Approximate Capacity of the General Gaussian Parallel Relay Network

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

NICOLAS SCHRAMMAR, MATTIAS ANDERSSON, AND MIKAEL SKOGLUND

Stockholm 2011

Communication Theory Department
School of Electrical Engineering
KTH Royal Institute of Technology

IR-EE-KT 2011:028
Approximate Capacity of the General Gaussian Parallel Relay Network

Nicolas Schrammar, Mattias Andersson, and Mikael Skoglund
School of Electrical Engineering and ACCESS Linnaeus Center
KTH Royal Institute of Technology, Stockholm, Sweden
Email: {nisc, amattias, skoglund}@kth.se

Abstract—We approximate the capacity of the Gaussian parallel relay network with general channel gains. Our strategy is to find capacity approximations for the corresponding network in the discrete superposition model and to use the fact that those are an approximation for the Gaussian capacity. The gap between our approximation and the Gaussian capacity is a constant depending only on the number of relays, hence it is a valuable characterization for the regime of high SNR and high rate.

I. INTRODUCTION

We consider the parallel relay network (PRN) or K-relay diamond network depicted in Fig. 1. The source s transmits information to the destination d. There exists no direct link between source and destination, but a layer of K relays 1, ..., K can assist in the transmission. The Gaussian PRN has been introduced by Schein and Gallager [1], however the capacity was found only for special cases. Improved achievable rates were found in [2] and [3], and the capacity was found for a special class of the PRN with two relays [4]. However, the capacity for the general Gaussian PRN remains unknown even for the two relay case.

Recently, it has been suggested that capacity approximations can be a valuable characterization for communication networks. In the high SNR regime it is desirable to approximate capacity within a constant additive gap. The approximation becomes tight for increasing SNR and, hence, increasing rate. This type of approximations have been first studied in [5] for Gaussian relay networks, of which the Gaussian PRN is one example. The result is that the capacity can be approximated by the cut-set bound within a constant gap. An approximation for a specific network topology has been found in [6]. The authors consider the class of symmetric Gaussian PRN, in which the channel gains from source to each of the relays are equal, and also the channel gains from each of the relays to the destination are equal. Both additive and multiplicative approximations with gaps independent of K have been found.

In this paper we find the approximate capacity for the Gaussian PRN as considered by Schein and Gallager. In particular, we consider the general case where all channel gains in the network can be different. The approximation depends only on the channel gains, and the additive gap to the capacity is of order K log K. Our approach utilizes the discrete superposition model (DSM) proposed by Anand and Kumar [7], which is a variation of the truncated deterministic model [8]. The capacity in the DSM was shown to be an additive capacity approximation for relay networks and the interference channel [9]. We find an additive approximation for the capacity of the PRN in the DSM by using our previous results on the capacity region of the multiple-access channel in the DSM presented in [10]. Since the PRN is a relay network, our approximation is also an additive approximation for the Gaussian capacity.

The remainder of this paper is organized as follows. After introducing the two channel models in Sec. II we find the approximate capacity region of the broadcast channel in the DSM in Sec. III. Furthermore, we consider the example of the broadcast channel with two receivers. In Sec. IV we present approximations on the capacity of the PRN in the DSM. Our final result, the approximate capacity of the Gaussian PRN, is stated in Th. 1 at the end of the section. We conclude the paper in Sec. V.

II. CHANNEL MODELS

A. Gaussian Channel Model

Fig. 1 depicts the topology of the PRN. We identify a broadcast phase from source s to the relays and a multiple-access phase from the relays to the destination d. In the Gaussian channel model the relay k receives

\[ y_k = h_{sk} x_s + z_k, \quad k = 1, \ldots, K, \]

where \( x_s, h_{sk} \) and \( z_k \) are the signal transmitted by the source, the channel gain from source to relay \( k \) and the additive Gaussian noise at relay \( k \), respectively. All variables are complex valued. We normalize the system according to \( \mathbb{E}(|x_s|^2) \leq 1 \).
and \( z_k \sim CN(0, 1) \), hence the system is completely described by the channel gains.

The multiple-access phase is modeled accordingly,

\[
y_d = \sum_{k=1}^{K} h_{kd} x_k + z_d, \tag{2}
\]

with \( E(x_k^2) \leq 1 \) and \( z_d \sim CN(0, 1) \).

**B. Discrete Superposition Model**

The discrete superposition model (DSM) as introduced in [7] serves as an approximation for the Gaussian model in the sense that the capacity \( C \) of the DSM is within a constant gap to the capacity \( C_G \) of the Gaussian model for some network topologies including the PRN [9].

