Cognitive Radio with Partial Channel State Information at the Transmitter

Pin-Hsun Lin, Shih-Chun Lin, Chung-Pi Lee, and Hsuan-Jung Su

Abstract

Cognitive radio (CR) has been proposed as an efficient method to reuse the licensed spectrum. It recognizes the primary (licensed) users’ signals and adapts its own to minimize the interference it generates. When perfect channel state information is known at the transmitters, the capacity of CR system can be achieved by utilizing the dirty paper coding (DPC). In this paper, we consider the performance of CR systems under both fast and slow fading channels with only channel statistics known at the transmitters. Due to the limited channel state information, the original DPC fails and is replaced by the so-called linear-assignment Gel’fand-Pinsker coding. By carefully designing the parameters based on this precoding, we show that significant rate gains over naively treating primary users’ signals as interference can be obtained for both fast and slow fading scenarios. A nested-lattice based coding/decoding scheme is also proposed to implement this precoding in practice, and it validates our theoretical claims.

I. INTRODUCTION

As the demands for data rates steadily increase, efficient spectrum usage becomes a critical issue. Recent measurements from the Federal Communications Commission (FCC) have indicated that 90 percent of the time, many licensed frequency bands remain unused. A radio technology that promises to solve such problems is the cognitive radio (CR) [1]. This technology is capable of dynamically sensing and locating unused spectrum segments in a target spectrum pool, and communicating via the unused spectrum segments.
without causing harmful interference to the primary (licensed) users of the spectrum. Primary users are defined as users of existing commercial standards. If a higher priority entity, e.g., a primary user demands the channel, the CR users should vacate and find an alternative channel. In [2], Devroye et. al. categorized three kinds of CR behaviors: 1) Interference mitigation: Both CR and primary users can simultaneously transmit at the same time or frequency bands. 2) Collaboration: CR users act as relays. 3) Interference avoidance: CR users transmit over a certain time or frequency band only when there are no other users.

Instead of simple interference avoidance, Devroye et. al. [3] applied the concept of dirty paper coding (DPC) [4] to implement the behavior of interference mitigation. Dirty paper coding is a promising precoding technique for cancelling arbitrary interference perfectly known only at the transmitter but not at the receiver. Surprisingly, with DPC the interference-free rate is achievable. To make the application of DPC to CR systems feasible, [3] [5] assume that the codewords of the primary users, as well as their channel gains are non-causally known at the CR transmitter (TX). This assumption may hold when CR and primary TX are close enough to each other and is discussed in [6]. The CR TX not only transmits its own signals but also relays those of the primary users. This is the key to keep primary users’ transmission rates unchanged under the coexistence of CR users since CR users’ signals will interfere the primary users. For CR receiver (RX), signals from primary TX and those relayed by CR TX are interferences. These interferences are known at the CR TX and can be precoded by the DPC such that the CR RX may not see these interferences. Devroye et. al. [3] derived the achievable rate of the DPC based interference mitigation CR systems. In [5], Jovicic and Viswanath further showed that this DPC based CR system indeed achieves the capacity by using the same technique used in proving the capacity region of the multiple input multiple output (MIMO) broadcast channel (BC).

Due to the limited feedback bandwidth and the speed of channel estimation, feedback of partial channel state information (CSI) is more practical in many situations. We thus focus on the fading CR problem where only the channel statistics are available at the transmitter. However, the interference-free rate achieved by DPC relies on the perfectly known CSI at the transmitter. Results from other systems such as MIMO BC with partial CSI [7] showed that naively applying DPC in fading channels might cause a significant performance loss. Thus we use the linear-assignment Gel’fand-Pinsker coding (LA-GPC) [7] which has
been proved having good (sometimes interference free) performance in a variety of fast and slow fading channels [7], [8]. Based on this coding, we propose design methods [9] to optimize the performance of CR with partial CSI at the TX for both fast and slow fading scenarios. According to our simulations, significant gains from precoding over naively treating primary user’s signals as interference at CR RX can be achieved. In addition to the results based on unstructured random Gaussian codebook, we also use nested-lattice codebook with lattice decoding to implement this precoding in practice [10]. Again, significant performance is obtained and this verifies our theoretical results.

This paper is organized as follows. Section II reviews the LA-GPC and shows the model of the partial CSI CR system. The proposed application of LA-GPC to the partial CSI CR system are also given in this section. The proposed parameters design under fast and slow fading channels are discussed in Section III and Section IV. A nested-lattice based linear-assignment Gel’fand-Pinsker coding scheme using the proposed design parameters is illustrated in Section V. Simulation results are illustrated in Section VI. Finally Section VII concludes this paper.

II. BACKGROUND AND SYSTEM MODEL

In this section we first introduce the considered system model. We then review the problem of DPC and the LA-GPC under fading channels. After that we will show how to apply the LA-GPC in the fading cognitive channel in Section II-C.

A. The cognitive channel

The cognitive channel coined by [3] is shown in Fig. 1. The channel gains between primary TX and RX, primary TX and CR RX, CR TX and primary RX, and CR TX and RX are denoted by \( H_{11}, H_{21}, H_{12}, \) and \( H_{22} \) respectively. The power constraints of the primary TX and CR TX signals \( X_p \) and \( X_c \) are denoted by \( P_p \) and \( P_c \) respectively. The unidirectional arrow from the primary TX to the CR TX means that the

*In this paper, the superscript * denotes the complex conjugate. Boldface capital letters denote deterministic matrices. \( A(i,j) \) denotes the entry at the \( i \)th row and the \( j \)th column of the matrix \( A \). \( \Sigma_{A,B} \) denotes the covariance matrix of the random vector \( [AB]^T \) where the superscript \( T \) denotes the transpose. \( |A| \) and \( |a| \) represent the determinant of the square matrix \( A \) and the absolute value of the scalar variable \( a \) respectively. All the logarithm operations are of base 2 such that the unit of rates is in bit.
CR TX knows the primary user’s codewords non-causally. The feasibility of this assumption is discussed in [5]. The signals at the primary and CR receivers are

\[
\begin{bmatrix}
Y_p \\
Y_s
\end{bmatrix} = \begin{bmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{bmatrix} \begin{bmatrix}
1 & 0 \\
\sqrt{\alpha_1 P_c / P_p} & 1
\end{bmatrix} \begin{bmatrix}
X_p \\
\hat{X}_c
\end{bmatrix} + \begin{bmatrix}
Z_p \\
Z_s
\end{bmatrix},
\]

(1)

where \(\hat{X}_c\) is CR user’s own signal and \(X_p\) is primary user’s signal known by CR TX. The CR TX relays primary user’s signal to maintain the performance of the primary link. \(\alpha_1\) is the relaying ratio which is defined as the ratio of the CR transmitted power used for relaying primary signal to the CR transmit power constraint. Thus the factor \(\sqrt{\alpha_1 P_c / P_p}\) is to ensure the relaying power is \(\alpha_1 P_c\). \(Z_p \sim cn(0, \sigma_{zp}^2)\) and \(Z_s \sim cn(0, \sigma_{zs}^2)\) are the complex Gaussian noises at primary and CR RX, respectively.

