Optimal Placement of Known Symbols for Slowly Varying Frequency Selective channels

Srihari Adireddy and Lang Tong†

Abstract

The problem of placing known symbols in a data stream for a slowly varying frequency selective channel is considered from an information-theoretic perspective. Given the amount of redundancy associated with known symbols, placement schemes that minimize the outage probability are derived by assuming that the transmitted codewords consist of packets that are constrained to have the same known symbol placement. Under the assumption that each known symbol cluster is at least as large as $\alpha \geq 2L + 1$ (where L is the channel order), we show that the optimal placement is obtained by arranging the known symbols into as many clusters as possible and placing them such that the unknown symbol blocks are as equal as possible. It is shown that the optimal placement of known symbol clusters does not depend on the probability density of the channel. Numerical examples are used to illustrate the ideas and potential gains of using optimal known symbol placement.

Index Terms

Known Symbols, Placement Schemes, Compound Channels, Outage Probability, Error Exponents.

†Corresponding author

S.Adireddy and L.Tong are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (email: {srihari,ltong}@ece.cornell.edu). This work was supported in part by the National Science Foundation under Contract CCR-9804019, the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564 and North American Philips Corp.

I. INTRODUCTION

Although information theory does not mandate the separation of the transmitted signal into known and unknown symbols, training symbols are inserted into the data stream to enable low complexity receiver implementations. For wireless communication, where the channel can vary significantly over time, a large portion of the resources might be dedicated to training symbols. It is therefore important for a system designer to optimize the design of training. This includes optimization of parameters such as percentage of training symbols and placement of training symbols.

In this paper, we consider the problem of optimal placement of training for a slowly varying frequency selective channel. We are particularly interested in the case when the decoding delay of the data application is such that it is not possible to code over multiple fades. The information theoretic performance of such a system is usually analyzed under a composite channel model, which is a compound channel with apriori probabilities on the channel states [1], [2]. This model, first analyzed in [3], has gathered a lot of attention. See for example [4], [5], [6]. In this model, it is assumed that the whole codeword sees a single fade irrespective of its length. The *practical* justification of this model is that the decoding delay is such that the codeword cannot be longer than the coherence time of the channel and the coherence time is long enough for information theoretic results to be meaningful. How long is long enough depends on the symbol error rate required and error exponents of this channel model. See [2] for a good discussion about these issues. Probability of outage is an important measure of performance for this channel model. We will therefore use outage probability as the performance metric to compare different placement schemes.

We assume that the receiver forms an estimate of the channel based on training alone and the estimate is then used by the decoder in order to perform decoding. It is assumed that the number of known symbols inserted in a codeword is sufficient to form a *reliable* estimate of the channel. The coherence time of the channel places a limit on the length of the codeword and thus in order to obtain a *reliable* estimate of the channel, it might be necessary to dedicate a significant percentage of the time to transmitting known symbols. The assumption that the receiver forms a reliable estimate of the channel allows us to utilize the techniques available on outage probability for known channels to address the problem of training symbol placement.

The drawback of this assumption is that it does not allow us to determine the optimal percentage of known symbols. The optimal percentage of known symbols can be however be obtained by analyzing the problem in the framework of error exponents because it allows us to take into account the error due to channel estimate. Problem formulation in terms of error exponents is explored in [7].

We assume that the codebooks used to transmit the bits consist of *packetized codewords* (PCW). Packetized codewords are constrained to consist of an integral number of *packets*. Each packet is assumed to have the same placement scheme and we are interested in optimizing this placement strategy. The problem of optimal placement for general codewords where there is no constraint on the placement of known symbols is considered in [7].

For the PCW case, we show that under the constraint that known symbols are placed in clusters of length at least $\alpha \geq 2L+1$ where L is the order of the channel, the outage probability is minimized by the family of placement schemes referred to as QPP-α. In this family, the known symbols are broken into as many clusters as possible under the constraint that each of them is at least α and they are placed such that the lengths of unknown symbol clusters are as equal as possible. Surprisingly, the optimal known symbol placement does not depend on the depend on the probability density of the channel.