The DSM comprises two key properties. Firstly, the channel inputs \( x_k, k \in \{d, 1, \ldots, K\} \) can be transmitted to any subset of receivers. Secondly, we consider a subset \( S \subseteq \{1, \ldots, K\} \) that the system is completely described by the channel gains. The length of that square is \( 1/\sqrt{2} \). The square gets scaled and rotated by multiplying \( h_{k} \) and sliced in unit squares by the rounding operation in (4).

\[
|S| \times |S| - \log 9 \text{ bits of information per transmitted symbol.}
\]

For brevity. The slicing grid of the second receiver can slice each of those pieces in at most 6 pieces, because \( |h_k| \leq |h_{\max}| \). Hence, using all \( |S| \) grids simultaneously, the number of pieces \( m \) is less than

\[
m \geq 6^{|S|-1} \cdot m_1.
\]

Plugging into (6) and observing that \( 1 < \log 6 \) yields the result.

**Lemma 1.** The sum capacity of any subset \( S \subseteq \{1, \ldots, K\} \) of receivers in the DSM can be upper bounded as

\[
\sum_{k \in S} C_k \leq 2 \log |h_{\max}| + |S| \log 6,
\]

where \( |h_{\max}| = \max_{k \in S} |h_k| \).

**Proof:** We allow joint processing of the received signals \( y_k, k \in S \); this can only increase the sum capacity. Due to (3), the transmitted signal \( x \) takes values from a square in the complex plane. The length of that square is \( 1/\sqrt{2} \). The square gets scaled and rotated by multiplying \( h_k \) and sliced in unit squares by the rounding operation in (4). Equivalently, the square of the transmitted signal \( x \) gets sliced by a rotated grid consisting of squares of length \( 1/|h_{\max}| \). Each receiver \( k \) provides such a slicing grid. Using all grids simultaneously slices the square of \( x \) into \( m \) pieces. Hence, the joint processing cannot distinguish more than \( m \) different transmit symbols, and therefore

\[
\sum_{k \in S} C_k \leq \log m. \tag{6}
\]

To find the number of pieces \( m \) consider the receiver with strongest channel gain \( h_{\max} \) first. It is easy to see that a grid consisting of squares of length \( 1/|h_{\max}| \) can slice a square of length \( 1/\sqrt{2} \) in at most

\[
m_1 \leq \left( \frac{|h_{\max}|}{\sqrt{2}} \right)^2 + 4 \left( \frac{|h_{\max}|}{\sqrt{2}} \right)^2 + 4 \leq 2 |h_{\max}|^2
\]

pieces. The slicing grid of the second receiver can slice each of those pieces in at most 6 pieces, because \( |h_k| \leq |h_{\max}| \). Hence, using all \( |S| \) grids simultaneously, the number of pieces \( m \) is less than

\[
m \leq 6^{|S|-1} \cdot m_1.
\]

Plugging into (6) and observing that \( 1 < \log 6 \) yields the result.

**Lemma 2.** Consider a subset \( S \subseteq \{1, \ldots, K\} \) of \( m = |S| \) receivers and sort them in the order of increasing channel gains \( |h_1| \leq \cdots \leq |h_m| \). The rates at which the transmitter can simultaneously convey independent information to these receivers can be lower bounded as

\[
R_k \geq 2 \log |h_k| - 2 \log |h_{k-1}| - 7, \quad k = 2, \ldots, m,
\]

\[
R_1 \geq 2 \log |h_1| - 7, \quad k = 1.
\]

**Proof:** The transmission strategy that achieves the above rate tuples relies on grouping transmit symbols into blocks. The idea is that receivers with low channel gain perceive a block as a single symbol. However, receivers with high channel gain can distinguish the symbols within the block. The transmitter can therefore convey different information to the receivers.

Fig. 2 depicts a transmission to three receivers. The transmitter uses the symbol constellation depicted as a dashed box on the left. The first and weakest receiver is supposed to detect which one of the solid blocks was received, hence it receives \( \log 9 \) bits of information per transmitted symbol. The second receiver has a stronger channel gain and can therefore detect details within the solid block magnified in the middle of the figure. The second receiver is supposed to detect which one
of the dash-dotted blocks was transmitted. Along with the log 9 bits intended for the first receiver it receives 2 additional bits. Finally, the third and strongest receiver can detect details within the dash-dotted blocks. It is supposed to detect which one of the transmit symbols (crosses) was transmitted and receives log 9 bits of additional information.