In the following discussion we assume that primary RX knows \(H_{11}\) perfectly, but only the statistics of \(H_{11}\) is known at primary TX. Similarly, the CR RX knows all channel gains perfectly, but only statistics of them are known at CR TX. The CR TX/RX can get these channel information by the methods described in [5]. Note that since the exact channel gains are unknown to the CR TX, the cognitive \((1, a, b, 1)\) channel [5] where channel gains are used to normalize users’ signals can not be used to derive the rate under partial CSI. We will derive the CR rate function without such a simplification. We assume that primary and CR users operate under the same environment, i.e., both are subject to the same (slow and fast) fading channel. In the slow fading scenario, all channel gains are random but fixed within several channel coding codeword lengths.

For simplicity, we consider the complex Rician channel, with mean and variance of \(H_{ij}\) denoted by \(\mu_{ij}\) and \(\sigma_{ij}^2\), \(1 \leq i, j \leq 2\), respectively. Also \(|\mu_{ij}|^2 + \sigma_{ij}^2 = 1\) to keep the calculation of the received signal-to-noise ratio (SNR) simple. The \(K\)-factor is defined as \(K \triangleq |\mu_{ij}|^2 / \sigma_{ij}^2\). However, our results can be extended to more general channel distributions.

B. Review of linear-assignment Gel’fand-Pinsker coding

Since only the statistics of the channels are known at the transmitter, the CR user cannot use the classical DPC with known \(X_p\) to cancel the impact of \(X_p\) at its receiver. Instead, we can use LA-GPC which is more general than DPC for this task. To illustrate the concept of this coding, we use the following channel
in the Shannon random-coding settings

\[ Y = H_x X + H_s S + Z, \]  

(2)

where \( X \) is the transmitted signal, \( S \) is the known interference (side-information) at the transmitter, and \( Z \) is the noise. All three terms are Gaussian random variables. The fading channel gains \( H_x \) and \( H_s \) are known perfectly at RX but only their statistics are known at TX. For illustration purpose, we first review the original non-fading dirty-paper coding setting with \( H_x = H_s = 1 \) [4] as \( Y = X + S + Z \). Based on the results of Gel’fand and Pinsker [11] where encoding in cosets is used, the capacity of this channel with non-causally known side information \( S \) at the transmitter is

\[ \sup \{ I(U;Y) - I(U;S) \}, \]  

(3)

where \( U \) is an auxiliary random variable with distribution obtained via the conditional distribution \( f(u|s) \), and \( f(\cdot) \) is a deterministic function such that \( X = f(U, S) \). For any particular choices of \( f(u|s) \) and \( f(\cdot) \), the rate \( I(U;Y) - I(U;S) \) is achievable. It is shown in [4] that the following “linear-assignment” strategy is optimal for this non-fading dirty-paper channel

\[ U = X + \alpha_2 S, \]  

(4)

where \( X \) is independent of \( S \) and \( \alpha_2 \) is a coefficient to be defined later. Note that strategy (4) specifies the conditional distribution \( f(u|s) \) and the function \( f(U, S) = U - \alpha_2 S \). By choosing \( \alpha_2 \) as the linear minimum mean square error (MMSE) filter to estimate \( X \) from the interference-free channel \( Y = X + Z \), the interference-free rate is achievable in this non-fading case. As in [7], we call the coding scheme using the strategy (4) (\( \alpha_2 \) is not restricted to the MMSE coefficient) as LA-GPC.

We now introduce the applications of LA-GPC in fading paper channels (2). Note that in our setting, the selection of the filter \( \alpha_2 \) must rely only on the channel statistics, and the MMSE filter used in Costa’s DPC is not applicable, nor optimal in this case. We first focus on the ergodic fast fading channels. Similar to (3), the capacity of this channel [11] is

\[ \sup \{ I(U;Y, H_x, H_s) - I(U;S) \}. \]  

(5)
The strategy to optimize (5) is still an open problem, thus we focus on the following achievable rate

\[ I(U; Y, H_x, H_s) - I(U; S) \]  

(6)

with the “linear-assignment” selection (4). This rate is called the *linear-assignment capacity* in [7]. The optimality of linear assignment in this ergodic fading channel is still unknown. However, when \( H_x = H_s \), this linear-assignment strategy is shown to have close to optimal (interference-free) rate performance in some SNR regions [8]. Also it has good performance in the MIMO setting [7].

For the quasi-static slow fading channels, the decoding error probability cannot be arbitrarily small since the transmitter does not know the reliable transmission rate from the limited channel knowledge. In this channel, the outage probability [12] for certain transmission rate \( R_0 \) is a better metric than the Shannon capacity to measure the performance. It is defined as

\[ P\{H_x, H_s : R(H_x, H_s) < R_0\}, \]  

(7)

where

\[ R(h_x, h_s) \triangleq I(U; Y|H_x = h_x, H_s = h_s) - I(U; S). \]

The explicit formula of \( R(h_x, h_s) \) in CR settings is calculated as (9) in Section II-C. In [8], it is shown that if \( H_x = H_s \), linear-assignment strategies achieve interference-free outage performance in the scalar channel with a properly selected \( \alpha_2 \).