There has been some prior work reported on the effect of training and channel estimation errors on mutual information that does not consider the placement issue[8], [9], [10]. Medard [8] has obtained lower and upper bounds on mutual information that are a function of the variance of the error in the channel estimate formed at the receiver. Hassibi and Hochwald [9] have optimized training in multiple-antenna systems with quasi-static flat fading by maximizing a tight lower bound on the ergodic capacity. They considered issues such as amount of training, choice of training symbols and power allocation. For the channel model considered in their work, the performance is independent of placement. The training issues for quasi-static frequency selective fading were addressed in [10]. In this paper, the placement issue was not considered and the known symbols were placed at the beginning of the packet.

We have addressed the problem of optimal placement of known symbols for ergodic block frequency selective fading with i.i.d Gaussian taps in [11]. We show that for OFDM systems, placing known symbols periodically in frequency is optimal where as periodic placement in time (QPP-α placement schemes) turns out to be optimal for single carrier systems. Optimal placement of known symbols for minimizing the variance of the error in channel estimate for OFDM systems has been addressed in [12]. The optimal placement for the more general setting of block precoded transmission systems with cyclic prefix was addressed with the channel estimate as the metric in [13] and at high SNR, with block length going to infinity, and with ergodic capacity as the metric in [14]. Training issues for tracking in Gauss-Markov channels was considered in [15], [16], [17]. The placement that minimizes the Cramer-Rao Lower Bound (CRLB) for semi-blind channel estimators was found in [18]. It was shown here that QPP-α placement is optimal under some constraints. It is quite surprising that the QPP-α placement schemes turn out to be optimal for a variety of metrics.

This paper is organized as follows. In Section II, we introduce the channel model, the model for the codebook, the receiver structure and training structure and define outage probability. In Section III we formulate the problem of optimal known symbol placement. In Section IV, we obtain optimal placement schemes. Section V illustrates the ideas proposed in the paper through different simulations. We finally conclude in VI.

II. System Model

In this section we first define the channel model. We then describe in detail the PCW model. We also give the structure of the receiver employed by each user. We then introduce the metric that is used for optimizing placement in each case.

A. Channel Model

We assume that the channel $\mathbf{h} = [h_0 h_1 \cdots h_L]^t$ to a user is random and is governed by the density function $p_{\mathbf{h}}(\cdot)$. We also assume that the channel stays constant over the duration of the codeword. We assume that neither the receiver nor the transmitter knows the propagation coefficients. The channel output to each user is corrupted by the additive, zero mean, white Gaussian noise w_k with variance 1. We assume that the average energy of the unknown symbols is equal to ρ.

B. Packetized Codewords

We assume that the codeword consists of packets that belong to the class $\mathcal{P}_\alpha$. A packet is in the class $\mathcal{P}_\alpha$ if

A1: The length of each packet is $(N + P + L)$ where N is the number of unknown symbols and $(P + L)$ is the number of known symbols.

A2: The known symbols come in clusters of length equal to at least $\alpha \geq 2L + 1$.

A3: Each packet starts with at least L known symbols.

The assumption A2 is introduced primarily from the point of view of channel estimation. A2 makes it possible to employ channel estimators based on training only. A3 implies that there is no inter-packet interference.

As shown in Fig. 2, every placement in the packet can be specified by $\mathbf{r} = (\mathbf{m}, \mathbf{n})$, that is, two tuples $\mathbf{m} = (m_1, \cdots, m_J)$ and $\mathbf{n} = (n_1, \cdots, n_{J+1})$ where $\mathbf{m}$ gives the lengths of unknown symbol blocks and $\mathbf{n}$ the lengths of known symbol clusters. For placements that end with unknown symbols we have $n_{J+1} = 0$. Further A3 implies that $n_1 \geq L$. We also have $\sum_{i=1}^{J} m_i = N$ and $\sum_{i=1}^{J+1} n_i = (P + L)$. It should be noted that the number of elements in each of these tuples is a variable and depends on the placement scheme. We refer to the symbols between any two consecutive known symbol clusters as unknown symbol blocks.

