna eavesdropper (MIMOME) wiretap channels. The authors develop an iterative algorithm for secrecy rate maximization via a gradient method which achieves substantial rate gains over the precoding design in [11].

This paper studies the achievable rates for the Gaussian broadcast channel with confidential message using M-PAM constellations. Unlike prior works, where the secrecy rate for the wiretap channel is studied massively assuming uniform input distribution or/and standard symbol positions, we investigate in this work the maximal achievable rate region for Gaussian BCCM, by optimizing over both symbol positions and the joint probability distribution, subject to the availability of a suitable initial guess. To our knowledge, no work investigated the maximization of achievable rate regions of the BCCM under finite alphabet constraint. The symbol positions in our work are allowed to take arbitrary values and are not necessarily proportional to those of standard constellation as in [10]. This leads to the determination of the maximal achievable rates with any constellation of M symbols. The achievable rate regions are also given for various broadcast transmission strategies which differ in their complexity of implementation. Preliminary and partial results were published in [15] by the same authors. The whole picture is given here. Additional contributions of this paper compared to [15] are specified hereafter. Regarding the achievable rate regions for the BCCM, comparisons between the various strategies are conducted, in this paper, in terms of SNR savings for target achievable rates and percentage of gain in achievable rates. The corresponding trade-off between complexity and efficiency is discussed. The goal is to know whether using practical schemes is sufficient to achieve good rates or it leads to significant losses. This contribution is a first step towards a practical implementation of secure communication at the physical layer.

II. Achievable rates for the BCCM

This section recalls classical results on the achievable rates of the BCCM [3], i.e. a broadcast channel with two receivers for which a sender attempts to send two messages simultaneously: a common message w_0 to both receivers and a secret message w_1 for receiver 1. A discrete-memoryless BCCM (DM-BCCM) consists of an input alphabet $\mathcal{X}$, two output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ and transition probabilities $P_{Y_1Y_2|X}$ such that $P_{Y_1^nY_2^n|X^n}(y_1^n, y_2^n|x^n) = \prod_{i=1}^n P_{Y_1Y_2|X}(y_{1i}, y_{2i}|x_i)$ (Figure 1). Conventionally, random variables (RV) are written in upper case letters and particular realizations are written in corresponding lower case letters.

A $(2^{nR_0}, 2^{nR_1}, n)$ code for the DM-BCCM consists of the following elements.

- Two message sets $\mathcal{W}_0 = \{1, .., 2^{nR_0}\}$ and $\mathcal{W}_1 = \{1, .., 2^{nR_1}\}$. We assume throughout that the messages W_0 and W_1 are uniformly distributed over the message sets $\mathcal{W}_0$ and $\mathcal{W}_1$ respectively.

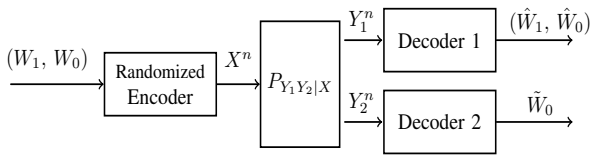

Figure 1. The broadcast channel with confidential message

- A randomized encoder that maps a message pair $(w_0, w_1) \in (\mathcal{W}_0, \mathcal{W}_1)$ to a codeword x^n.
- Two decoders: Decoder 1 maps a received sequence $y_1^n \in \mathcal{Y}_1^n$ to a message pair $(\hat{w}_0, \hat{w}_1)$ or an error message e, the second one at receiver 2 maps a received sequence $y_2^n \in \mathcal{Y}_2^n$ to a message $\tilde{w}_0$ or an error message e.

The secrecy level of W_1 at the eavesdropper is measured by the *equivocation rate*. The average error probability is $P_e^{(n)}$ with expression given below

$$\frac{1}{2^{nR_0}2^{nR_1}} \cdot \sum_{w_0=1}^{nR_0} \sum_{w_1=1}^{nR_1} \Pr\{(\hat{w}_0, \hat{w}_1) \neq (w_0, w_1) \text{ or } \tilde{w}_0 \neq w_0\}$$

The rate-equivocation triple (R_0, R_1, R_e) is achievable if there is a sequence of $(2^{nR_0}, 2^{nR_1}, n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$ and with equivocation rate satisfying $R_e \leq \liminf_{n\to\infty} \frac{1}{n} H(W_1|Y_2^n)$.

Throughout this work, we focus on the case in which *perfect secrecy* is achieved ($R_1 = R_e$) , i.e. the confidential messages transmitted are entirely hidden to the eavesdropper. The secrecy-capacity region is the set of all rate pairs (R_0, R_1) such that $(R_0, R_1, R_e = R_1)$ is achievable. The secrecy-capacity region which has been provided in [3] is the closure of the set that includes all (R_0, R_1) such that:

$$0 \leq R_1 \leq I(V; Y_1|U) - I(V; Y_2|U) \quad (1)$$

$$R_0 \leq \min\{I(U; Y_1), I(U; Y_2)\} \quad (2)$$

for some $P_{UVXY_1Y_2}$, and where U and V are auxiliary random variables satisfying $U \leftrightarrow V \leftrightarrow X \leftrightarrow Y_1Y_2$. U serves as a cloud center distinguishable by both receivers. In other terms, it carries the common information. V is an auxiliary random variable for additional randomization at the encoder side. The cardinality of the set $\mathcal{U}$ can be limited to $|\mathcal{U}| \leq |\mathcal{X}| + 3$.

The channel to receiver 2 is called a physically degraded version of the channel to receiver 1 if $p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$ i.e. $X \leftrightarrow Y_1 \leftrightarrow Y_2$ is a Markov chain. In this case, it is shown in [4] that $I(V; Y_1|U) - I(V; Y_2|U) \leq I(X; Y_1|U) - I(X; Y_2|U)$. Moreover, we have $I(U; Y_1) \geq I(U; Y_2)$ due to the Markov chain condition $U \leftrightarrow V \leftrightarrow X \leftrightarrow Y_1 \leftrightarrow Y_2$. Thus, the achievable rates in (1) and (2) satisfy for the degraded BCCM $U \leftrightarrow V \leftrightarrow X \leftrightarrow Y_1 \leftrightarrow Y_2$ [4]:

$$R_1 \leq I(X; Y_1|U) - I(X; Y_2|U) \quad (3)$$

$$R_0 \leq I(U; Y_2) \quad (4)$$

where $V = X$ in this case. It can be shown that the secrecy-capacity region depends only on the conditional marginals. Hence, this result generalizes to stochastically degraded DM-BCCM. In the case of degraded BCCM, the cardinality of $\mathcal{U}$ can be limited to $|\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\}$ which follows

from Caratheodory's theorem [16]. Comparing to the capacity region of the degraded broadcast channel without confidential messages [17], it may seem that the secrecy constraint leads to sacrifice a significant portion of the available capacity to confuse the eavesdropper. However, this is misleading because it is possible to send an additional private message to the legitimate receiver in addition to the confidential message to achieve the capacity region of the degraded broadcast channel without confidential message. In this paper, we will focus on the study of the secrecy rate for the legitimate receiver and the common message rate only.

Throughout this work, we consider the (degraded) Gaussian BCCM channel. The channel outputs are $Y_i = X + Z_i$, where $i \in \{1, 2\}$, $Z_i \sim \mathcal{N}(0, N_i)$ and $N_2 > N_1$. We consider also an input power constraint $\mathbf{E}[X^2] \leq P$. In [4], the secrecy-capacity region of the Gaussian BCCM with input power constraint P is given as:

$$= \bigcup_{\beta \in [0,1]} \left\{ (R_0, R_1) : \begin{array}{l} R_0 \leq C(\frac{(1-\beta)\cdot P}{N_2 + \beta \cdot P}) \\ R_1 \leq C(\frac{\beta \cdot P}{N_1}) - C(\frac{\beta \cdot P}{N_2}) \end{array} \right. \quad (5)$$

where $C(x) = \frac{1}{2} \cdot \log_2(1 + x)$. The achievability of the secrecy-capacity region follows from the previous definition of achievable rates for degraded BCCM with the following choice of random variables: $U \sim \mathcal{N}(0, (1-\beta) \cdot P)$, $X = U + X'$ with $X' \sim \mathcal{N}(0, \beta \cdot P)$.

