On Gaussian Channels with MMSE Interference

Ronit Bustin and Shlomo Shamai (Shitz)
Department of Electrical Engineering, Technion-Israel Institute of Technology,
Technion City, Haifa 32000, Israel, Email: bustin@tx.technion.ac.il, sshlomo@ee.technion.ac.il

Abstract—We examine codes, over the additive Gaussian noise channel, designed for or for reliable communication at some specific signal-to-noise ratio (snr) and constrained by the permitted MMSE at some lower snr. We show that the maximum possible rate is the one attained by superposition codebooks. Moreover, the MMSE function of codes attaining this maximum rate under the MMSE constraint is completely defined for all snr. The problem is also extended to the maximization of the rate under two MMSE constraints. The optimal rate is again achieved by three-layers superposition codebooks.

I. INTRODUCTION

Capacity and capacity achieving codes have been the main concern of Information Theory from the very beginning. Trying to design capacity achieving codes is a central goal of many researchers in this field. Recently some emphasis has been given to the research of non-capacity achieving point-to-point codes [1], [2]. These codes, referred to as “bad” point-to-point codes [2], are heavily used in many multi-terminal wireless networks. In [3] it was shown that the mutual information and thus also the minimum mean square error (MMSE) of “good” (capacity achieving) point-to-point codes is known exactly, no matter the specific structure of the code. Furthermore, it is known that “bad” codes can obtain lower MMSE at low signal-to-noise ratios (snrs) [2]. This advantage is meaningless in point-to-point communication, where all that matters is the performance at the receiver. However, in multi-terminal wireless networks, such as a cellular network, the case is different. In such networks there are two fundamental phenomena: interference from one node to another (an interference channel), and the potential cooperation between nodes (a relay channel). In the interference channel, where a message sent to an intended receiver acts as interference to other receivers in the network, a lower MMSE implies better interference cancelation, and thus improved rates for the interfered user. The performance of optimal point-to-point codes in the interference setting was the investigation in [4].

The best known achievable region for the two-user interference channel is given by the Han and Kobayashi (HK) scheme [5]. This scheme uses partial decoding of the interfering message at the receiver. Rate splitting (that is, superposition coding) is a special case of the HK scheme, and is also point-to-point “bad” (see [1, Appendix VIII-C]). It was shown in [6] that these codes are close to optimal for the Gaussian interference channel, and in fact are within one bit from capacity. In this work we show that these codes are in fact optimal MMSE-wise.

Although it is known that we can obtain an advantage, MMSE-wise, when using “bad” codes, the question “how much better can we do, given a looser requirement on the rate?” is still open. In this work we answer this question for the additive Gaussian noise channel showing that superposition codebooks, optimal for a specific Gaussian broadcast channel (BC) are optimal in this sense. We further extend the question to the case of two MMSE constraints in lower snrs, and show that three-layers superposition codebooks attain the maximum possible rate. This result can also be extended to the general K receiver case.

A similar question has been raised in the work of Bandemer and El Gamal [7], where they provide the rate-disturbance region: for any given rate that can be transmitted reliably to the intended receiver, what is the minimum possible disturbance that can be attained at some interfered user. In [7] the authors measure the disturbance using the mutual information between the codeword and the output at the interfered user, rather than the minimum possible MMSE, as done here. We further discuss and compare the two measures in our concluding remarks (section IV).

II. PROBLEM FORMULATION AND PRELIMINARY RESULTS

In this work we are looking at the transmission of codewords, of length n, through a discrete memoryless standard Gaussian channel:

\[ Y = \sqrt{\gamma}X + N \]  

where N is standard additive Gaussian noise. The codewords are constrained by the standard average power constraint: \( \forall x \in C_n \) \( \frac{1}{n} \sum_{i=1}^{n} x_i^2 \leq 1 \), where \( C_n \) stands for a code of n-dimensional codewords. We examine codebooks designed for reliable transmission at \( \gamma = \text{snr}_2 \) (reliable decoding of the codeword from \( Y(\gamma = \text{snr}_2) \)). Our main interest will be in examining non-optimal codes, alternatively known as “bad” codes [2], defined using code-sequences, as follows:

Definition 1: A non-optimal code-sequence \( C = \{C_n\}_{n=1}^{\infty} \), for a channel with capacity \( C \), is a code-sequence with vanishing error probability and rate satisfying \( \lim_{n \to \infty} \frac{1}{n} \log M_n < C \) where \( M_n \) is the size of code \( C_n \).