A code of rate R whose codewords consist of k packets each is denoted as (k, R). A rate R is said to be achievable with a placement $\mathbf{r}$ if there exists a sequence of codes $\{(k, R)\}, k = 1, 2 \cdots$ such that the placement of known symbols in each packet is $\mathbf{r}$ and the probability of error tends to zero.

C. Receiver Structure and Performance

We assume that the transmitter uses a codebook $\mathcal{C} = \{\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \cdots, \mathbf{s}^{(M)}\}$ to transmit the data. Fig. 3 illustrates the receiver structure, the channel estimator is given by $\hat{\mathbf{h}} = g(\mathbf{y}, \mathbf{s}_t)$ where $\mathbf{y}$ is the vector containing the received output due to the codeword and $\mathbf{s}_t$ is the vector containing the non-zero known symbols. The codeword is decoded as

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \mathcal{C}} p(\mathbf{y}|\mathbf{s}, \hat{\mathbf{h}}), \tag{1}$$

where $p(\mathbf{y}|\mathbf{s}, \hat{\mathbf{h}})$ is the conditional probability density of $\mathbf{y}$ conditioned on the transmitted codeword and the estimated channel.

D. Outage Probability as Performance Metric

Due to the constraint that each codeword undergoes a single fade, it is possible that the Shannon capacity of this channel is equal to zero. A meaningful metric for this channel is therefore outage probability. The interpretation of outage probability is the one given in [6] and the references therein. Given a transmitted SNR ρ and rate R, $\Theta(R,\rho)$ is the largest possible set for which C_Θ, the capacity of the compound channel with the parameter $\mathbf{h} \in \Theta(R,\rho)$, satisfies $C_\Theta \geq R$. The outage probability is then defined as $\mathrm{P}_{out}(R,\rho) = \Pr(\mathbf{h} \notin \Theta(R,\rho))$. Hence there exists a sequence of codes $\mathcal{C}_n$ of rate R that satisfies the power requirement ρ, and for which the supremum of the probability of error over all channels $\mathbf{h} \in \Theta(R,\rho)$ tends to zero. In this paper, we consider outage probability only for i.i.d inputs. This makes the problem tractable and it also implies that the capacity achieving codes will have a flat transmit spectrum.

III. Problem Formulation

Let $\mathrm{P}_{out}(R,\rho,\mathbf{r})$ be the probability of outage as a function of the placement scheme $\mathbf{r}$. If $\mathbf{r} = (\mathbf{m},\mathbf{n})$, it can be shown that the outage probability depends only on $\mathbf{m}$ [7]. Further, the placements corresponding to all the permutations of the elements of the tuple $\mathbf{m}$ have the same outage probability. We can thus conclude that the order in which the unknown symbol blocks or the known symbol clusters are transmitted is immaterial. In fact the exact values of $\mathbf{n}$ are also irrelevant as long as A1 is satisfied and the channel estimator is consistent. Please refer to [7] for the details.

The objective is to minimize the outage probability with respect to the size of unknown symbol blocks $\mathbf{m}$. Formally, our objective is to examine the following optimization :

$$\mathbf{m}^* = \arg\min_{\mathbf{m}\in\mathcal{P}^\alpha} \mathrm{P}_{out}(R,\rho,\mathbf{r}). \tag{2}$$

IV. Quasi Periodic Placement and its Optimality

In this section, we define a family of placement schemes for the PCW model called Quasi Periodic Placement (QPP) and state their optimality properties. The family of QPP placement schemes is divided into different classes on the basis of the smallest allowable length of any known symbol cluster. The class of schemes for which α is the smallest allowable known symbol cluster length is denoted as QPP-α. Formally, we define $\mathcal{Q}_\alpha$ as the set of all placement schemes belonging to the class QPP-α.