The wiretap channel is a special case of BCCM where U is a constant, and R_0 is equal to zero. The secrecy capacity of the discrete memoryless WTC is obtained by taking $U = const$ in the BCCM case [3].

III. Broadcast Transmission Strategies

The common rate R_0 in (4) and the secrecy rate R_1 in (3) are achieved using superposition coding (SC) to transmit simultaneously both messages. Stochastic encoding [3], [18] is used to ensure security. This paper considers various broadcast strategies which differ in their complexity of implementation and performance. A detailed description of these strategies can be found in [19], [20]. They are listed below in ascending order of complexity and performance:

- Time Sharing (TS). Messages w_0 and w_1 are transmitted in different time-slots. Here, transmitted symbols belong to a standard M-PAM constellation ($\mathcal{X} = \{M - 1 - 2 \cdot (i-1) \text{ for } i = 1, .., M\}$).
- Superposition Modulation (SM) M symbols are obtained by adding two random variables X_1 and X_2 of respective cardinality M_1 and M_2, i.e. $M = M_1 M_2$. This corresponds to a separable labeling. Two situations are considered: (i) equiprobable symbols and optimized symbol positions, denoted as $SM_{\mathcal{X}, \overline{P_{UX}}, \overline{P_X}}$, which allows a separate encoding for both common and secret information and (ii) full optimization of symbol positions and joint probability distribution P_{UX}, a scheme denoted as $SM_{\mathcal{X}, P_{UX}, P_X}$.
- Superposition Coding (SC). P_{UX} takes the most general form, i.e. U has the largest cardinality: $|\mathcal{U}| = |\mathcal{X}|$ for Gaussian BCCM. The auxiliary variable U serves as a cloud center for the information. Thus labeling does not allow to distinguish between the common and the secret information. The encoding of both messages is done jointly using the joint distribution of probability P_{UX} and the decoding is based on large block typicality [17]. Superposition coding is here assumed to correspond to the optimization of the symbol position and P_{UX}. This scheme is denoted as $SC_{\mathcal{X}, P_{UX}, P_X}$.

IV. Achievable Rates With M-PAM

This section shows how to compute the maximal achievable rate region (i.e. R_0 as a function of R_1) of two-user power-constrained (degraded) Gaussian BCCM when the transmitted signal is modulated using an M-PAM constellation. This work is easily extended to complex Gaussian channel models using M-PSK and M-QAM constellations.

A. Problem Formulation

Consider a Gaussian BCCM in which the transmitter attempts to send a common message to two receivers (1 and 2) and a confidential message to receiver 1 at rates R_0 and R_1 respectively. The channel additive white Gaussian noise (AWGN) of receiver $k \in \{1, 2\}$ follows a normal distribution of zero mean and variance N_k. The channel input X is subject to a practical average power constraint $\mathbb{E}[X^2] \leq P$. The input alphabet $\mathcal{X}$ consists of M real valued symbols : $|\mathcal{X}| = M$. We study the case where the receiver SNRs verifies $SNR_2 < SNR_1$, with $SNR_k = \frac{P}{N_k}$: the output at receiver 2 is a degraded version of the output at receiver 1. The optimal rates R_0 and R_1 for some broadcast strategy satisfy the right hand side inequalities of (3) and (4). Thus the achievable rates in our system model can be computed for some $\theta \in [0, 1]$, by solving the following weighted sum rate maximization problem:

$$\max_{P_{UX}, \mathcal{X}} \quad \theta \cdot \Big[I(X; Y_1|U) - I(X; Y_2|U) \Big] + (1 - \theta) \cdot I(U; Y_2)$$
$$s.t. \quad \left\{ \begin{array}{l} p_{ij} \geq 0 \quad \forall (i,j) \in \mathcal{I} \times \mathcal{J} \\ \sum_{ij} p_{ij} \cdot x_j^2 \leq P \\ \sum_{ij} p_{ij} = 1 \end{array} \right. \quad (6)$$

where $p_{ij} = \Pr\{U = u_i, X = x_j\}$, $j \in \mathcal{J} = \{0, .., M-1\}$ and $i \in \mathcal{I} = \{0, .., |\mathcal{U}| - 1\}$. When P_X is constrained to be uniform, the last constraint is replaced by $\sum_i p_{ij} = \frac{1}{M}$. $I(X; Y_k|U)$ where $k \in \{1, 2\}$, and $I(U; Y_2)$ can be written for the Gaussian channel with finite input alphabet case as [1]

$$I(X; Y_k|U) = \sum_{i,j} \int_{-\infty}^{+\infty} p_{ij} P_{Y_k|X}(y_k|x_j) \cdot \log \frac{(\sum_{j'} p_{ij'}) P_{Y_k|X}(y_k|x_j)}{\sum_{j'} p_{ij'} P_{Y_k|X}(y_k|x_{j'})} dy_k \quad (7)$$

$$I(U; Y_2) = \sum_i \int_{-\infty}^{+\infty} (\sum_j p_{ij} P_{Y_2|X}(y_2|x_j)) \cdot \log \frac{\sum_{j'} p_{ij'} P_{Y_2|X}(y_2|x_{j'})}{(\sum_{j'} p_{ij'})(\sum_{i',j'} p_{i'j'} P_{Y_2|X}(y_2|x_{j'}))} dy_2 \quad (8)$$

[1] All logarithms are taken base 2.

Here also, one can note that the optimization for the wiretap channel is equivalent to the one for BCCM with $\theta = 1$ and constant U (6).

Clearly, the non-concave problem (6) can hardly be solved using exhaustive search especially when M increases. An iterative method is proposed below.

B. Numerical Solution

In order to solve the problem (6), we use an alternative maximization of the Lagrangian with respect to $\mathcal{X}$ and P_{UX}. A similar method was proposed in [19] for the broadcast channel without secrecy constraint. The Lagrangian L of problem (6) can be written as:

$$L(P_{UX},\mathcal{X},s) = \theta \cdot \Big[I(X;Y_1|U) - I(X;Y_2|U)\Big] + (1-\theta)\cdot I(U;Y_2) + s\cdot\Big(P - \sum_{ij} p_{ij}\cdot x_j^2\Big) \quad (9)$$

For a given value of s, the maximization of L with respect to P_{UX} and to $\mathcal{X}$ is done iteratively until convergence:

$$P_{UX}^{(\ell)} = \arg\max_{P_{UX}\in\mathcal{C}} L(P_{UX},\mathcal{X}^{(\ell-1)},s) \quad (10)$$

$$\mathcal{X}^{(\ell)} = \arg\max_{\mathcal{X}} L(P_{UX}^{(\ell)},\mathcal{X},s) \quad (11)$$

where ℓ is the iteration index and $\mathcal{C}$ denotes the set of constraints on P_{UX} and can be defined either as $\mathcal{C} = \{P_{UX} : p_{ij} \geq 0, \sum_{i,j} p_{i,j} = 1\}$ or as $\mathcal{C} = \{P_{UX} : p_{ij} \geq 0, \sum_i p_{i,j} = \frac{1}{M}\}$ if symbols are used with equal probability. It is observed in [15] that $L(P_{UX}^{(\ell)},\mathcal{X},s)$ is a concave function in $\mathcal{D}$ where $\mathcal{D}$ is the set of input alphabets with a minimum spacing between symbols greater than d and d is a function of the SNR and of the constellation size [15]. This condition was observed in experiments for most values of studied SNR (except when the value of s is very high such that $s\cdot\Big(P - \sum_{ij} p_{ij}\cdot x_j^2\Big)$ becomes the dominant part in L, then L is concave in $\mathcal{X}$ and the optimal $\mathcal{X}$ contains null symbols.). A simplex method is then used to solve (11) on $\mathcal{D}$. Fig. 2 shows an example of the contour of L as a function of symbol positions x_0 and x_1 of a 4-PAM constellation ($\mathcal{X} = \{x_0, x_1, -x_1, -x_0\}$) for fixed s and P_{UX}. We observe that the Lagrangian has one global maximum and one local maximum which are located in the regions $x_0 > x_1$ and $x_0 < x_1$. The Lagrangian also has multiple minima in the region where x_0 is close to x_1. From a practical point of view, this corresponds to a case where the modulation has a tendency to degenerate to a smaller size, therefore denoting a poor match between the SNR and the constellation size. Thus, in simulations, we make multiple initializations for x_0 and x_1 in the case of 4-PAM (in the regions $x_0 > x_1$ and $x_0 < x_1$). In the same manner, for other constellation cardinalities, we can identify the regions where the function is concave. Then, we run the simplex algorithm for multiple initializations and choose the result that achieves the maximum value of the Lagrangian.