**C. Linear-assignment Gel’fand-Pinsker coding in fading CR channel**

Since the CR user only perfectly knows the primary user’s signal \( X_p \) but not the channel gains, we let the side information of the LA-GPC be \( S = X_p \). Other variables in (2) are replaced by

\[ H_s = H_{21} + \sqrt{\frac{\alpha_1 P_c}{P_p}} H_{22}, \quad H_x = H_{22}, \quad Z = Z_s \]  

(8)

\[ X = \tilde{X}_c, \quad U = \tilde{X}_c + \alpha_2 X_p. \]
\[ \Sigma_{U,Y_s} = \begin{pmatrix} \sigma^2_{x_c} + |\alpha_2|^2 P_p & h^*_{22} \sigma^2_{x_c} + \alpha_2 (h_{21} + h^*_{22} \sqrt{\frac{\alpha_1 P_c}{P_p}}) P_p \\ h_{22} \sigma^2_{x_c} + \alpha_2 (h_{21} + h_{22} \sqrt{\frac{\alpha_1 P_c}{P_p}}) P_p & |h_{22}|^2 \sigma^2_{x_c} + |h_{21} + h_{22} \sqrt{\frac{\alpha_1 P_c}{P_p}}|^2 P_p + \sigma^2_{Z_s} \end{pmatrix} \] (11)

Then the mutual information with given channel realizations can be computed as follows

\[ R(h_{22}, h_{21}) \triangleq I(U; Y_s|H_{22} = h_{22}, H_{21} = h_{21}) - I(U; S) \]

\[ = h(Y_s) - h(U, Y_s) - h(S) + h(U, S) \]

\[ = \log(\Sigma_{U,Y_s}(2, 2)) - \log(|\Sigma_{U,Y_s}|) - \log(\Sigma_{U,S}(2, 2)) + \log(|\Sigma_{U,S}|) \]

\[ = \log \left( (|h_{22}|^2 \sigma^2_{x_c} + |h_{21} + h_{22} \sqrt{\frac{\alpha_1 P_c}{P_p}}|^2 P_p + \sigma^2_{Z_s}) \sigma^2_{x_c} \right) - \log(|\Sigma_{U,Y_s}|), \] (9)

where

\[ \Sigma_{U,S} = \begin{pmatrix} \sigma^2_{x_c} + |\alpha_2|^2 P_p & \alpha_2 P_p \\ \alpha_2 P_p & P_p \end{pmatrix}, \]

\[ \sigma^2_{x_c} = (1 - \alpha_1)P_c, \] and \( \Sigma_{U,Y_s} \) is shown in (11). Similarly, we define the primary user’s rate in the existence of CR user with certain channel realizations as

\[ R(h_{11}, h_{12}) \triangleq I(X_p; Y_p, H_{11} = h_{11}, H_{12} = h_{12}) \]

\[ = \log \left( (1 - \alpha_1)P_c|h_{12}|^2 + \sqrt{P_p} h_{11} + \sqrt{\alpha_1 P_c} h_{12} |^2 + \sigma^2_{Z_p} \right). \] (10)

**III. Fast fading scenario**

In Viswanath’s setting [5] the CR TX must ensure the primary user’s rate unchanged. The definition of the rate varies depending on the fading effect. For fast fading channels, a meaningful definition of the rate is the *ergodic capacity*. Therefore, we take the liner-assignment achievable rate (6) averaged over all channel realizations [12, P.533] to form the following design criteria for this scenario

\[ \text{maximize } E[R(H_{22}, H_{21})] \] (12)

subject to \( E[R(H_{11}, H_{12})] \geq E[\log(1 + \frac{|H_{11}|^2 P_p}{\sigma^2_{Z_p}})]. \) (13)
A. Derivation of the relaying ratio $\alpha_1$

Since the constraint (13) is hard to solve analytically, we use the known statistics of the distributions of the channels to generate a set of realization of $H_{11}$ and $H_{12}$. $\alpha_1$ is then found by computer search that ensures

$$
\frac{1}{N} \sum_{k=1}^{N} \log \left(1 - \alpha_1 \right) P_c |h_{11}^{(k)}|^2 + \sqrt{\frac{P_p}{P_c} h_{11}^{(k)}} + \sqrt{\alpha_1 P_c h_{12}^{(k)}}^2 + \sigma_{Z_p}^2 = \frac{1}{N} \sum_{k=1}^{N} \log \left(1 + \frac{|h_{11}^{(k)}|^2}{\sigma_{Z_p}^2} \right)
$$

(14)

where $h_{12}^{(k)}$ and $h_{11}^{(k)}$ are the $k$th realizations of random variables $H_{12}$ and $H_{11}$ respectively and $N$ is the number of realization generated.

B. Derivation of the precoding coefficient $\alpha_2$

First note that only the second term in $R(h_{22},h_{21})$ is related to $\alpha_2$. Thus to derive $\alpha_2$ for maximizing $E[R(H_{22},H_{21})]$, we may use the following approximation

$$
\alpha_2 = \arg\min_{\alpha_2} E[|\Sigma_{U,Y_s}|].
$$

(15)

Since $E[|\Sigma_{U,Y_s}|]$ is a quadratic function of $\alpha_2$, we can get the optimal $\alpha_2$ by solving

$$
\frac{\partial E[|\Sigma_{U,Y_s}|]}{\partial \alpha_2} = (\sigma_{22}^2 + |\mu_{22}|^2)(1 - \alpha_1) P_c + \sigma_{Z_p}^2) \alpha_2 - \left(\mu_{22}^* \mu_{21} + (\sigma_{22}^2 + |\mu_{22}|^2) \sqrt{\frac{\alpha_1 P_c}{P_c}} (1 - \alpha_1) P_c \right) = 0.
$$

Then the optimal precoding coefficient under partial CSI is

$$
\alpha_2 = \frac{(\mu_{22}^* \mu_{21} + (\sigma_{22}^2 + |\mu_{22}|^2) \sqrt{\frac{\alpha_1 P_c}{P_c}} (1 - \alpha_1) P_c \right)}{(\sigma_{22}^2 + |\mu_{22}|^2)(1 - \alpha_1) P_c + \sigma_{Z_p}^2}.
$$

(16)

IV. Slow fading scenario

When CR systems operate under this scenario, typically an outage probability $P_{\text{out}}^{\text{CR}}$ is given, and it is desired to maximize the rate $R^{\text{CR}}$ such that this outage probability is satisfied, i.e., $P(R(H_{21},H_{22}) < R^{\text{CR}}) = P_{\text{out}}^{\text{CR}}$. Since CR systems must ensure that the primary user’s outage capacity does not decrease, we have the following formulation

$$
\text{maximize } R^{\text{CR}}
$$

subject to $R^{\text{primary}} \leq F^{-1}(1 - P_{\text{out}}^{\text{primary}})$,
where \( P_{\text{out}}^{\text{primary}} \) and \( R^{\text{primary}} \) are the outage probability and outage capacity, respectively, of the primary user in the absence of CR TX, and \( F \) is the complementary cumulative density function (CDF) of \( R(H_{11}, H_{12}) \).