Definition 1: Given an α and a packet with N unknown symbols and $(P+L)$ known symbols, let $J = \lfloor \frac{P+L}{\alpha} \rfloor$. A placement scheme belongs to $\mathcal{Q}_\alpha$ if and only if

1) $\mathbf{n} \in \mathcal{N}$ where $\mathcal{N} = \{(n_1,\cdots,n_J,0) : \sum_i n_i = (P+L), \min(\{n_1,\cdots,n_J\}) \geq \alpha\}$

2) $\mathbf{m} \in \mathcal{M}$ where $\mathcal{M} = \{(m_1, \cdots, m_J) : \sum_i m_i = N, m_i \in \{\lfloor \frac{N}{J} \rfloor, (\lfloor \frac{N}{J} \rfloor + 1)\}\}$

In other words, in a QPP-α placement scheme, the known symbols are divided into as many clusters as possible under the constraint that each of them is no less than α, and these clusters are placed such that the unknown symbol blocks are as "equal" as possible. An element in $\mathcal{Q}_\alpha$ is denoted as $\mathbf{r}^\alpha$.

It can be shown [7] that under assumptions A1 and A2, given $\alpha \geq 2L+1$, all the placement schemes in $\mathcal{Q}_\alpha$ have the same outage probability (For all other α, such a claim is not true in general.). The optimality of QPP-α schemes is summarized in the following theorem.

Theorem 1: If $\mathbf{r}^\alpha \in \mathcal{Q}_\alpha$, then under A1-A3, for any given $p_{\mathbf{h}}(\cdot)$, R and ρ

$$\mathrm{P}_{out}(R, \rho, \mathbf{r}^\alpha) = \min_{\mathbf{r} \in \mathcal{P}^\alpha} \mathrm{P}_{out}(R, \rho, \mathbf{r}). \tag{3}$$

Furthermore, $\mathrm{P}_{out}(R, \rho, \mathbf{r}^\alpha)$ is a monotonically increasing function of α. Hence

$$\mathrm{P}_{out}(R, \rho, \mathbf{r}^L) = \min_{\alpha \geq L, \mathbf{r} \in \mathcal{P}^\alpha} \mathrm{P}_{out}(R, \rho, \mathbf{r}). \tag{4}$$

Proof : Please refer to [7].

Theorem 1 shows that if we allow all possible $\alpha \geq 2L+1$, the placement schemes belonging to QPP-$2L+1$ are optimal. Conventionally known symbols have been placed in big clusters. Theorem 1 indicates that there is some gain to be achieved by spreading them. The algorithm for placing the known symbols is also quite simple. The optimal placement is independent of the probability density of the channel coefficients. This property makes the scheme highly attractive for the broadcast scenario. The intuition in placing the known symbols in small clusters is that known symbol clusters reduce the inter-symbol interference (ISI) and one should thus maximize the number of known symbol clusters in the data stream.

V. SIMULATIONS

The composite model that is used in this paper can be used to model a broadcast communication system where the transmitter is transmitting common information to all the receivers. Outage probability in this context refers to the probability that a user is unable to receive the transmission.

We assume that the receiver might belong to one of three different geographical locations, each of which has a different multi-path structure. Each geographical region is assumed to have line of sight but distinctly different kind of ISI channel. The specular component in region A is assumed to be flat. The specular component in region B is assumed to have a deep null in the spectrum. The goal of the transmitter is to minimize the outage probability for a given rate.

We assume that the channel in each region has $L+1$ taps. The channel in region A is given by

$$h_A(l) = \sqrt{\beta} g_A(l) + \sqrt{1-\beta} r_A(l) \quad l = 0, 1, \cdots, L \tag{5}$$

where β gives the power of the specular component. If $g_A(z)$ denotes the z-transform of $g_A(l)$, then we have $g_A(z) = 1$. Hence the specular component is just the delta function. The channel $\{r_A(l)\}_{l=0}^{L}$ is generated from i.i.d. complex Gaussian with zero mean and variance equal to $\frac{1}{L+1}$. For region B, the z-transform of the specular component is given by $g_C(z) = k_2 \left(z^{-1} + 1\right)^L$ where once again the constant k_2 is selected so that the norm of the channel is equal to one.