Now, we turn to the optimization problem in (10) which is used when P_{UX} is not constrained to be uniform. In the literature, there exists a Blahut-Arimoto type algorithm which enables to maximize the secrecy rate $R_1 = I(X;Y_1) - I(X;Y_2)$ for the case of wiretap channel in which the eavesdropper's channel is noisier than the main channel. This algorithm proposed in [21] is guaranteed to converge to a global maximum since the function $I(X;Y_1) - I(X;Y_2)$ is concave in P_X for a fixed $\mathcal{X}$ in this case [22]. For the general case of Gaussian BCCM, the following lemma focuses on the (non)-concavity of (10) when $0 \leq \theta < 1$.

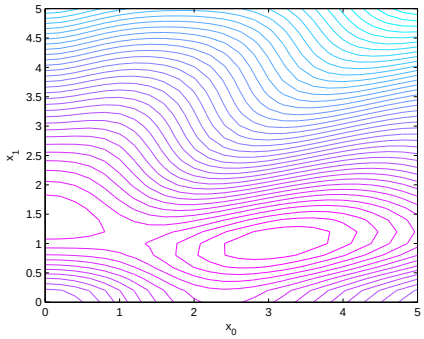

Figure 2. Contour of Lagrangian L. $SNR_1 = 10$ dB, $SNR_2 = 4$ dB, s=0.03, P_{UX} arbitrarily chosen, $\theta = 0.7$. The maximum corresponds to $x_0 = 3$ and $x_1 = 1$.

Lemma 1: (i) The receiver 1's channel $X \rightarrow Y_1$ is less noisy than the receiver 2's channel $X \rightarrow Y_2$ if and only if $I(X;Y_1|U) - I(X;Y_2|U)$ is a concave function of P_{UX}, (ii) $I(U;Y_2)$ is a difference of concave functions of P_{UX}.

Proof: (i) is proven in the Appendix and (ii) is demonstrated in [23, Appendix A]. ■

Thus (10) is a non-concave optimization problem but it is similar to the non-concave problem without secrecy constraint considered in [19]. From the expressions of the mutual information $I(X;Y_k|U)$ and $I(U;Y_2)$, where $k \in \{1,2\}$, we have also the following lemma.

Lemma 2: Consider the case of superposition coding where the alphabet of the transmitted signal is not a sum of two alphabets for the common and the secret information respectively. In this case, if $P_{UX}^{*(l)}(s)$ is a solution of problem (10), then any joint probability distribution P_{UX} obtained by permuting the rows of $P_{UX}^{*(l)}(s)$ is also a solution of problem (10).

Proof: Lemma 2 comes from (7), (8) and the constraints in (6) in which permuting the rows of the joint distribution of probability does not change the function value in (10). Hence, problem (10) has multiple solutions. However Lemma 2 does not hold for superposition modulation, since in this scheme $\mathcal{U}$ represents the alphabet of the common information. Thus the constellation symbol positions in $\mathcal{X}$ will depend on the values of $\mathcal{U}$, i.e. $X = U + X_1$ where X_1 represents the signal carrying the secret information. Consequently, permuting the rows of P_{UX} will change the mutual information values in (7), (8) for superposition modulation strategy. ■

Therefore, obviously, in some of the considered situations, the problem of interest has multiple solutions, and the uniqueness of a global maximum cannot hold. This is indicated below.

In order to solve the optimization problem in (10) with constraint set $\mathcal{C}$ we used a Blahut-Arimoto type algorithm which can be done for the Gaussian BCCM using the same method in [24] for the degraded broadcast channel without

Step 0	$s \leftarrow s^{(0)}$	
Step k	Step 0	$\mathcal{X} \leftarrow \mathcal{X}^{(0)}$ where $\mathcal{X} = (x_0, x_1, .., x_{M-1})$
	Step ℓ	$P_{UX}^{(\ell)} = \arg\max_{P_{UX} \in \mathcal{C}} L(P_{UX}, \mathcal{X}^{(\ell-1)}, s^{(k-1)})$ $\mathcal{X}^{(\ell)} = \arg\max_{\mathcal{X}} L(P_{UX}^{(\ell)}, \mathcal{X}, s^{(k-1)})$
	Stopping criterion	$\lvert L(P_{UX}^{(\ell)}, \mathcal{X}^{(\ell)}, s^{(k)}) - L(P_{UX}^{(\ell-1)}, \mathcal{X}^{(\ell-1)}, s^{(k-1)})\rvert \leq \epsilon_L$
	$s^{(k)} = \Big[s^{(k-1)} - \beta\Big(P - \sum_{i,j} p_{ij}^*(s^{(k-1)}) \cdot (x_j^*(s^{(k-1)}))^2\Big)\Big]^+$ where $[.]^+ = \max(., 0)$	
Stopping criterion	$\lvert s^{(k)} - s^{(k-1)}\rvert \leq \epsilon_s$	

Table I
NUMERICAL SOLUTION FOR SOLVING (6)

secrecy constraint. However since (10) is not concave in P_{UX}, the Blahut-Arimoto type algorithm can be demonstrated to converge only when some specific conditions hold [24]. These conditions are given in Theorem 2 of [24]. Indeed, if the solution of (10), $P_{UX}^{*(l)}(s)$, lies in a set $T_{k,\theta}(\tilde{P}_{UX})$ and the function $L(P_{UX}, \mathcal{X}^{(\ell-1)}, s)$ is concave in $T_{k,\theta}(\tilde{P}_{UX})$ and the initial guess $P_{UX}^{(0)(l)}(s) \in T_{k,\theta}(\tilde{P}_{UX})$, the Blahut-Arimoto type algorithm is shown to converge to the optimal value. $T_{k,\theta}(\tilde{P}_{UX})$ is defined in [24] as the set of all the points $P_{UX} \in S_{k,\theta} \triangleq \{P_{UX} | L(P_{UX}, \mathcal{X}^{(\ell-1)}, s) \geq k\}$ such that P_{UX} is reachable from $\tilde{P}_{UX} \in S_{k,\theta}$ by a continuous path. Therefore, the problem is now to choose an appropriate initial point. It is observed in [15] that the size of the region $T_{k,\theta}(\tilde{P}_{UX})$ where the objective function in (10) is concave in P_{UX} is larger when θ increases. Thus we have more chance that the algorithm converges from a random initial guess in this case. In our experiments, the initial guesses are chosen randomly (avoiding "Degenerate cases" such as uniform distribution, distribution with similarities [15] and distribution with null elements) and the Blahut-Arimoto type algorithm is observed to converge to reasonable solutions, since the resulting rate regions have a very smooth shape. In the case of general superposition coding, the algorithm converges to one of the $M!$ solutions (Lemma 2). Note that when $\theta = 0$, the maxima of $I(U;Y_2)$ are obtained when $U \equiv X$. Note also that the algorithm proposed in [21] is a particular case of the Blahut-Arimoto type algorithm for the Gaussian BCCM when $\theta = 1$.