### A. Derivation of the relaying ratio \( \alpha_1 \)

Assume the primary user’s target outage probability is \( P_{\text{out}}^{\text{primary}} \) and the channel statistics are known at CR TX. Then CR TX can derive the outage capacity \( R^{\text{primary}} \) of the primary user as following. We know that the outage capacity of the primary user in the absence of CR TX is defined as

\[
P_{\text{out}}^{\text{primary}} \triangleq P(\log(1 + |H_{11}|^2 \frac{P_p}{\sigma^2_z}) \leq R^{\text{primary}})
\]

\[
= P(|H_{11}|^2 \leq (2^{R^{\text{primary}} - 1}) \frac{\sigma^2_z P_p}{P_p}).
\]

\( R^{\text{primary}} \) can then be found by searching the CDF of \( |H_{11}|^2 \). When the CR user emerges, the rate function of primary user is replaced by \((10)\). Since this is a function of two random variables, it is hard to have an analytical form of \( \alpha_1 \). Thus we find \( \alpha_1 \) by computer search as following

\[
\alpha_1 : P(R(H_{11}, H_{12}) \leq R^{\text{primary}}) = P_{\text{out}}^{\text{primary}}.
\]

### B. Derivation of the precoding coefficient \( \alpha_2 \)

It is hard to solve \( \alpha_2 \) from \((17)\) for the CR user analytically and we may use the criteria in Section [III-B] instead under high Rician \( K \)-factors, that is, choosing \( \alpha_2 \) which maximizes \( E[R(H_{22}, H_{21})] \). This is because when the \( K \)-factor of the channel is large, the channel gains are almost fixed and knowing the channel statistics is almost the same as knowing the channel gains. Thus the rate with precoding coefficient \( \alpha_2 \) in \((15)\) may coincide with that with full CSI under large \( K \)-factor. On the otherhand, when the \( K \)-factor is small, channel gains vary dramatically and the rate \( R(H_{22}, H_{21}) \) may widely spread. Thus maximizing the mean value of \( R(H_{22}, H_{21}) \) may not lead to maximized outage capacity. This phenomenon is shown in Fig. 2 where (a) and (b) are the rate distributions of treating interference as noise (IaN) and precoded CR with \( \alpha_2 \) from \((16)\) when \( P_c = P_p = 10 \) and \( K=2 \text{dB} \). The rate of treating interference as noise \( R_{\text{IaN}} \) is as following

\[
R_{\text{IaN}} = \log \left( \frac{|h_{22}|^2 P_c + |h_{21}|^2 P_p + \sigma^2_z}{|h_{22}|^2 \alpha_1 P_c + |h_{21}|^2 P_p + \sigma^2_z} \right).
\]
The horizontal and vertical axes are the rate and frequency of occurrence, respectively. It can easily be seen that $R_{IaN}$ has a much smaller mean value than that of $R(h_{22}, h_{21})$. Note that the conditional rate $R(h_{22}, h_{21})$ is negative for a certain range of $\{h_{22}, h_{21}\}$ as calculated from (9). This implies that for that range of $\{h_{22}, h_{21}\}$, the LA-GPC with the chosen $\alpha_2$ can not generate nonnegative achievable rates. Similar phenomenon was also observed in [8] which applied LA-GPC to the broadcast channel. According to [11], whenever $R(h_{22}, h_{21}) < 0$, it loses no generality to set $R(h_{22}, h_{21}) = 0$. However, if $P_{out}^{CR}$ is not large enough, the outage capacity of CR user will be zero. On the contrary, (19) is always positive and thus implying a positive outage capacity. Thus only maximizing the mean of rate given $P_{out}$ may result in a worse performance than simply treating interference as noise.

In order to find good precoding coefficient $\alpha_2$ for all $K$-factors, we propose two methods which improve the results in the following subsections.

1) Gaussian approximation: From numerical experiments we found that under low $K$-factor the rate distribution is approximately Gaussian. Assuming that the rate $R(H_{22}, H_{21}) \sim N(\mu_R(\alpha_2), \sigma_R^2(\alpha_2))$, for any $R_{out}$ there exists an $a$ such that

$$R_{out} = \mu_R(\alpha_2) - a\sigma_R(\alpha_2), \quad (20)$$

and

$$P_{out} \triangleq P(R(H_{22}, H_{21}) < R_{out})$$

$$= P(R(H_{22}, H_{21}) < \mu_R - a\sigma_R)$$

$$\approx \Phi(-a) \quad (21)$$

where $\Phi$ is the CDF of the normal distribution. Since (21) is independent of $\mu_R$ and $\sigma_R$, and $\Phi$ is monotonically increasing, when $P_{out}$ is given, $a$ can be obtained by

$$a \approx \Phi^{-1}(1 - P_{out}). \quad (22)$$

To maximize $R_{out}$ given $P_{out}$, the following approximate objective function can be used

$$\text{maximize } \mu_R(\alpha_2) - a\sigma_R(\alpha_2). \quad (23)$$

The selection of $\alpha_2$ which maximizes (23) is also good for moderate and high $K$-factors, as shown in our simulation results.
2) **Joint optimization of mean and variance:** We may also consider the mean and variance of \( R(\mathcal{H}_{22}, \mathcal{H}_{21}) \) simultaneously. Denote the mean and variance of \( R(\mathcal{H}_{22}, \mathcal{H}_{21}) \) by \( \mu_R(\alpha_2) \) and \( \sigma_R^2(\alpha_2) \), respectively, for a certain \( \alpha_2 \). Heuristically we would like to maximize \( \mu_R(\alpha_2) \) and minimize \( \sigma_R^2(\alpha_2) \) (or maximize \( -\sigma_R^2(\alpha_2) \)) simultaneously as the following multi-objective optimization problem,

\[
\alpha_2 = \arg \max_{\alpha_2'} \{ \mu_R(\alpha_2'), -\sigma_R^2(\alpha_2') \}. \tag{24}
\]