Fig. 4 compares the performance of the QPP-$2L+1$ placement scheme with placing all the known symbols at the beginning of the packet(preamble scheme). We assume that $N = 112, (P+L) = 48$, $L = 1$ and the transmitted SNR is equal to 20 dB. The user is assumed to belong to one of regions with equal probability. As expected the QPP-$(2L+1)$ scheme is better than the preamble scheme at every rate. We also find that the gain of the QPP-$(2L+1)$ scheme is higher at lower outage probabilities. This is because at lower outage probabilities, the bottle neck channels belong to region B, which is where the ISI is greatest. The optimal known symbol placement, being primarily a measure to decrease ISI, provides the maximum gain for these channels.

Fig. 5 shows the variation of the outage probability with SNR for the rate R = 3.5 bps per hertz and $L = 1$. We see that it is possible to obtain large gains in outage probability using the optimal placement scheme. This has a direct implication on the coverage that can be obtained at a given SNR. Please refer to [19] for additional simulations.

VI. Conclusions

In this paper we studied the optimization of placement of known symbols in the data stream for a slowly varying frequency selective channel. The performance metric used is outage probability with i.i.d inputs. We examined the optimization of the position of known symbols in the data stream under the assumption that each codeword consists of packets and each packet contains the same number of known symbols and the same placement. We show that under the constraint each known symbol cluster is at least of length $\alpha > (2L+1)$ the outage probability is minimized by breaking the known symbols into as many clusters as possible and by doing the best for placing these clusters periodically in the data stream. In particular the placement schemes belonging to the class QPP-α are optimal.

Simulations indicate that there is gain to be obtained by optimizing the position of known symbols. The gain in optimization increases with SNR since the optimizing the placement is primarily a tool to decrease ISI. For the same reason, we see that the gain is higher in channel ensembles that are more severely affected by ISI. The placement schemes are optimal given any probability density $p_{\mathbf{h}}(\cdot)$ governing the channel realization.

REFERENCES

[1] M.Effros and A.Goldsmith, "Capacity definitions and coding strategies for general channels with receiver side information," in *Proceedings of ISIT*, (Cambridge, MA.), pp. 39–39, August 1998.

[2] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, pp. 1135–1149, May 2002.

[3] S. L.H.Ozarow and A.D.Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Trans. Vehicular Technology*, vol. 43, pp. 359–378, May 1994.

[4] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1996.

[5] I. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Trans. Telecomm*, vol. 10, pp. 585–596, Nov-Dec 1999.

[6] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, Oct 1998.

[7] S.Adireddy, *The Use of Channel State in Wireless Communication*. PhD thesis, Cornell University, Ithaca, NY, May 2003.

[8] M. Medard, "The effect upon channel capacity in wireless communication of perfect and imperfect knowledge of the channel," *IEEE Trans. Information Theory*, vol. 46, pp. 933–946, May 2000.

[9] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links," *IEEE Transactions on Information Theory*, vol. 49, pp. 951–963, April 2003.

[10] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, "Optimal training for frequency-selective fading channels," in *Proc. of Intl. Conf. on ASSP*, (Salt Lake City, Utah), pp. 2105–2108, May 2001.

[11] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of known symbols for frequency-selective block-fading channels," *IEEE Trans. Info. Theory*, vol. 48, pp. 2338–2353, August 2002.

[12] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM System," *IEEE Trans. on Consumer Electronics*, vol. 44, pp. 1122–1128, Aug. 1998.

[13] S. Ohno and G. Giannakis, "Optimal training and redundant precoding for block transmissions with application to wireless ofdm," in *Proc. ICASSP*, (Salt Lake City, UT), May 2001.

[14] S. Ohno and G.B.Giannakis, "Capacity maximizing pilots for wireless OFDM over rapidly fading channels," *submitted to IEEE Trans. Info. Theory*, April 2001.