Clearly, each iteration of the alternative maximization method increases the objective function. In the experiments, we have observed that this method converges at least to a local maximum (denoted $p_{i,j}^*(s)$, $x_j^*(s)$, $0 \leq j \leq M-1$, $0 \leq i \leq |\mathcal{U}| - 1$).

Finally, the function $g(s) = \max_{P_{UX}, \mathcal{X}} L(P_{UX}, \mathcal{X}, s)$ is convex in s even $L(P_{UX}, \mathcal{X}, s)$ is not concave. This is because $L(P_{UX}, \mathcal{X}, s)$ is linear in s for each $(P_{UX}, \mathcal{X})$, and $g(s)$ is the maximum of linear functions, and is therefore convex [25]. Since $g(s)$ is convex, a gradient-type search is guaranteed to converge to the global optimum s^*. Thus in order to update the value of s, we use a gradient search method as follows:

$$s^{(k+1)} = \Big[s^{(k)} - \beta\Big(P - \sum_{i,j} p_{ij}^*(s^{(k)}) \cdot (x_j^*(s^{(k)}))^2\Big)\Big]^+ \quad (12)$$

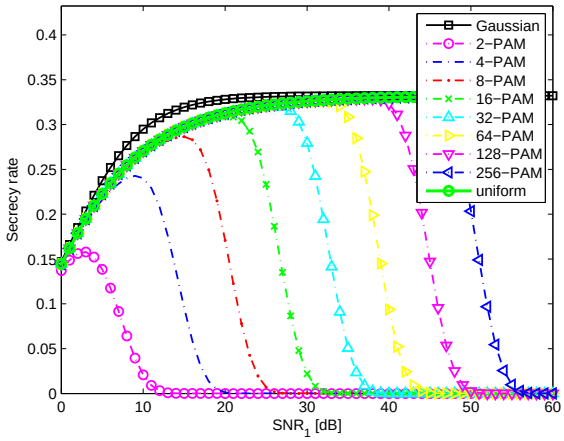

Figure 3. Secrecy rate for a Gaussian wiretap channel using Gaussian alphabet or M-PAM standard constellation where P_X is uniform. SNR_2 [dB]=SNR_1 [dB]-2 dB

where $[.]^+ = \max(., 0)$. We use a constant step size, i.e., $\beta^{(k)} = \beta$ whose value is chosen in experiments to be sufficiently small and such that the value of s does not change very much from an iteration to another. The optimal value of s is found when $|s^{(k)} - s^{(k-1)}| \leq \epsilon_s$, where ϵ_s is the target resolution.

The algorithm used to solve the optimization problem (6) is summarized in Table I.

V. RESULTS AND DISCUSSION

This section provides an evaluation of the achievable rate regions for Gaussian BCCM using various transmission strategies.

A comparison between the achievable rate regions for Gaussian BCCM using time sharing, superposition modulation and superposition coding is provided. The effect of constellation shaping is evaluated by analyzing the achievable rate region curves obtained for an M-PAM constellation ($M \in \{4, 8, 16, 32\}$) and for several pairs (SNR_1, SNR_2). The comparisons of achievable rates are conducted in terms of SNR savings for target achievable rates (Maximum Shaping Gain) and in terms of Maximum Percentage of Gain on the

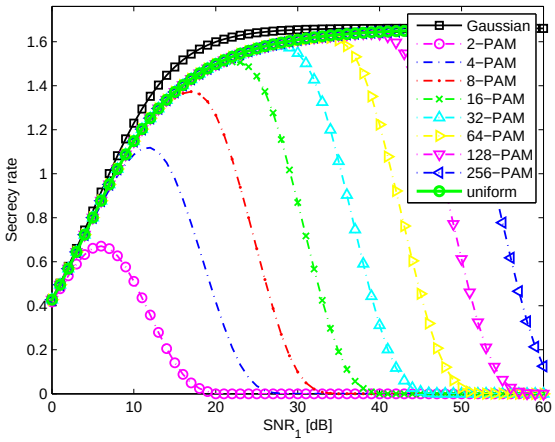

Figure 4. Secrecy rate for a Gaussian wiretap channel using Gaussian alphabet or M-PAM standard constellation where P_X is uniform. SNR_2 [dB]=SNR_1 [dB]-10 dB

common message rate R_0 or the secrecy message rate R_1 or the sum $R_0 + R_1$. These quantities are defined below.

Definition 1: Consider two transmission strategies (A and B). The pair of rates (R_1, R_0) is achieved for (SNR_1, SNR_2) with A and for $(SNR_1 + \Delta SNR, SNR_2 + \Delta SNR)$ with B. The shaping gain (with A compared to B) is ΔSNR. The maximum shaping gain is defined as:

$$MG_{SNR_{dB}}(A|B) = max_{R_0} \Delta SNR \tag{13}$$

The maximum percentage of gain on the secrecy message rate is defined below and can be defined in the same way for the other cases.

Definition 2: Consider two transmission strategies (A and B). For a given pair of SNR (SNR_1, SNR_2) and a fixed value of R_0, the achievable pair of rates is (R_1^A, R_0) resp. (R_1^B, R_0) with A resp. B. The gain on the achievable secrecy rate for user 1 is given by

$$G_{R_1}(A|B) = \frac{R_1^A - R_1^B}{R_1^B} \cdot 100 \ (\%) \tag{14}$$

The maximum gain on the achievable secrecy rate for user 1 (with A compared to B) is given by

$$MG_{R_1}(A|B) = max_{R_0} G_{R_1}(A, B) \tag{15}$$

A. Analysis of the secrecy rate

To understand the behavior of the achievable rate region curves, we begin by analyzing the secrecy rate for the wiretap channel, i.e. BCCM when $U = const.$ ($\theta = 1$). The conclusions obtained here also apply in the presence of the common message as shown in the next section.

Figures 3 and 4, show the achievable secrecy rate using standard M-PAM constellations whose symbols are used with equal probability, where $M \in \{2, 4, 8, 16, 32, 64, 128, 256\}$, the secrecy capacity achieved using Gaussian input, and the performance of a PAM constellation with uniform distribution input. Both figures depict the secrecy rate as a function of SNR_1, Fig. 3 corresponds to the case where the eavesdropper channel SNR is 2 dB below SNR_1, while in Fig. 4, the difference is 10 dB.

Obviously, the secrecy rate should increase when the gap between SNR_1 and SNR_2 increases for fixed SNR_1, this is observed in the corresponding figures. It is also observed in Fig. 3 and 4 that when the SNR for both receivers is "high", the secrecy rate is null. This is in line with the results in [8] , [9] where it is shown that when a standard finite constellation of M symbols is used and when symbols are chosen with equal probability, the optimal transmission power may not be given by the total available power, since when $P \to \infty$, both $I(X;Y_1)$ and $I(X;Y_2)$ converge to $\log_2 M$. Thus, the transmitter should use a cardinality M sufficiently high and adapted to the target SNRs n order to obtain sufficient secrecy capacity. It is also observed that the uniformly distributed input always introduces a gap with the secrecy capacity even when the cardinality of the input alphabet tends to ∞. It is seen below that constellation shaping does not result in the same drawback.