Using the additional minimization of the variance of the rate, we may avoid the case of large mean value with small outage capacity caused by large variance. By standard multi-objective optimization theory we can recast (24) as a single objective optimization problem with the *aggregating function method* [13], that is

\[
\alpha_2 = \arg \max_{\alpha_2'} \mu_R(\alpha_2') - a\sigma_R^2(\alpha_2'), \tag{25}
\]

where the parameter \( a > 0 \). In the following we discuss how to determine \( a \) and find \( \alpha_2 \) respectively. Simultaneously choosing the pair \( \{\alpha_2, a\} \) to maximize (25) does not guarantee the maximization of \( R^{CR} \) since there is no explicit relation between \( \mu_R(\alpha_2) - a\sigma_R^2(\alpha_2) \) and \( R^{CR} \). Both (23) and (25) are approximations to (24) since \( \sigma_R \) is neither convex nor concave.\(^\dagger\) And (23) can be seen as the following problem after applying the aggregation method which is similar to (24): \( \alpha_2 = \arg \max_{\alpha_2'} \{ \mu_R(\alpha_2'), -\sigma_R(\alpha_2') \} \). However, (23) and (25) have different objective functions of \( \alpha_{28} \) and result in different optimal \( \alpha_{28} \)s.

The above two methods have their own advantages. The use of (22) as a suboptimal solution of \( a \) can reduce large computational complexity. Meanwhile, its performance approximates that of the computer full search. On the contrary, under low \( K \)-factor region, \( \alpha_2 \) from (25) has an approximately analytical form which can be solved very efficiently [9]. We can combine these two methods to have both advantages while having approximately the same performance as that of optimal LA-GPC.

\(^\dagger\)If \( \max_{\alpha_2} \{ \mu_R(\alpha_2), -\sigma_R(\alpha_2) \} \) is a convex problem, then for any Pareto optimal solution \( \alpha_{28}^* \) for it, there exists a weighting vector \( \mathbf{a} \) \((a_i \geq 0, i = 1, 2, a_1 + a_2 = 1) \) such that \( \alpha_{28}^* \) is a solution of the problem: \( \max_{\alpha_2'} a_1\mu_R(\alpha_2) - a_2\sigma_R(\alpha_2) \). Please refer to [14, Theorem 3.1.4] for the details.
C. Discussion

As indicated in [8], the outage capacity of BC (with $H_x = H_s = H$ in (2)) with the statistics of CSI at the TX is the same as that of the cases with no interference or interference known at the RX. Following is an intuitive explanation for this result. Denote the linear assignment achievable rate function when $H_x = H_s = H$ by $R(H)$ ((15) in [8]). For a given outage capacity $R_{out}$, let $h_{IF}$ be the corresponding channel gain of the interference free scenario obtained by

$$R_{out} = \log(1 + |h_{IF}|^2SNR).$$

For the case where only the statistics of CSI are known at the TX, let us obtain $\alpha_2$ assuming the channel gain is always $h_{IF}$. With this $\alpha_2$, when the channel gain is really $h_{IF}$, we will have $R(h_{IF}) = R_{out}$. In addition, since $R(H)$ is a non-decreasing function of $|H|$, we can easily find that

$$P(R(H) < R_{out}) = P(R(H) < R(h_{IF})) = P(|H| < |h_{IF}|),$$

where $P(|H| < |h_{IF}|)$ is the interference free outage probability. Therefore, the outage probability of an interference channel is the same as that of the interference free channel if the achievable rate function is non-decreasing of the absolute value of the channel gain.

Unfortunately, the same procedure can not be applied to (7) and the outage capacity of our CR channel setting is still an open problem. As the simulation results in Section VI will show, for the CR channel, there is an apparent rate loss at low and moderate $K$-factor regions if only the statistics of CSI are known at the TX.

V. NESTED-LATTICE BASED LINEAR-ASSIGNMENT GEL’FAND-PINSKER CODING

We will introduce a nested-lattice based coding structure for the linear-assignment Gel’fand-Pinsker coding. Since lattice code has an algebraic structure, we can bring the previous information-theoretic result into practice. To illustrate this coding, we focus on the following channel corresponding to (2)

$$y_t = H_x(t)x_t + H_s(t)s_t + z_t,$$

where $t$ is the discrete time index and $1 \leq t \leq T$, $T$ is the number of symbols in a code block. To emphasize the differences between the unstructured Shannon random codebook setting and the lattice codebook
setting, signal variables in the former are denoted in italic capitals as in previous sections while the corresponding ones in the latter are denoted in bold-face lower-cases, such as $X$ and $x$, respectively. Note that in the previous sections, we neglect the time index $t$ for simplicity. Focusing on the CR application, the channel gains and signals are specified as in Section II-C. The dirty-paper channel input is limited by an input power constraint $(1 - \alpha_1)P_c$. Note that the CR TX also uses $\alpha_1 P_c$ to relay the primary user’s signal, thus the total CR Tx power is $P_c$.

We can rewrite (27) in an equivalent real super channel to present the lattice coding more easily. By concatenating all $T$ symbols, (27) becomes

$$y = H_x x + H_s s + z,$$

where $x = (x_1^T, \ldots, x_T^T)^T$ and $x_t^T = [\text{Re}\{x_t\}, \text{Im}\{x_t\}]$. The noncasually known side-information at the transmitter $s$ and the noise term $z$ are obtained similarly from $s_t$ and $z_t$ respectively as $x$ from $x_t$. The covariance matrix of $z$ is denoted by $\Sigma_z$ and it is assumed to be $\sigma_z^2 I_{NT}$. The $2T \times 2T$ block-diagonal real channel matrix $H_x$ is

$$
\begin{pmatrix}
H_x(1) & 0_{2 \times 2} & \cdots & 0_{2 \times 2} \\
0_{2 \times 2} & H_x(2) & \cdots & 0_{2 \times 2} \\
\vdots & \vdots & \ddots & \vdots \\
0_{2 \times 2} & 0_{2 \times 2} & \cdots & H_x(T)
\end{pmatrix},
$$

where the $t$th diagonal term $H_x(t)$ is

$$
\begin{bmatrix}
\text{Re}\{H_x(t)\} & -\text{Im}\{H_x(t)\} \\
\text{Im}\{H_x(t)\} & \text{Re}\{H_x(t)\}
\end{bmatrix}.
$$

$H_s$ is formed from $H_x(t)$ in the same way as $H_x$ from $H_x(t)$. The channel input power constraint is $(1 - \alpha_1)P_c/2$ per real dimension. We will first give a brief review of nested-lattice codebook then introduce our proposed lattice coding in subsection V-A and V-B, respectively.