Their associated MMSE defined as:

\[ \text{MMSE}_C(\gamma) = \lim_{n \to \infty} \text{MMSE}_C(\gamma) = \lim_{n \to \infty} \frac{1}{n} \text{Tr}(E_X(\gamma)) \]  

where \( E_X(\gamma) \) is the MMSE matrix when estimating the codeword \( X \) from the output of the channel \( Y = \sqrt{\gamma}X + N \). We further define the following for abbreviation:

\[ I(\gamma) = \lim_{n \to \infty} I_n(\gamma) = \lim_{n \to \infty} \frac{1}{n} I(X;Y(\gamma)) \]  

20
Surely, for any such codes, the error probability for any \( \gamma > \text{snr} \) is zero, when \( n \to \infty \), since reliable transmission is guaranteed at \( \text{snr} \). As a result MMSE\( ^c(\gamma) \) for these snrs is also zero. On the other hand, for \( \gamma < \text{snr} \) the value of the error probability is not guaranteed to be any specific value. For an optimal code, it was shown in [3], that \( I(\gamma) \) for \( \gamma < \text{snr} \) follows that of the Gaussian i.i.d. input and thus MMSE\( ^c(\gamma) \) is also known exactly and descends gradually according to \( \frac{1}{\text{snr}} \).

Our goal is to find a code-sequence which both attains a minimum required rate at \( \text{snr} \), denoted as \( I(\text{snr}) \geq \frac{1}{2} \log (1 + \alpha \text{snr}) \), for some predetermined \( \alpha \in (0, 1] \), and secondly, obtains the minimum possible MMSE at some \( \text{snr}_1 < \text{snr} \). Clearly, if \( \alpha = 1 \) we have the maximum rate at \( \text{snr} \), attainable only by an optimal code-sequence.

As mentioned above, this is a well understood case [3]. For \( \alpha \leq \frac{\text{snr}}{\text{snr}_1} \) we also have an obvious solution, since for these low rates we can obtain zero MMSE\( ^c(\text{snr}_1) \) by simply choosing an optimal code-sequence of rate \( \frac{1}{2} \log (1 + \text{snr}) \). Thus, the interesting range of parameters is \( \frac{\text{snr}}{\text{snr}_1} < \alpha < 1 \), for which the minimum attainable MMSE\( ^c(\text{snr}_1) \) is unknown. This problem is equivalent to maximizing the rate at \( \text{snr}_1 \), given some constraint on the maximum allowed MMSE at \( \text{snr}_1 \). In this formulation the obvious extension, motivated by multi-user interference channels, is to maximize the rate at \( \text{snr}_2 \), given two MMSE constraints. In this work we proceed with this view of the problem.

A. The I-MMSE approach

The approach used in order to provide insight into the above mentioned problem is the I-MMSE approach, this to say that we make use of the fundamental relationship between the mutual information and the MMSE in the Gaussian channel and its generalizations [8], [9]. Even though we are examining a scalar Gaussian channel, the \( n \)-dimensional version of this relationship is required since we are looking at the transmission of \( n \)-dimensional codewords through the channel. In our setting the relationship is as follows:

\[
I_n(\text{snr}) = \frac{1}{2} \int_0^{\text{snr}} \text{MMSE}^c(\gamma) \, d\gamma.
\]  

(4)

Taking the limit of \( n \to \infty \) on both sides results with:

\[
I(\text{snr}) = \frac{1}{2} \int_0^{\text{snr}} \text{MMSE}^c(\gamma) \, d\gamma.
\]

(5)

The main property of the I-MMSE used for these proofs is an \( n \)-dimensional “single crossing point” property derived in [10] given here for completeness. This property is an extension of the scalar “single crossing point” property shown in [11]. In [10] the following function is defined for an arbitrary random vector \( \mathbf{X} \):

\[
q_A(\mathbf{X}, \sigma^2, \gamma) = \frac{\sigma^2}{1 + \sigma^2 \gamma} \text{Tr} (A) - \text{Tr} (A \text{EX}(\gamma))
\]

where \( A \) is some \( n \times n \) general weighting matrix. The following theorem is proved in [10].

**Theorem 1 ([10]).** Let \( A \in \mathbb{S}^n_+ \) be a positive semidefinite matrix. Then, the function \( \gamma \mapsto q_A(\mathbf{X}, \sigma^2, \gamma) \), defined in (6), has no nonnegative-to-negative zero crossings and, at most, a single negative-to-nonnegative zero crossing in the range \( \gamma \in [0, \infty) \). Moreover, let \( \text{snr}_1 \in [0, \infty) \) be that negative-to-nonnegative crossing point. Then,

1. \( q_A(\mathbf{X}, \sigma^2, 0) \leq 0 \).
2. \( q_A(\mathbf{X}, \sigma^2, \gamma) \) is a strictly increasing function in the range \( \gamma \in [0, \text{snr}_1] \).
3. \( q_A(\mathbf{X}, \sigma^2, \gamma) \geq 0 \) for all \( \gamma \in [\text{snr}_1, \infty) \).
4. \( \lim_{\gamma \to \infty} q_A(\mathbf{X}, \sigma^2, \gamma) = 0 \).