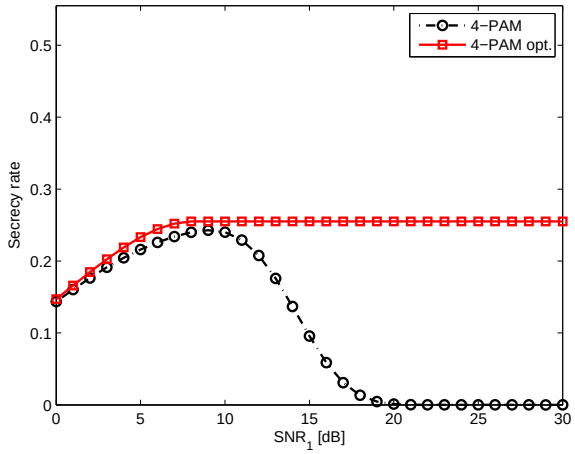

Figure 5. Secrecy rate for a Gaussian wiretap channel using 4-PAM standard constellation where P_X is uniform and when $\mathcal{X}$ and P_X are optimized (4-PAM opt.). SNR_2 [dB]=SNR_1 [dB]-2 dB

Figures 5 and 6 show also the secrecy rate using standard 4-PAM constellation and the maximal secrecy rate obtained by maximizing over both symbol positions and their probabilities as a function of SNR_1, where the eavesdropper channel SNR is 2 dB and 10 dB respectively below SNR_1. In Fig. 5 and 6, it can be observed that the joint optimization of symbol positions and their probabilities seems to bring moderate gains in secrecy rate for small SNRs. However, when translated in SNR improvement, this gain is far from negligible. For example, in Fig. 5, the 4-PAM standard with uniform P_X achieves a secrecy rate $R_1 = 0.23$ bit/ch.use when $SNR_1 = 7$ dB ($SNR_2 = 5$ dB) which is achieved using the optimized 4-PAM at almost $SNR_1 = 5$ dB ($SNR_2 = 3$ dB); thus the shaping gain is close to 2 dB in terms of SNR. The optimal symbol positions when $SNR_1 = 5$ dB ($SNR_2 = 3$ dB) are given by $\mathcal{X} = \{4.03, 1.29, -1.29, -4.03\}$ and the optimal $P_X = \{0.114, 0.386, 0.386, 0.114\}$. We observe that optimal probabilities of symbols near origin are higher than the ones of

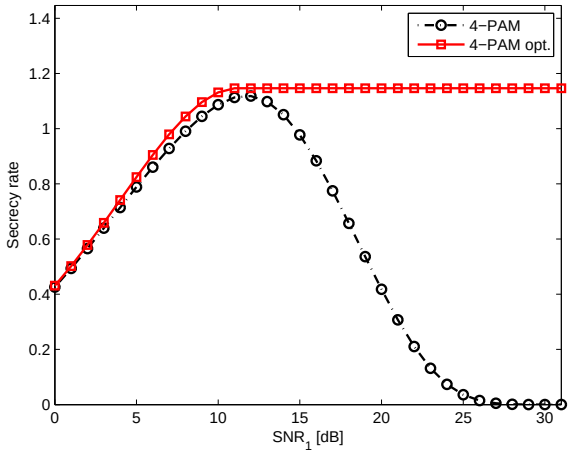

Figure 6. Secrecy rate for a Gaussian wiretap channel using 4-PAM standard constellation where P_X is uniform and when $\mathcal{X}$ and P_X are optimized (4-PAM opt.). SNR_2 [dB]=SNR_1 [dB]-10 dB

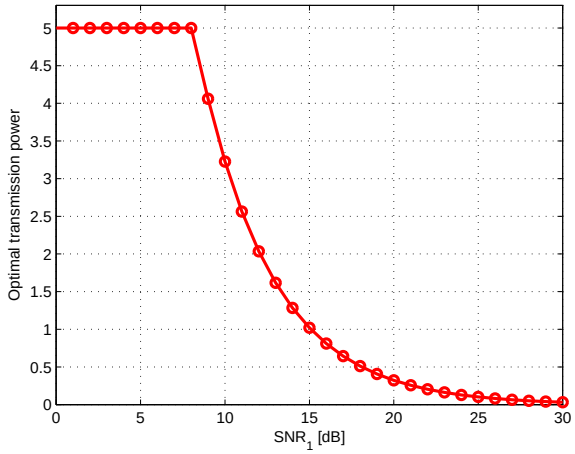

Figure 7. Optimal transmission power for a Gaussian WTC channel using 4-PAM constellation given that the maximal allowed power is equal to $P = 5$. $SNR_2 = SNR_1 - 2$ dB.

symbols far from origin. Thus, the optimal distribution is very similar to the sampling of a Gaussian distribution. We note that in Fig. 5 and 6, the optimal transmission power when $SNR_1 < 9$ dB and $SNR_1 < 12$ dB respectively is given by the maximal available power as shown in Fig. 7 when $SNR_2 = SNR_1 - 2$ dB.

The main difference between uniform and optimized constellations is that when SNR_1 is higher than a certain value, full optimization of the secrecy rate prevents the secrecy rate to vanish and brings significant gains compared to standard 4-PAM with equally probable symbols (eg. $SNR_1 \geq 9$ dB in Fig. 5). The optimal transmit power is less than the maximal available power when $SNR_1 \geq 9$ dB in Fig. 5 and it decreases with SNR_1 as shown in Fig. 7. This was already explained in [8], [9] for uniform constellations. The novelty here is the improvement brought by constellation shaping. Note also that when $SNR_1 \geq 9$ dB in Fig. 5, the optimal probability distribution is given by $P_X = \{0.169, 0.331, 0.331, 0.169\}$, $\forall SNR_1$ and only symbol positions change with SNR_1 in order to conserve the maximum value of the secrecy rate when SNR_1 and SNR_2 increase.

The next subsections are concerned not only with the secrecy rate, but also with the tradeoff between the achievable common rate and the corresponding secrecy rate.

B. Superposition modulation using M-PAM

The achievable rate region computation for superposition modulation with $M = 4$ and using equiprobable symbols ($SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$) does not require to solve any optimization problem. In $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ strategy, the total power P is split such that $\alpha \cdot P$ is used for the alphabet of the secret information and $(1-\alpha) \cdot P$ for the alphabet of the common information, with $\alpha \in [0, 1]$. Thus, the four transmitted signal constellation symbols can be expressed as a function of α only [26]. Consequently, obtaining the maximal achievable rate region for $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ and with $M = 4$, involves the computation of $I(X;Y_1|U) - I(X;Y_2|U)$ and $I(U;Y_2)$

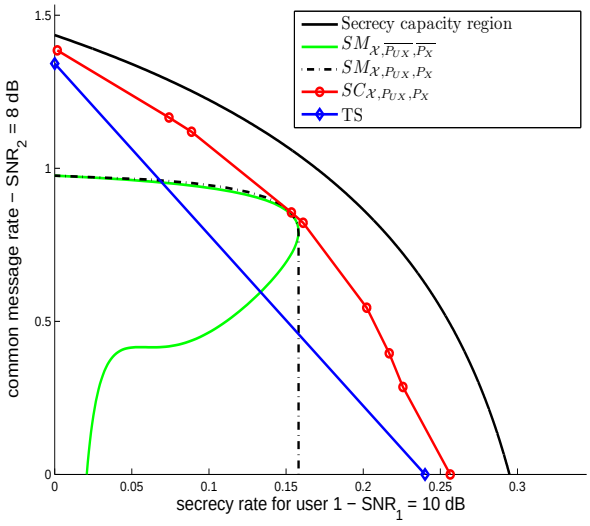

Figure 8. Achievable rate regions with $M = 4$ and $(SNR_1, SNR_2) = (10, 8)$ dB

in function of α, knowing that P_{UX} is uniform, and then to vary α between 0 and 1. In Figs. 8 and 9, the achievable rate regions are illustrated for the Gaussian BCCM with 4-PAM inputs using various broadcast strategies when SNR_1= 10 dB and $SNR_2 \in \{0, 8\}$ dB.

The particular shape of the achievable rate region using superposition modulation scheme in Fig. 8, comes also from the fact that the common rate is not necessarily increasing when the portion of power allocated to the common message increases using finite alphabet inputs as explained in [27]. For a fixed secrecy rate, the curve of achievable rate region using superposition modulation with equally probable symbols can have two possible values of the common message rate which is not the case of the secrecy capacity region achieved with Gaussian inputs. Thus, it is necessary to choose the good portion of power for each message in order to avoid the "bad" part of the rate region.