A. Review of lattices and lattice quantization noise

An $m_L$-dimensional real lattice $\Lambda$ is defined as $\Lambda = \{Gb : b \in \mathbb{Z}^{m_L}\}$, where $G$ is the $m_L \times m_L$ generator matrix of $\Lambda$. The Voronoi region $\nu$ is the set of points $g \in \mathbb{R}^{m_L}$ which are closest to $0$ in Euclidean distance.
than to other lattice points \( \lambda \in \Lambda \). Every \( g \in \mathbb{R}^{m_L} \) can be uniquely written as \( g = \lambda + n_g \), where \( \lambda \in \Lambda \) and \( n_g \in \nu \). With quantizer input \( g \), the lattice quantizer associated with \( \nu \) is defined as \( Q(g) = \lambda \), if \( g \in \lambda + \nu \).

The modulo-\( \Lambda \) operation associated with \( \nu \) is then

\[
g \mod \Lambda = g - Q(g).
\]

(29)

Let \( u \) be a dither uniformly distributed in \( \nu \) and independent of \( g \), it is proved in [15, Lemma 1] that the dithered quantization error \((g + u) \mod \Lambda\) is also uniformly distributed in \( \nu \) as \( u \), and independent of \( g \).

We define the nested lattice codes as

**Definition 1:** Let \( \Lambda_c \) be a lattice and \( \Lambda_q \) be a sublattice of it, that is, \( \Lambda_q \subseteq \Lambda_c \). The codeword set of the nested lattice code is

\[
c_c = \{ \Lambda_c \mod \Lambda_q \} \triangleq \{ \Lambda_c \cap \nu_q \}.
\]

And the code rate of this nested lattice code is

\[
R_c = \frac{1}{T} \log \frac{||\nu_q||}{||\nu_c||},
\]

where \( \nu_c \) is the fundamental Voronoi region of \( \Lambda_c \). A sequence of lattices \( \{\Lambda_q^T\} \) of increasing dimension \( T \) is good for covering [15] if the covering efficiency

\[
\eta_{cov}(\Lambda_q^T) \triangleq \frac{R_c^c(\Lambda_q^T)}{R_c^c(\Lambda_q^T)} \to 1,
\]

(30)

where the covering radius \( R_c^c(\Lambda_q^T) \) is the radius of the smallest sphere centered at the origin that contains the Voronoi region \( \nu \), and the effective radius \( R_c^e(\Lambda_q^T) \) is the radius of the sphere with volume equal to \( ||\nu|| \). The goodness for covering implies goodness for MSE quantization [16]. We also need the following good nested lattice sequences [15]

**Lemma 1 (Erez, Zamir):** There exist sequences of nested lattices \( \Lambda_q^T \subseteq \Lambda_c^T \) of dimension \( 2T \) such that

1) The sequence of coarse lattices \( \{\Lambda_q^T\} \) is a sequence of lattices that are good for covering and thus good for MSE quantization.

2) For each \( 2T \), \( \Lambda_c^T \) is randomly selected in an ensemble that asymptotically satisfies the Minkowski-Hlawka theorem in the form given in [17], in the limit of \( T \to \infty \).
B. Nested-lattice based linear-assignment Gel’fand-Pinsker coding

As in Section II-C, without loss of generality, to achieve the outage performance (7), we focus on the achievable rate of certain channel realizations with \( H_x = H_x \) and \( H_s = H_s \). The channel becomes

\[
y = H_x x + H_s s + z, \tag{31}
\]

and we focus on the achievability of \( R(h_{22}, h_{21}) \) in (9). The encoder works as follows

**Encoder:** The encoder selects a codeword \( c \in c_c \) according to the message index and sends

\[
x = \sqrt{(1 - \alpha_1) P_c ((c - Ws - u) \mod \Lambda_q)}, \tag{32}
\]

where the dither signal \( u \), uniformly distributed in \( V \) and independent of the channel, is known to both the transmitter and receiver. This dither plays a critical role to make \( x \) independent of \( c \) and \( s \) [18], as introduced in subsection V-A. We assume that the second moment \( P(\nu_q) \) of \( \Lambda_q \) equals \( 1/2 \) to satisfy the power constraint. The \( 2T \times 2T \) block-diagonal matrix \( W \) is formed from \( \alpha_2 \) as

\[
I_T \otimes \begin{bmatrix}
\text{Re}\{\alpha_2\} & -\text{Im}\{\alpha_2\} \\
\text{Im}\{\alpha_2\} & \text{Re}\{\alpha_2\}
\end{bmatrix},
\tag{33}
\]

where \( \otimes \) denotes the Kronecker product. The signal \( x \) is then transmitted through the channel (31) with channel output \( y \).

**Decoder:** The decoder performs signal processing on the received signal and gets

\[
\hat{y} = L(W_r y + u), \tag{34}
\]

where the matrix filter \( W_r \) is

\[
\left(\frac{1}{2} \tilde{H}^T + W \Sigma_x H_s^T \right) \left(\frac{1}{2} \tilde{H} \tilde{H}^T + H_s \Sigma_x H_s^T + \Sigma_z \right)^{-1}, \tag{35}
\]

where \( \tilde{H} = \sqrt{(1 - \alpha_1) P_c H_x} \), \( \Sigma_x \) and \( \Sigma_z \) are the covariance matrix of \( s \) and \( z \), respectively. And the matrix filter \( L \) must satisfy

\[
L^T L = \Sigma_E^{-1}, \tag{36}
\]
where
\[
\Sigma_E = \frac{1}{2}I_{mT} + W\Sigma_s W^T \left( \frac{1}{2} \tilde{H}^T + W\Sigma_s H_s^T \right) \left( \frac{1}{2} \tilde{H}\tilde{H}^T + H_s\Sigma_s H_s^T + \Sigma_z \right)^{-1} \left( \frac{1}{2} \tilde{H} + H_s\Sigma_s W^T \right).
\]

From [18], \( \hat{y} \) in (34) can be rewritten as
\[
\hat{y} = L(c'_c) + e',
\]
where \( e' \) will approach white Gaussian noise as \( T \to \infty \), and \( c'_c \in \Lambda_q + c_c \subset \Lambda_c \). Thus we can use the generalized minimum Euclidean distance lattice decoder [17] to decode \( c_c \). First the decoder finds
\[
\hat{b} = \arg\min_{b \in \mathbb{Z}^{2T}} |\hat{y} - LG_c b|^2,
\]
where \( G_c \) is the generator matrix of the channel coding lattice \( \Lambda_c \). And the decoded codeword is
\[
\hat{c}_c = [G_c\hat{b}] \mod \Lambda_q.
\]

According to [18], we have the following result.