The objective of the scheme is twofold. Firstly, the symbols inside a block should merge at the intended receiver due to the rounding operation. Secondly, the intended receiver should be able to distinguish which block was sent.

To ensure that a block of length $a_k$ merges at receiver $k$, two opposing corner points need to be at distance of at most 1, hence

$$|[h_k]x - [h_k](x + a_k + ja_k)| = |[h_k]|\sqrt{2}a_k \leq |h_k|\sqrt{2a_k} \leq 1.$$  

We choose the block size $a_k = (\sqrt{2} |h_k|)^{-1}$. Additionally, we have to shift the block by $s_k$ so that $[h_k](x + s_k) \in \mathbb{Z} + j\mathbb{Z}$. Consider that

$$|[h_k]x - [[h_k]x]| \leq |[h_k]x - ([h_k]x - 1 - j)| = \sqrt{2},$$

hence a shift of

$$s_k = \frac{\sqrt{2}}{|[h_k]|} \leq \frac{\sqrt{2}}{|h_k| - \sqrt{2}} \leq \frac{2}{|h_k|}$$

is sufficient to ensure merging.

In order to meet the second objective, every two symbols $x_1, x_2$ from two different blocks have to fulfill [10, Eq. (5)]

$$|[h_k]x_1 - [h_k]x_2| \geq 2\sqrt{2}.$$ 

We find

$$|[h_k]|x_1 - x_2| \geq |h_k| - (1 + j)|x_1 - x_2| $$

$$\geq |h_k| - \sqrt{2}|x_1 - x_2| $$

$$\geq \frac{|h_k|}{\sqrt{2}}|x_1 - x_2| \geq 2\sqrt{2},$$

where we have used $|h_k| \geq 5$ in the third inequality. Hence, a distance of $4/|h_k|$ between two blocks ensures that the receiver $k$ can detect which block was sent. Considering the shift $s_k$, we find that the total required gap between two neighboring blocks is

$$b_k = \frac{4}{|h_k|} + s_k \leq \frac{6}{|h_k|}.$$ 

Finally, the number of blocks that can be used for transmitting information to receiver $k$ is

$$m_k = \left(\frac{a_{k-1}}{a_k + b_k}\right)^2 \geq \frac{|h_k|^2}{|h_{k-1}|^2} \left(\frac{1}{1 + 6\sqrt{2}}\right)^2.$$ 

Thus, we can achieve a rate of

$$R_k = \log m_k \geq 2\log |h_k| - 2\log |h_{k-1}| - 7.$$ 

The blocks for the weakest receiver $k = 1$ are fitted into the square of length $a_0 = 1/\sqrt{2}$ due to the definition of the DSM. This yields lower bound for $R_1$ stated above.

**Corollary 1.** The achievable sum rate of information conveyed to the subset of receivers $\mathcal{S} \subseteq \{1, \ldots, K\}$ can be lower bounded as

$$\sum_{k \in \mathcal{S}} R_k \geq 2\log |h_{\text{max}}| - 7|\mathcal{S}|,$$

where $|h_{\text{max}}| = \max_{k \in \mathcal{S}} |h_k|$. 

Note that this achievable sum rate is within a constant gap to the upper bound found in Lem. 1.

Next, we consider the special case of $K = 2$ receivers as an example. Fig. 3 depicts the outer bound due to Lem. 1 (dashed line) and the inner bound due to Lem. 2 (dash-dotted line) for channel gains of $h_1 = 30e^j$ and $h_2 = 1000e^{2j}$. For the case $K = 2$ we can find a tighter outer bound. The two capacities $C_1$ and $C_2$ are upper bounded due to [10, Lem. 2]

$$C_i \leq \log \left(\frac{|h_i|^2}{2} + 3\sqrt{2}|h_i| + 9\right) \leq 2\log |h_i| \quad (7)$$

for $|h_i| \geq 11$. The sum capacity is bounded by the number of possible transmit symbols

$$C_1 + C_2 \leq 2n \leq 2\max(\log |h_1|, \log |h_2|) \leq 2\log |h_2|. \quad (8)$$

The resulting outer bound corresponds to the dotted line in Fig. 3.
We obtain a tight inner bound by simulation. For our specific channel gains we find sets $X_i$ of transmit symbols $x$ that merge at the weaker receiver $k = 1$ due to the rounding operation. For each set $X_i$ we count the number of symbols $n_i$ that can be distinguished by the stronger receiver $k = 2$. In our example we have $i = 1, \ldots, 471$ sets containing between $n_i \geq 2$ and $n_i \leq 1192$ distinguishable symbols. The transmitter sends information to the weaker and the stronger receiver by choosing the set index $i$ and one of the $n_i$ distinguishable symbols within the set, respectively. It is optimal to send the symbols using a numerical solver like CVX. The corresponding rate pairs are marked with crosses in Fig. 3.