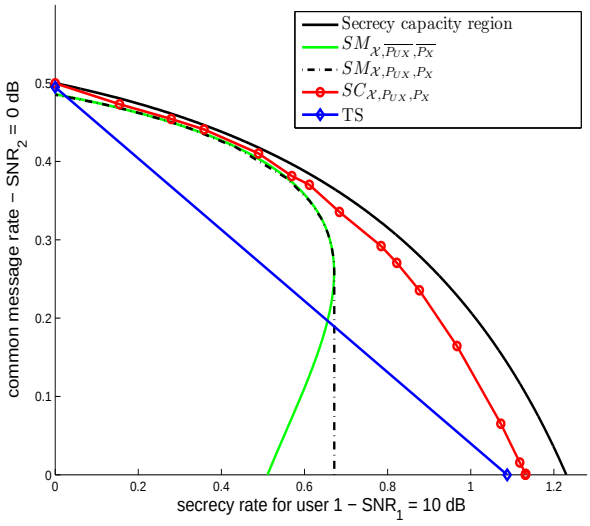

Figure 9. Achievable rate regions with $M = 4$ and $(SNR_1, SNR_2) = (10, 0)$ dB

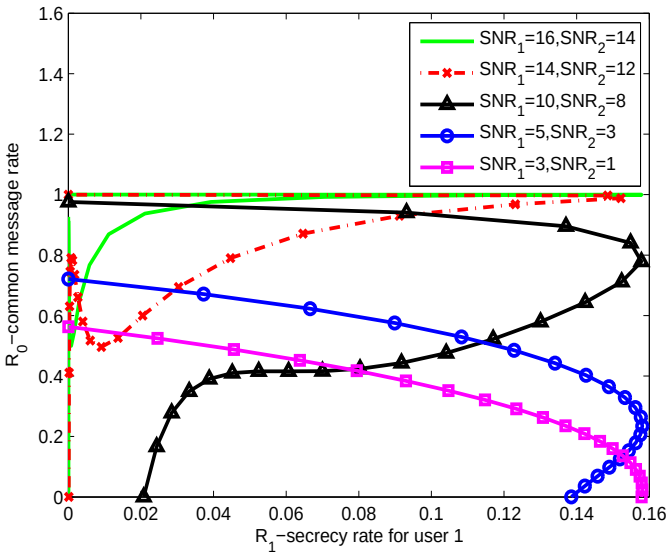

Figure 10. Achievable rate regions with $M = 4$ and for superposition modulation where symbols are used with equal probability. The SNRs are in dB.

Consider for example the case where SNR_1= 10 dB and $SNR_2 = 0$ dB. Users 1 and 2 receive the secret information with a SNR equal to $SNR'_1 = \alpha \cdot \frac{P}{\sigma_1^2}$ and $SNR'_2 = \alpha \cdot \frac{P}{\sigma_2^2}$ respectively. When $\alpha = 1$, $SNR'_1 = 10$ dB and $SNR'_2 = 0$ dB, the secrecy capacity is equal to 0.51 bit/ch.use using a 2-PAM constellation according to Fig. 4. We observe also in this figure that the maximal secrecy rate is obtained when $SNR_1 = 6$ dB ($SNR_2 = -4$ dB) and is equal to 0.6711 bit/ch.use. Thus the optimal $\alpha = \alpha^*$ which maximize the secrecy rate in the case of superposition modulation is such that $\alpha^* \cdot \frac{P}{\sigma_1^2} = 6$ dB. Obviously if we solve (6) we cannot obtain the region when $\alpha > \alpha^*$ because it is not optimal, in other terms, it does not correspond to the solution of any $\theta \in [0, 1]$. This is what we can observe also from the achievable rate regions using $\{8, 16, 32\}$-PAM.

M	SNR_1	SNR_2	$MG_{R_{0\|1}}(A\|B)$	$MG_{SNR_{dB}}(A\|B)$
4	10	8	0.06%	0.24
		6	0.477%	0.1
		4	0.34%	0.03
		2	0.14%	0
8	16	14	$5.15\%^{(M_1=4,M_2=2)}$	0.36
		12	$5.3\%^{(M_1=4,M_2=2)}$	0.43
		10	$5.14\%^{(M_1=4,M_2=2)}$	0.4
		8	$5.02\%^{(M_1=4,M_2=2)}$	0.38
16	20	18	$7.06\%^{(M_1=4,M_2=4)}$	0.61
		16	$5.93\%^{(M_1=4,M_2=4)}$	0.57
		14	$8\%^{(M_1=4,M_2=4)}$	0.54
		12	$8.48\%^{(M_1=4,M_2=4)}$	0.43

Table II
COMPARISON OF $SM_{\mathcal{X},P_{UX},P_X}$ (A) AND $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ (B) WITH RESPECT TO MG_{R_1} OR MG_{R_0} AND $MG_{SNR_{dB}}(A|B)$

In Fig. 10, achievable rate region with 4-PAM using $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ is given for several pairs (SNR_1,SNR_2) such that $SNR_1 - SNR_2 = 2$ dB. We observe that the maximal secrecy rate is the same for all pairs and is achieved for $\alpha < 1$ when $SNR_1 > 3$ dB. However when $SNR_1 = 3$ dB, the maximal achievable secrecy rate is when $\alpha = 1$ as $SNR_1 = 3$ dB maximizes the secrecy rate for a 2-PAM with a gap equal to 2 dB between user SNRs (see Fig. 3).

Regions of achievable rates (Fig 8 to 15) show the improvement obtained by optimizing symbol positions and the joint probabilities ($SM_{\mathcal{X},P_{UX},P_X}$ (full optimization) compared to $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ (optimization of $\mathcal{X}$ only)). In table II the maximum gain in achievable rate on R_1 or R_0 is given ($MG_{R_1}(A|B)$ or $MG_{R_0}(A|B)$), depending if the full optimization provides an horizontal or vertical gain in achievable rates. The maximum SNR savings ($MG_{SNR_{dB}}$) are also given in table II for the 4-PAM, the 8-PAM when $M_1 = 4, M_2 = 2$ and for the 16-PAM when $M_1 = 8, M_2 = 2$. For the other cases of 8-PAM ($M_1 = 2, M_2 = 4$) and 16-PAM ($M_1 = 4, M_2 = 4$ and $M_1 = 2, M_2 = 8$), we did not evaluate the gain in SNR because the maximum secrecy rate obtained by full optimization ($SM_{\mathcal{X},P_{UX},P_X}$) can not be reached by superposition modulation using equally probable symbols even when we increase the user SNRs. This is due to the fact that in these cases and for the considered values of user SNRs, the maximum secrecy rate will not necessarily increase when the user SNRs increase as can be seen in Fig. 10. One can observe that the maximum shaping gain increases with the constellation size. Thus, constellation shaping for SM strategy seems more useful for high values of M. Moreover, we observe that independently of M, the maximum shaping gain is very small when the gap between the user SNRs increases. This is also the case for a broadcast channel model without secrecy constraints [19]. The analysis of the optimal matrix P_{UX} (results not reported) when $X = X_1 + X_2$, such that X_1 and X_2 are two signals carrying the secret information and the common information respectively, leads to the conclusion that X_1 and X_2 are not independent in general when using finite-size constellations.