**Theorem 1:** With selected filters \( W, W_r \), and \( L \) as (33), (35), and (36), respectively, and based on sequences of good nested lattices, the nested lattice coding specified in (32)-(38) is able to achieve the linear-assignment rate \( R(h_{22}, h_{21}) \) under power constraint \( (1 - \alpha_1)P_c \) when \( T \to \infty \).

The proof can be easily derived from [18] and is omitted here. To achieve the rate performance specified in Theorem 1, good nested-lattice tailored for very long codeword length (that is \( T \to \infty \)) are needed. In our simulation, we will show that good performance can still be obtained at large enough SNR with reasonable codeword length (and decoding latency).

**VI. Simulation results**

In this section we demonstrate the performance of the LA-GPC with the proposed parameter design under both fast and slow fading channels. We also show the performance of the proposed lattice encoder/decoder which implements the proposed precoding.
A. Fast fading

Fig. 3 shows the simulation results of the fast fading scenario. CR channels with full CSI and without CSI are used as the upper and lower bounds of the performance. In the former case it is interference free (IF) for CR users since original DPC in [4] can eliminate the interference at CR RX. In the latter case CR users do not use any precoding and treat primary users’ signals as interferes. Meanwhile, CR TX still uses partial power to relay the primary signal to ensure primary user’s rate performance unchanged. The rate formula of the lower bound is shown in (19) and that of the upper bound for given channel realizations is

\[ R_{IF} = \log\left(1 + |h_{22}|^2 \sigma_{\hat{x}}^2 / \sigma_{z_t}^2\right). \]  

(39)

It can be found that the ergodic capacity with \( \alpha_2 \) derived in (16) is almost the same as the optimal one of the LA-GPC for all \( K \)-factor. The optimal LA means the value of \( \alpha_2 \) is found by computer search which maximizes the ergodic capacity. However, large rate loss of LA-GPC happens in the low \( K \)-factor region compared to the full CSI case. Treating interference as noise incurs dramatic rate loss for all \( K \)-factor. Rate loss in fast fading channels caused by partial CSI at the TX also happens in the BC case [8].

B. Slow fading

In this scenario, we consider two kinds of outage probabilities: \( P_{out}^{primary} = P_{CR}^{out} = 0.1 \) and 0.01. For simplicity the \( K \)-factors of the four channels in Fig. 1 are set as the same. For comparison, we also show the cases with full CSI (no interference from primary user) and treating interference as noise. The rate \( R_{CR} \) satisfying \( P(R(H_{21}, H_{22}) < R_{CR}^{out}) = P_{out}^{CR} \) versus different \( K \)-factors is used as the performance index of the slow fading channel. We also use computer search to find the optimal \( \alpha_2 \) which results in the highest outage capacity using LA-GPC. Comparing the computer search result to (39) we can find the rate loss caused by the fading channel with LA-GPC. The computer search result can also serve as an performance upper-bound of the proposed method (25) for finding \( \alpha_2 \). The results with small transmission power are shown in Fig. 4 and Fig. 5 while the high transmission power results are shown in Fig. 6 and Fig. 7 respectively. From these figures we can see that at low \( K \)-factor region, curves of \( \alpha_2 \) from (16) and treating interference as noise intersect. This is because (19) is non-negative and thus the outage capacity of treating
interference as noise is always nonnegative. On the other hand, the outage capacity with $\alpha_2$ derived from (16) may be negative when the $K$-factor is small. As the $K$-factor increases, the performance using $\alpha_2$ from (16) is improved as discussed in Section IV-B. Note that for $K$-factor $\to \infty$ the value of channel gain is the same as the mean value and the histogram of the rate becomes a delta function. Then the outage capacity is the same as the ergodic capacity and independent of the outage probability. Therefore Fig. 4 and Fig. 5 should converge to the same rate (about 1.8 bits per channel use) as the $K$-factor $\to \infty$. And the performance of the proposed criterion (25) is very close to the one obtained from computer searching optimal $\alpha_2$ for all $K$-factor. From simulation we also find that the optimal value of $a$ in (25) increases with decreasing outage probability where the values of $a$ in the considered condition range from 0 to 4.8. This is reasonable because as $P_{out}$ becomes smaller the corresponding $R_{out}$ is closer to the left tail of the rate distribution. Thus to have higher outage capacity it is more important to reduce the variance than increasing the mean of the rate when the outage probability is smaller. From simulation results we can also find that both the two objectives (23) and (25) with $a$ derived from Gaussian approximation (G. A.) have similar performances. And the losses from that of optimal $\alpha_2$ are also negligible.

C. Lattice implementation

As stated in Section V, we examine the error performance at large enough SNR with reasonable codeword length (and decoding latency) in this section. To generate the lattice, as in [17], we use the pair of self-similar nested lattices. The Gosset lattice $E_8$ [19] which has the densest packing in 8-dimension is used as the fine lattice $\Lambda_c$. The reason to use the Gosset lattice $E_8$ is that in moderate dimensions, it is well-known that the best lattices with respect to coding gain are also good quantizers, i.e., have good shaping gains. To have a good shaping gain, the transmitted signal must have a Gaussian like distribution. The Gaussianity of signals at the CR TX output is shown in Fig. 8. The green and blue lines are the CDFs of the Gaussian distribution and the TX output. We can easily find that the TX output distribution generated by the adopted code is almost Gaussian although there are small deviation around the tail. Thus it verifies that we have good shaping gain. Lattices with longer length will have better shaping gain but the decoding latency will increase. The coarse lattice or the shaping lattice $\Lambda_q$ is generated by
\[ \Lambda_q = Q \Lambda_c \] where \( Q \) is the nesting ratio and the coding rate is given by \( R = 2 \log Q \). The lattice codeword length is \( T = 8 \). A Fano sequential-decoding based lattice decoder [20] is used to solve (38) with a good performance.