[15] M. Medard, I. Abou-Faycal, and U. Madhow, "Adaptive coding with pilot signals," in *Proc. 38th Annual Allerton Conf. on Communication, Control and Computing.*, (Allerton, IL), October 2000.

[16] M. Dong, S.Adireddy, and L. Tong, "Optimal pilot placement for semi-blind channel tracking of packetized transmission over time-varying channels," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E86-A, pp. 550–563, March 2003.

[17] M. Dong, L. Tong, and B. Sadler, "Impact of MAC design on estimation of spatial markov process in sensor networks." to be submitted to *IEEE Trans. on Signal Processing*.

[18] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. on Signal Processing*, vol. 50, pp. 3055–3069, December 2002.

[19] S.Adireddy and L. Tong, "Optimal placement of known symbols for slowly varying frequency selective channels," Tech. Rep. ACSP-TR-04-03-01, Cornell University, April 2003. `http://acsp.ece.cornell.edu/pubR.html`.

PLACE
PHOTO
HERE

Srihari Adireddy (S'00) Srihari Adireddy was born in India in 1977. He received the B.Tech degree from the Department of Electrical Engineering, Indian Institute of Technology, Madras, 1998, and M.S and Ph.D. degrees from the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, in 2001 and 2003 respectively. Currently, he is working for Silicon Labs, Austin. His research interests include signal processing, information theory, and random-access protocols.

PLACE
PHOTO
HERE

Lang Tong (S'87,M'91,SM'01) Lang Tong is a Professor in the School of Electrical and Computer Engineering. He received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Notre Dame, Indiana. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University in 1991. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology.

Dr. Tong received *Young Investigator Award* rom the Office of Naval Research in 1996, and the *Outstanding Young Author Award* from the IEEE Circuits and Systems Society. His areas of interest include statistical signal processing, wireless communications, communication networks, sensor networks and information theory.