We would like to find rate pairs that are located on the boundary of the region of achievable rates. This can be done by choosing a rate $R_1$ and solving the convex maximization problem

$$
\max_{p_i} R_2
$$

subject to $p_i \geq 0$, $\sum_i p_i = 1$, and $R_1 \geq \hat{R}_1$

using a numerical solver like CVX. The corresponding rate pairs are marked with crosses in Fig. 3.

Two additional rate pairs are marked with circles. They correspond to the rates achievable by sending information to only one of the receivers. The achievable rate region is the convex hull over all achievable rate pairs; it is depicted by a solid line.

IV. Parallel Relay Network

In this section we turn to the capacity of the PRN in the DSM. Like in the previous section we derive upper and lower bounds partly relying on results we presented in [10].

**Lemma 3.** The capacity of the PRN in the DSM is upper bounded by

$$
C \leq 2 \min_{S \subseteq \{1, \ldots, K\}} \left( \max_{k \in S} |h_{sk}| + \max_{k \in S} |h_{kd}| \right) + K \log 6 + 1
$$

where $\overline{S} = \{1, \ldots, K\} \setminus S$ is the complement of $S$.

**Proof:** Consider the separation of relays into sets $S$ and $\overline{S}$. We allow joint processing within each of the sets, that is we create two super-relays having multiple receptions and multiple transmissions. This can only increase the capacity. The resulting network consists of two parallel branches. The capacity of each branch is bounded by its broadcast phase and its multiple-access phase. We use the broadcast bound due to Lem. 1 on the capacity $C_1$ of the first branch,

$$
C_1 \leq 2 \max_{k \in S} \log |h_{sk}| + |S| \log 6.
$$

We bound the capacity $C_2$ of the second branch by using [10, Lem. 8] for the multiple-access phase,

$$
C_2 \leq 2 \log \sum_{k \in S} |h_{kd}| + 1
$$

$$
\leq 2 \log \left( |\overline{S}| \max_{k \in S} |h_{kd}| \right) + 1
$$

$$
= 2 \max_{k \in S} \log |h_{kd}| + 2 \log |S| + 1.
$$

The capacity of the PRN cannot be larger than the sum $C_1 + C_2$. Furthermore, each partition $S \subseteq \{1, \ldots, K\}$ yields an upper bound, of which we choose the lowest one. Observing that

$$
\min_{S \subseteq \{1, \ldots, K\}} (|S| \log 6 + 2 \log |\overline{S}| + 1) \leq K \log 6 + 1
$$

concludes the proof.

**Lemma 4.** The achievable rate over the PRN is lower bounded by

$$
R \geq 2 \min_{S \subseteq \{1, \ldots, K\}} \left( \max_{k \in S} \log |h_{sk}| + \max_{k \in S} \log |h_{kd}| \right) - 11K - 2(K - 1) \log(K - 1) - 2.
$$

**Proof:** Consider an ideal PRN having a capacity region of the multiple-access phase and the broadcast phase equal to our approximation, that is

$$
\sum_{k \in S} C_{sk} = 2 \max_{k \in S} \log |h_{sk}|, \quad \sum_{k \in \overline{S}} C_{kd} = 2 \max_{k \in \overline{S}} \log |h_{kd}|.
$$

A cut that separates source from destination divides the relays into two sets $S \subseteq \{1, \ldots, K\}$ and $\overline{S}$. The cut-set bound for the PRN is hence

$$
C_{\text{cut-set}} = 2 \min_{S \subseteq \{1, \ldots, K\}} \left( \max_{k \in S} \log |h_{sk}| + \max_{k \in \overline{S}} \log |h_{kd}| \right).
$$

We show that this cut-set bound is achievable in the ideal PRN by transforming our sum-capacity constrained network into a link-capacity constrained network and using the min-cut max-flow theorem [11], [12]. The transformation is performed stepwise and essentially replaces constraints on the sum-capacity (9) by additional nodes and links with constrained link-capacity.