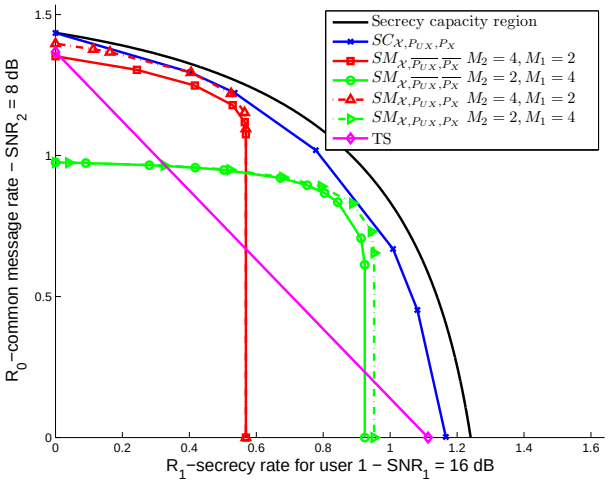

Figure 11. Achievable rate regions with $M = 8$ and $(SNR_1, SNR_2) = (16, 8)$ dB

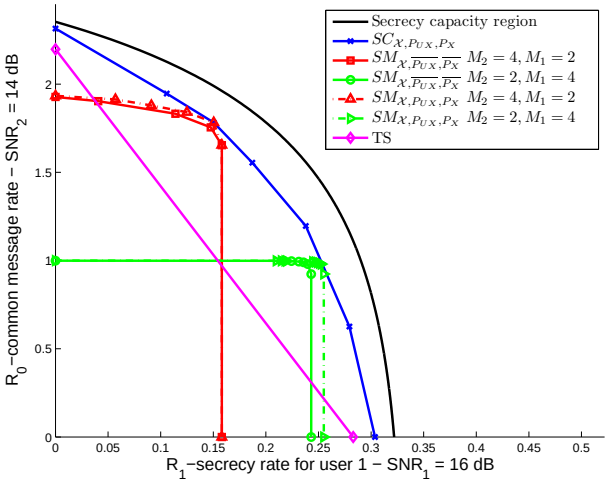

Figure 12. Achievable rate regions with $M = 8$ and $(SNR_1, SNR_2) = (16, 14)$ dB

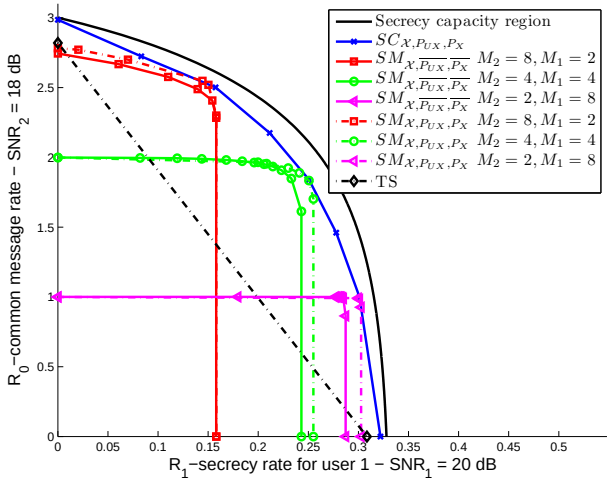

Figure 13. Achievable rate regions with $M = 16$ and $(SNR_1, SNR_2) = (20, 18)$ dB

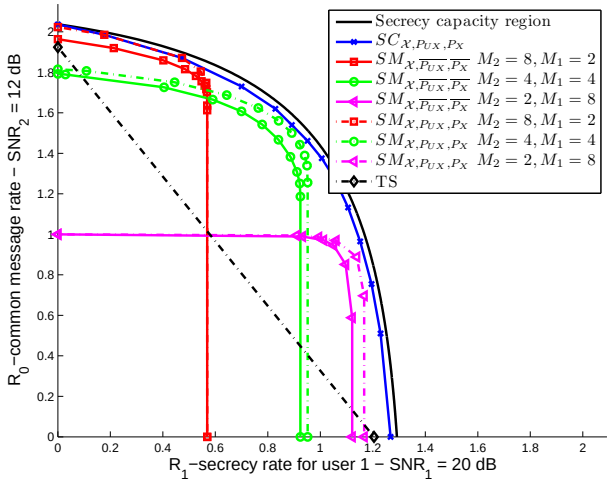

Figure 14. Achievable rate regions with $M = 16$ and $(SNR_1, SNR_2) = (20, 12)$ dB

C. Superposition modulation vs time sharing

This section compares the achievable rates for the Gaussian BCCM using two classical schemes: time sharing using standard constellation and superposition modulation. Moreover, we consider the case where symbols are used with equal probability for practical constraints. Figures 8-15 show that the achievable rate region can be divided into two parts, such that in each part, one strategy is more efficient than the other. This is also what is observed in [19] for a broadcast channel model without secrecy constraints. The efficiency of time sharing strategy increases with respect to superposition modulation when SNR_1 and SNR_2 become closer. Table III shows the maximum percentage of improvement in achievable rate by user 1 $(R_0 + R_1)$ using $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$, comparing to TS strategy, in the achievable rate area where $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ outperforms TS. It can also be observed that the best improvement happens when δ_{SNR} increases for all $M \in \{4, 8, 16\}$. Thus superposition modulation should be preferred to time sharing when users have very different SNRs.

D. Superposition coding

It is well known that the secrecy-capacity region for the Gaussian BCCM is achievable using superposition modulation scheme $(SM_{\mathcal{X},P_{UX},P_X})$, a.k.a. signal superposition, where $U = X_2$. However, in the finite-input alphabet case, the results show that the general case of superposition coding, $SC_{\mathcal{X},P_{UX},P_X}$, outperforms superposition modulation in terms of achievable rate region. A detailed discussion about this result for the two-user broadcast channel without secrecy constraint was given in [19]. In table III, the maximum percentage of improvement in achievable rate by user 1 $(R_0 + R_1)$ is given using $SC_{\mathcal{X},P_{UX},P_X}$, comparing to $SM_{\mathcal{X},P_{UX},P_X}$ (full optimization). It can be observed that the maximum gain is proportionally greater for small values of M since superposition modulation offers less flexibility in this case, while superposition coding keeps all its power.

M	SNR_1	SNR_2	$MG_{R_1+R_0}(A\|B)$	$MG_{R_1+R_0}(A\|C)$
4	10	8	4.51%	32.33%
		6	12.54%	11.57%
		4	24.06%	9.86%
		2	35.22%	9.7%
		0	48.23%	17.562%
8	16	14	$7.3\%^{(M_1=4,M_2=2)}$	17.48%
		12	$14.23\%^{(M_1=4,M_2=2)}$	5.4%
		10	$22.12\%^{(M_1=4,M_2=2)}$	1.03%
		8	$32.1\%^{(M_1=4,M_2=2)}$	2.99%
16	20	18	$7.09\%^{(M_1=8,M_2=2)}$	6.93%
		16	$13.73\%^{(M_1=8,M_2=2)}$	1.24%
		14	$19.94\%^{(M_1=8,M_2=2)}$	0.11%
		12	$30.31\%^{(M_1=4,M_2=4)}$	3.56%

Table III
COMPARISON OF $SM_{\mathcal{X},\overline{P_{UX}},\overline{P_X}}$ (A) VS TS (B). COMPARISON OF $SC_{\mathcal{X},P_{UX},P_X}$ (A) VS $SM_{\mathcal{X},P_{UX},P_X}$ (C).

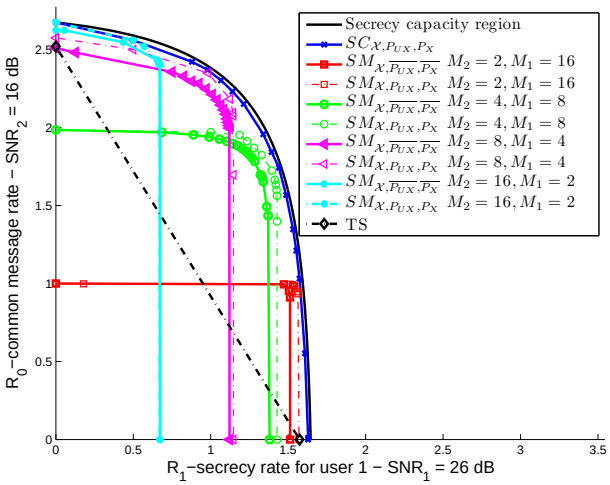

Figure 15. Achievable rate regions with $M = 32$ and $(SNR_1, SNR_2) = (26, 16)$ dB

VI. CONCLUSION

In this paper, we derived the achievable rate region for the Gaussian broadcast channel with confidential message using finite input constellations for various broadcast strategies. The simulation conducted in the previous section gives insights for choosing the best strategy for a given situation. For superposition modulation and the general case of superposition coding, the achievable rate regions are maximized by optimizing over symbol positions and over the joint distribution of probability. It is shown that, for finite modulations, the optimal transmission power which maximizes the secrecy rate may not be given by the total available power, whatever the allowed flexibility (superposition modulation or coding). In addition, full maximization of achievable rate region for superposition modulation (a suboptimal strategy with reasonable complexity) provides more significant improvements when the cardinality of the input alphabet increases compared to the case where only symbol positions are optimized. It is also observed that superposition modulation should be preferred to time sharing when users have very different SNRs. In that case, full optimization is not necessary.