In the simulation we consider two cases of \( K \)-factors: 0dB and 10dB and two code rates: 2 and 4 bits per channel use. The noise variance is normalized as 1 and the primary user power constraint is assumed as \( P_p = 100 \). We compare the lattice coding results with the theoretical linear-assignment Gel’fand-Pinsker coding results (with random Gaussian codebooks). The performance index considered here is the codeword error probability which is approximately the same as the outage probability. The results are shown in Fig. 9 and Fig. 10. For comparison, we also show the cases with full CSI (no interference from primary user) and treating interference as Gaussian noise. When the information of the primary user is not used and the interference is treated as noise, the frame error rate is 1 despite the SNR and the \( K \)-factor. It can be found that in the considered \( K \)-factor region, the performance of the used nested lattice coding/encoding is approximately the same as that derived theoretically. This also verifies the goodness of the used Gosset lattice code. Note that there is a gap between the theoretical Gaussian coding results with partial CSI and the full CSI case. However, the capacity of the partial CSI case is still an open problem. We also observe that in Fig. 9 under \( K \)-factor=0dB, precoding with partial CSI by lattice coding performs slightly better than the theoretical outage case. A similar phenomenon was also reported in [21].

VII. CONCLUSION

In this paper we considered the cognitive radio channel with partial channel side information at the transmitter. Using the linear-assignment Gel’fand-Pinsker coding, we proposed methods for finding the relaying ratios and the precoding coefficients for both fast and slow fading channels. We also used nested-lattice coding and decoding to implement our coding in practice. Simulation results showed that the proposed method with reduced-complexity computer search has performance approaching the optimal linear assignment rates.
REFERENCES

[1] J. Mitola, “Cognitive radio: An integrated agent architecture for software defined radio,” Ph.D. dissertation, KTH Royal Inst. Technology, Stockholm, Sweden, 2000.

[2] N. Devroye, P. Mitran, M. Sharif, S. Ghassemzadeh, and V. Tarokh, “Information Theoretic Analysis of Cognitive Radio Systems” in Cognitive Wireless Communications. V. Bhargava and E. Hossain Ed., Springer, 2007.

[3] N. Devroye, P. Mitran, and V. Tarokh, “Achievable rates in cognitive radio channels,” IEEE Trans. Inform. Theory, vol. 52, no. 5, pp. 1813–1827, May 2006.

[4] M. H. M. Costa, “Writing on dirty paper,” IEEE Trans. Inform. Theory, vol. 29, pp. 439–441, May 1983.

[5] A. Jovicic and P. Viswanath, “Cognitive radio: An information-theoretic perspective,” submitted to IEEE Trans. Inform. Theory, Apr. 2006.

[6] N. Devroye and V. Tarokh, “Fundamental Limits of Cognitive Radio Networks” in Cognitive Wireless Networks: Concepts, Methodologies and Vision. Springer, 2007.

[7] A. Bennatan and D. Burshtein, “On the fading paper achievable region of the fading MIMO broadcast channel,” in 44th Annual Allerton Conference on Communication, Control and Computing, Sep. 2006.

[8] W. Zhang, S. Kotagiri, and J. N. Laneman, “Writing on dirty paper with resizing and its application to quasi-static fading broadcast channels,” in IEEE Int. Symp. Information Theory (ISIT), Nice, France, Jun. 2007.

[9] P.-H. Lin, S.-C. Lin, and H.-J. Su, “Cognitive radio with partial channel state information at the transmitter,” IEEE International Conference on Communications, May 2008.

[10] P.-H. Lin, C.-P. Lee, S.-C. Lin, and H.-J. Su, “Cognitive radio in slow fading channels with partial channel state information at the transmitter,” Annual Conference on Information Sciences and Systems, Mar 2008.

[11] S. I. Gel’fand and M. S. Pinsker, “Coding for channels with random parameters,” Probl. Contr. and Inform. Theory, vol. 9, no. 1, pp. 19–31, 1980.

[12] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge University Press, 2005.

[13] A. Abraham, L. Jain, and R. Goldberg, Evolutionary Multiobjective Optimization: Theoretical Advances and Applications. Springer-Verlag London Limited, 2005.

[14] K. M. Miettinen, Nonlinear multiobjective optimization. Norwell, MA: Kluwer Academic Publishers, 1999.

[15] U. Erez and R. Zamir, “Achieving 1/2 log (1+SNR) on the AWGN channel with lattice encoding and decoding,” IEEE Trans. Inform. Theory, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[16] R. Zamir and M. Feder, “On lattice quantization noise,” IEEE Trans. Inform. Theory, vol. 42, no. 4, pp. 1152–1159, July. 1996.

[17] H. E. Gamal, G. Caire, and M. O. Damen, “Lattice coding and decoding achieve the optimal diversity-vs-multiplexing tradeoff of MIMO channels,” IEEE Trans. Inform. Theory, vol. 50, no. 6, pp. 968 – 985, Jun 2004.

[18] S.-C. Lin, P.-H. Lin, and H.-J. Su, “Lattice coding for the vector fading paper problem,” in IEEE Information Theory Workshop, Lake Tahoe, California, USA, Sep. 2-Sep. 6 2007.

[19] J. H. Conway and N. J. A. Sloane, Sphere packings, lattices, and groups. New York: Springer-Verlag, 1988.
[20] A. D. Murugan, H. El-Gamal, M. O. Damen, and G. Caire, “A unified framework for tree search decoding: rediscovering the sequential decoder,” *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 933–953, 2006.

[21] K. R. Kumar and G. Caire, “Construction of Structured LaST Codes,” in *IEEE International Symposium on Information Theory (ISIT)*, 2006, pp. 2834–2838.
THE MODEL OF THE COGNITIVE RADIO CHANNEL.
Fig. 2

Histograms of CR rate for (a) treating interferences as noise and (b) precoded CR using criterion (15).

Fig. 3

Comparison of the ergodic capacities with $P_c = P_p = 10$. 
Comparison of the outage capacities of different methods with $P_{out} = 0.1$ and $P_c = P_p = 10$.

Fig. 4

Comparison of the outage capacities of different methods with $P_{out} = 0.01$ and $P_c = P_p = 10$.

Fig. 5
**Fig. 6**

**Comparison of the Outage Capacities of Different Methods with $P_{out} = 0.1$ and $P_c = P_p = 100$.**

**Fig. 7**

**Comparison of the Outage Capacities of Different Methods with $P_{out} = 0.01$ and $P_c = P_p = 100$.**
The Gaussianity of the transmitted signal using nested lattice code.

Comparison of the outage probabilities with rate = 2 bits/channel use.
Fig. 10

Comparison of the outage probabilities with rate = 4 bits/channel use.