In this work, the matrix \( A \) can be set to the identity matrix. The above property is valid for all natural \( n \), thus we may also take \( n \to \infty \).

B. Superposition Coding

An important family of non-optimal codes, that is, a family of codes that do not attain the point-to-point capacity at \( \text{snr} \), is that of Gaussian superposition codes which are optimal for a degraded Gaussian BC [12]. As will be shown in the sequel these codes are optimal MMSE-wise. The analysis of this family was done by Merhav et al. in [13, section 5.3] from a statistical physics perspective. As noted in [13], the MMSE of this family of codebooks unfolds phase transitions, that is, it is a discontinuous function of \( \gamma \). The mutual information, \( I(\gamma) \), and MMSE\( ^c(\gamma) \) of this family of codebooks is known exactly and given in the next theorem. An example is depicted in Figure 1.

**Theorem 2 ([13] section 5.3):** A superposition codebook designed for \( (\text{snr}_0, \text{snr}_1, \ldots, \text{snr}_K) \) with the rate-splitting coefficients \( (\beta_0, \ldots, \beta_{K-1}) \) has the following \( I(\gamma) \):

\[
\frac{1}{2} \log (1 + \gamma), \quad \text{if } 0 \leq \gamma < \text{snr}_0
\]

\[
\frac{1}{2} \log \left( \frac{1 + \text{snr}_0}{1 + \text{snr}_0 + \gamma} \prod_{j=0}^{K-1} \frac{1 + \beta_j \text{snr}_j}{1 + \beta_j \text{snr}_j + \gamma} \right) + \frac{1}{2} \log (1 + \beta_K \gamma), \quad \text{if } \text{snr}_1 \leq \gamma < \text{snr}_{K+1}
\]

\[
\frac{1}{2} \log \left( \frac{1 + \text{snr}_0}{1 + \text{snr}_0 + \gamma} \prod_{j=0}^{K-1} \frac{1 + \beta_j \text{snr}_j}{1 + \beta_j \text{snr}_j + \gamma} \right) + \frac{1}{2} \log (1 + \beta_K - 1 \text{snr}_K), \quad \text{if } \text{snr}_1 < \gamma
\]

and the following MMSE\( ^c(\gamma) \):

\[
\text{MMSE}^c(\gamma) = \begin{cases} \frac{1}{\beta}, & 0 \leq \gamma < \text{snr}_0 \\ \frac{1}{1 + \beta \text{snr}_1}, & \text{snr}_1 \leq \gamma < \text{snr}_{K+1} \\ 0, & \text{snr}_2 < \gamma \end{cases}
\]

### III. MAIN RESULTS

The main result is given in the next theorem.

**Theorem 3:** Assuming \( \text{snr}_1 < \text{snr}_2 \) the solution of the following optimization problem,

\[
\max_{\text{snr}_2} I(\text{snr}_2)
\]

s.t. \( \text{MMSE}^c(\text{snr}_1) \leq \frac{\beta}{1 + \beta \text{snr}_1} \)

for some \( \beta \in [0, 1] \), is the following

\[
I(\text{snr}_2) = \frac{1}{2} \log (1 + \beta \text{snr}_2) + \frac{1}{2} \log \left( \frac{1 + \text{snr}_1}{1 + \beta \text{snr}_1} \right)
\]
Int. Zurich Seminar on Communications (IZS), February 29 – March 2, 2012

and is attainable when using the optimal Gaussian superposition codebook designed for \((\text{snr}_1, \text{snr}_2)\) with a rate-splitting coefficient \(\beta\).

**Proof:** It is simple to verify that the optimal Gaussian superposition codebook satisfies the above MMSE constraint and attains the maximum rate. Thus, the focus of the remainder of the proof is on deriving a tight upper bound on the rate. We prove an equivalent claim, by assuming a code of rate \(R_c = \frac{1}{2} \log (1 + \alpha \text{snr}_2)\), designed for reliable transmission at \(\text{snr}_2\), and deriving a lower bound on \(\text{MMSE}^c(\text{snr}_1)\).

If \(\alpha \text{snr}_2 \leq \gamma \leq 1\) the lower bound is trivially zero using the optimal Gaussian codebook designed for \(\alpha \text{snr}_2\). This is equivalent to setting \(\beta = 0\) in which case we obtain the rate \(I(\text{snr}_2) = \frac{1}{2} \log (1 + \text{snr}_1)\) using an optimal Gaussian codebook (no superposition). Thus, we assume \(\gamma \leq \alpha \text{snr}_2\).