In the case of BCCM with finite input alphabet, superposition modulation is not the optimal strategy, like in the Gaussian alphabet case. The general case of superposition coding can provide significant gains when compared to practical schemes.

However in many cases, using practical schemes can achieve near optimal rates and provides a good tradeoff between complexity of implementation and efficiency. This paper provides tools allowing to choose the appropriate complexity/efficiency tradeoff.

APPENDIX

Proof of Lemma 1 (i): First, we recall that a function $f(P_{UX})$ is a concave function of the probability distribution P_{UX} if for all $\alpha, 0 \leq \alpha \leq 1$, and all probability distributions P^a_{UX} and P^b_{UX},

$$\alpha f(P^a_{UX}) + (1-\alpha) f(P^b_{UX}) \leq f\big(\alpha P^a_{UX} + (1-\alpha) P^b_{UX}\big)$$

To prove Lemma 1 (i), we use the method for proving [22, Theorem 2]. Indeed, for an arbitrary V with finite input alphabet $\mathcal{V}$, consider the Markov chain $(U,V) \to X \to (Y_1, Y_2)$. Each $v \in \mathcal{V}$ specifies a probability distribution P^v_{UX} for X in the manner

$$P^v_{UX}(u,x) = P_{UX|V}(u,x|v), \qquad (u,x) \in \mathcal{U} \times \mathcal{X}$$

We first note that

$$\begin{aligned} I(XV;Y_1|U) &\overset{i)}{=} I(X;Y_1|U) + I(V;Y_1|UX) \\ &\overset{ii)}{=} I(X;Y_1|U) \end{aligned} \tag{16}$$

where i) follows from the chain rule for mutual information and where ii) follows from the fact that $(U,V) \to X \to (Y_1, Y_2)$ is a Markov chain so that $I(V;Y_1|UX) = 0$. We note also that

$$I(XV;Y_1|U) \overset{iii)}{=} I(V;Y_1|U) + I(X;Y_1|UV) \tag{17}$$

where iii) follows from the chain rule for mutual information. By combining (16) and (17), we can write:

$$I(V;Y_1|U) = I(X;Y_1|U) - I(X;Y_1|UV) \tag{18}$$

In the same way, we can show that

$$I(V;Y_2|U) = I(X;Y_2|U) - I(X;Y_2|UV) \tag{19}$$

From (18) and (19), we infer that

$$I(V;Y_2|U) \leq I(V;Y_1|U)$$

if and only if

$$I(X;Y_1|UV) - I(X;Y_2|UV) \leq I(X;Y_1|U) - I(X;Y_2|U) \tag{20}$$

But it can be shown, using the definition of mutual information, that

$$\begin{aligned} &I(X;Y_1|UV) - I(X;Y_2|UV) \\ &= \sum_{v\in\mathcal{V}} P_V(v) \cdot \Big[I(X;Y_1|U,V=v) - I(X;Y_2|U,V=v) \Big] \\ &= \sum_{v\in\mathcal{V}} P_V(v) \cdot \Big[I(X;Y_1|U) - I(X;Y_2|U) \Big]_{P^v_{UX}} \end{aligned} \tag{21}$$

and

$$\begin{aligned} &I(X;Y_1|U) - I(X;Y_2|U) \\ &= \Big[I(X;Y_1|U) - I(X;Y_2|U) \Big]_{\sum_{v\in\mathcal{V}} P_V(v)\cdot P^v_{UX}} \end{aligned} \tag{22}$$

Using the definition of a concave function, the part (i) of Lemma 1 follows immediately from (20), (21) and (22).

ACKNOWLEDGMENT

The authors would like to thank Dr. Maël Le Treust for the helpful discussions and Prof. Pablo Piantanida for the valuable comments and suggestions to improve the quality of the paper.

REFERENCES

[1] A. D. Wyner, "The wiretap channel," *Bell System Technical Journal*, vol. 54, pp. 1355–1387, 1975.

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, 1978.

[3] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[4] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[5] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, 2008.

[6] R. Liu, T. Liu, H. Poor, and S. Shamai, "New results on multiple-input multiple-output broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1346–1359, 2013.

[7] J. Li and A. Petropulu, "On ergodic secrecy rate for Gaussian MISO wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1176–1187, April 2011.

[8] G. D. Raghava and B. S. Rajan. Secrecy capacity of the Gaussian wiretap channel with finite complex constellation input. [Online]. Available: http://arxiv.org/abs/1010.1163

[9] F. Renna, N. Laurenti, and H. V. Poor, "Achievable secrecy rates for wiretap OFDM with QAM constellations," in *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools, VALUETOOLS '11*, Paris, France, 2011.

[10] M. R. D. Rodrigues, A. Somekh-Baruch, and M. Bloch, "On Gaussian wiretap channels with M-PAM inputs," in *European Wireless Conference (EW)*, 2010, pp. 774–781.

[11] S. Bashar, Z. Ding, and C. Xiao, "On secrecy rate analysis of MIMO wiretap channels driven by finite-alphabet input," *IEEE Trans. on communications*, vol. 60, no. 12, pp. 3816–3825, dec. 2012.

[12] ——, "On the secrecy rate of multi-antenna wiretap channel under finite-alphabet input," *IEEE Communications Letters*, vol. 15, no. 5, pp. 527–529, May 2011.

[13] H. Qin, Y. Sun, T.-H. Chang, X. Chen, C.-Y. Chi, M. Zhao, and J. Wang, "Power allocation and time-domain artificial noise design for wiretap OFDM with discrete inputs," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2717–2729, June 2013.

[14] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Trans. Veh. Tech.*, vol. 61, no. 6, pp. 2599–2612, July 2012.

[15] Z. Mheich, F. Alberge, and P. Duhamel, "The impact of finite-alphabet input on the secrecy-achievable rates for broadcast channel with confidential message," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, 2nd ed., ser. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, Jul. 2006.

[17] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, october 1998.

[18] M. Bloch and J. Barros, *Physical layer security: from information theory to security engineering*. Cambridge University Press, 2011.

[19] Z. Mheich, F. Alberge, and P. Duhamel, "Achievable rates optimization for broadcast channels using finite size constellations under transmission constraints," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, p. 254, 2013. [Online]. Available: http://jwcn.eurasipjournals.com/content/2013/1/254

[20] ——, "On the efficiency of transmission strategies for broadcast channels using finite size constellations," in *Proc. of the 21st European Signal Processing Conference*, Marrakech, sept. 2013.

[21] K. Yasui, T. Suko, and T. Matsushima, "An algorithm for computing the secrecy capacity of broadcast channels with confidential messages," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.

[22] M. van Dijk, "On a special class of broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 712–714, 1997.

[23] E. Calvo, D. P. Palomar, J. R. Fonollosa, and J. Vidal, "The computation of the capacity region of the discrete degraded BC is a nonconvex DC problem," in *IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 2008.

[24] K. Yasui and T. Matsushima, "Toward computing the capacity region of degraded broadcast channel," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, June 2010, pp. 570–574.

[25] D. P. Bertsekas, *Nonlinear Programming*, second edition ed. Athena Scientific, 1999.

[26] Z. Mheich, P. Duhamel, L. Szczecinski, and M.-L. Alberi-Morel, "Constellation shaping for broadcast channels in practical situations," in *Proc. of the 19th European Signal Processing Conference*, Barcelona, Spain, Aug. 2011.

[27] C. Huppert and M. Bossert, "On achievable rates in the two user AWGN broadcast channel with finite input alphabets," in *IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 2007.