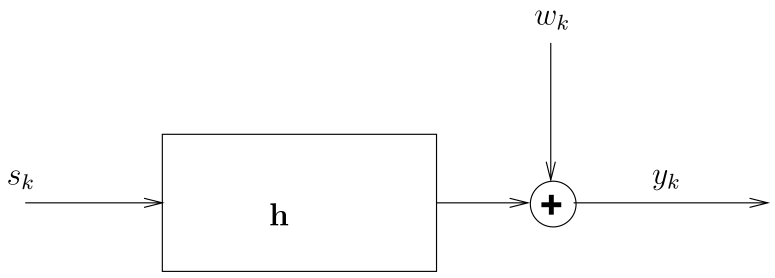

Fig. 1: System Model

figure

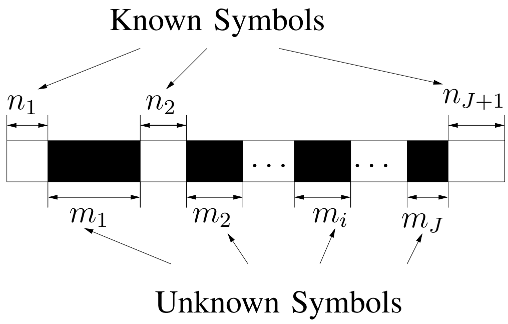

Fig. 2: Representation of Placement Schemes

figure

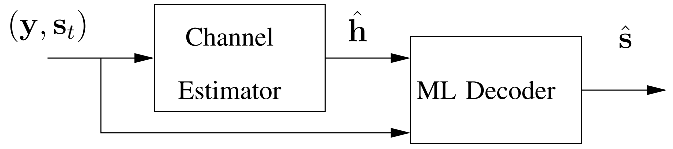

Fig. 3: Receiver Structure

figure

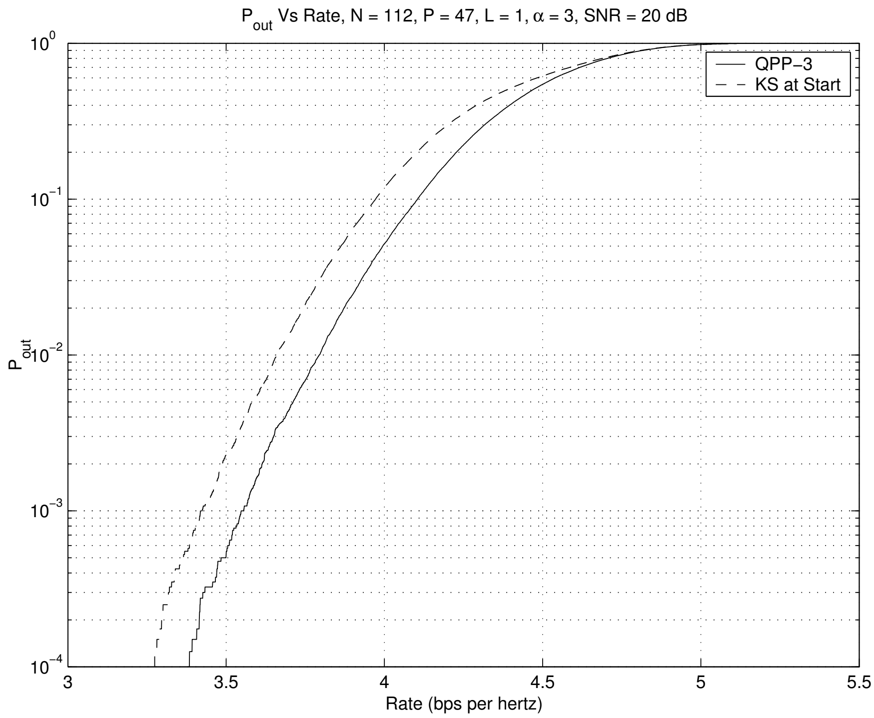

Fig. 4: Performance of Placement Schemes at SNR = 20dB, L=1

figure

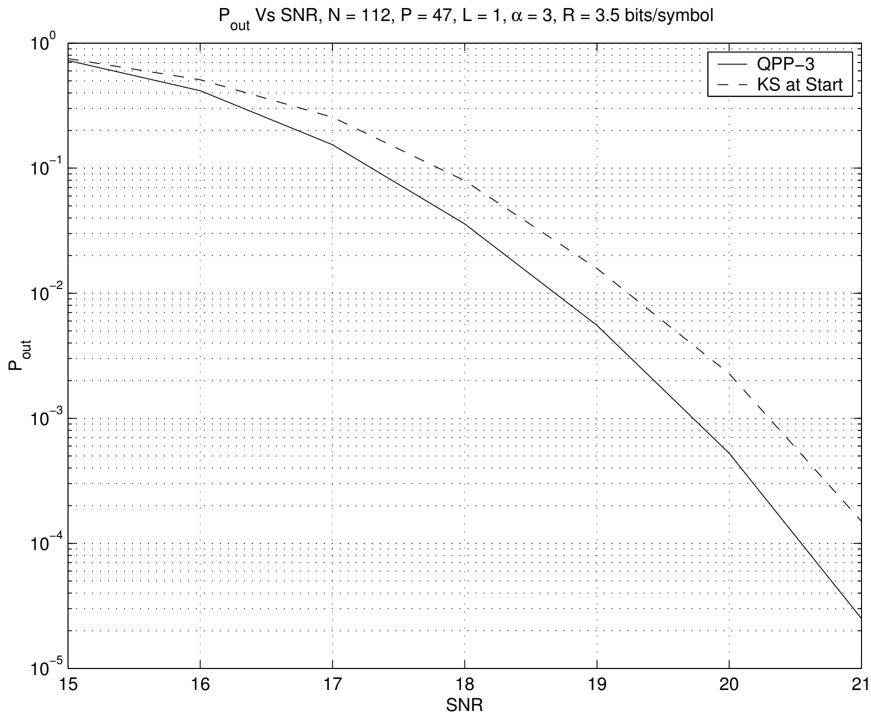

Fig. 5: Performance of Placement Schemes at R = 3.5 ,L = 1

figure