One step of the transformation for the broadcast phase of the PRN is depicted in Fig. 4. The left part shows the network before the transformation step. Black circles denote physical nodes, that is the source and the relays. White circles denote artificial nodes. Links with sum-capacity constraint and link-capacity constraint are depicted by solid and dashed lines, respectively. Among the solid links we find the link with strongest channel gain $|h_{\text{strong}}|$. This link is marked by a bold line. For better pictorial representation we reordered the relay nodes to bring the strongest link to the top.

A transformation step consists of the following operations:

1. Create an artificial node joining all solid links.
2. Create a link from the previous root note of the solid links to the new node.
the DSM is in turn an approximation on the Gaussian capacity, we arrived at our final result:

**Theorem 1.** The capacity \( C_G \) of the Gaussian PRN is approximately

\[
\tilde{C} = 2 \min_{S \subseteq \{1, \ldots, K\}} \left( \max_{k \in S} \log |h_{sk}| + \max_{k \in S} \log |h_{kd}| \right)
\]

in the sense that there is a gap

\[
|C_G - \tilde{C}| \leq \kappa,
\]

that depends only on the number of relays

\[
\kappa = O(K \log K).
\]

**Proof:** Lem. 3 and Lem. 4 yield \( |C - \tilde{C}| \leq \kappa_1 \) with \( \kappa_1 = O(K \log K) \). The constant gap result \( |C_G - \tilde{C}| \leq \kappa_2 \) with \( \kappa_2 = O(K \log K) \) is provided in [9, Th. 3.2]. Hence, \( |C_G - \tilde{C}| \leq \kappa_1 + \kappa_2 = \kappa \) and \( \kappa = O(K \log K) \).

Note that the value of \( \tilde{C} \) is relatively easy to compute, since there are only \( 2^K \) possible subsets \( S \).

**V. CONCLUSION**

We found the capacity of the Gaussian PRN with general channel gains within a constant additive gap of order \( K \log K \). The approach was to find bounds on the capacity in the corresponding PRN in the DSM. Since the capacity in the DSM is an additive approximation to the Gaussian capacity, our bounds are at the same time additive bounds on the Gaussian capacity. The fact that our upper and lower bounds are within a constant gap yielded the result.

**REFERENCES**


---

3) Assign the link-constraint \( C_{\text{link}} = 2 \log |h_{\text{strong}}| \) to the new link and to the strongest link.

4) Drop all sum-capacity constraints due to sets containing the strong link, \( S : \text{strong} \in S \).

The right part of Fig. 4 shows the result of the transformation. The new node is the white circle to the right, and the new link is the one pointing at it. The two dashed bold links are link-capacity constrained.

We show that neither the maximum flow nor the minimum cut is affected by the transformation. The maximum flow conservation is trivial. Replacing the sum-capacity constraint by an equivalent link-capacity constraint for the joining link does not change the constraint on any flow.

A cut is affected by the transformation step if and only if it crosses the strongest link, because the corresponding sum-capacity constraints are dropped. If a cut crosses the strongest link, it connects the points A and B in the left part of Fig. 4. The minimum cut from A to B is the one depicted in the figure, it has the value \( 2 \log |h_{\text{strong}}| \). Hence the global minimum cut connects A and B in this way. The minimum cut in the transformed network from A to B is still \( 2 \log |h_{\text{strong}}| \) cutting the new link (see right part). Hence the value of the global minimum cut is conserved.

The remaining solid links are the starting point for the next step. The multiple-access phase of the PRN can be transformed exactly the same way. It is obvious that the resulting network has only link-capacity constraints, and hence the min-cut max-flow theorem states that the cut-set bound is achievable.2

So far we considered the ideal PRN and showed that the cut-set bound (10) is achievable. However, in the DSM we can only achieve (10) minus a constant gap. Due to Corollary 1 and [10, Th. 3] we have

\[
R \geq C_{\text{cut-set}} - 7|S| - 2 \log((|S| - 1)! - 4|S| - 2
\]

\[
\geq C_{\text{cut-set}} - 11K - 2(K - 1) \log(K - 1) - 2.
\]

With Lem. 3 and Lem. 4 we found the approximate capacity of the PRN in the DSM. Since we know that the capacity in

---

2The achievable strategy can be obtained by using the flow augmenting Ford-Fulkerson algorithm [11] for the transformed network. We can find the flow assignment for the original network directly from the flows augmented in the transformed network. This is done by considering the flow in the links attached to the relay nodes. Assigning these flows to the corresponding links in the original network yields the cut-set bound achieving flows.