Using the trivial upper bound on \(I(\gamma) \leq \frac{1}{2} \log (1 + \gamma)\) (due to maximum entropy), we can lower bound the following difference, for any \(\gamma < \alpha \text{snr}_2\):

\[
I(\text{snr}_2) - I(\gamma) \geq I(\text{snr}_2) - \frac{1}{2} \log (1 + \gamma) .
\]  

(10)

Using the I-MMSE relationship (5), the above translates to the following inequality:

\[
\frac{1}{2} \int_\gamma^{\text{snr}_2} \text{MMSE}^c(\tau) \, d\tau \geq R_c - \frac{1}{2} \log (1 + \gamma) \tag{11}
\]

\[
= \frac{1}{2} \log (1 + \alpha \text{snr}_2) - \frac{1}{2} \log (1 + \gamma) .
\]

Defining \(d\) through the following equality:

\[
\frac{1}{2} \log (1 + \alpha \text{snr}_2) - \frac{1}{2} \log (1 + \gamma) = \frac{1}{2} \log (1 + d \text{snr}_2) - \frac{1}{2} \log (1 + d \gamma) .
\]  

(12)

It is simple to check that for \(\gamma < \alpha \text{snr}_2\), \(d\) is in the range of \((0, 1)\). Now we can continue with equation (11):

\[
\frac{1}{2} \int_\gamma^{\text{snr}_2} \text{MMSE}^c(\tau) \, d\tau \geq \frac{1}{2} \log (1 + d \text{snr}_2) - \frac{1}{2} \log (1 + d \gamma) = \frac{1}{2} \int_\gamma^{\text{snr}_2} \text{mmse}_C(\tau) \, d\tau .
\]  

(13)

where \(\text{mmse}_C(\tau)\) is the MMSE assuming a Gaussian random variable with variance \(d\) through the additive Gaussian channel at \(\text{snr} = \tau\). The single crossing point property (Theorem 1, with \(A = I\)) tells us that \(\text{MMSE}^c(\tau)\) and \(\text{mmse}_C(\tau)\) cross each other at most once, and after that crossing point \(\text{mmse}_C(\tau)\) remains an upper bound. From the inequality in (13) we can thus conclude that the single crossing point, if exists, will occur in the region \((\gamma, \infty)\). Thus, for \(\gamma\) we have the following lower bound:

\[
\text{MMSE}^c(\gamma) \geq \frac{\text{mmse}_C(\gamma) - \frac{1}{2} \log (1 + \gamma)}{d(\gamma)} = \frac{\alpha \text{snr}_2 - \gamma}{\text{snr}_2 - \gamma} + \frac{1}{1 + \gamma}.
\]

A similar derivation can be done for any \(\gamma < \alpha \text{snr}_2\), and will result with a different \(d(\gamma)\). Specifically for \(\gamma = \text{snr}_1\) we obtain

\[
\text{MMSE}^c(\text{snr}_1) \geq \frac{\alpha \text{snr}_2 - \text{snr}_1}{\text{snr}_2 - \text{snr}_1} + \frac{1}{1 + \text{snr}_1} .
\]

Deriving \(\alpha\) as a function of the constraining \(\beta\), and substituting it in \(R_c = \frac{1}{2} \log (1 + \alpha \text{snr}_2)\) results with the superposition rate given in (9).

An interesting question to ask is whether there could be a different code that can attain maximum rate under the MMSE constraint at \(\text{snr}_1\) and also provide better MMSE for other values of \(\text{snr}\). The answer is to the negative, and is given in the next theorem, proof of which is omitted and can be found in [14].

**Theorem 4:** From the set of reliable codes of rate \(R_c = \frac{1}{2} \log (1 + \beta \text{snr}_2) + \frac{1}{2} \log \left(\frac{1 + \text{snr}_1}{1 + \text{snr}_0}\right)\), complying with the MMSE constraint at \(\text{snr}_1\), the superposition codebook provides the minimum MMSE for all \(\text{snr}\).

Two MMSE Constraints: We now wish to extend the result of Theorem 3 to the case of two constraints:

**Theorem 5:** Assuming \(\text{snr}_0 < \text{snr}_1 < \text{snr}_2\) the solution of the following optimization problem,

\[
\max \quad I(\text{snr}_2)
\]

s.t. \(\text{MMSE}^c(\text{snr}_0) \leq \frac{\beta_1}{1 + \beta_1 \text{snr}_1}\)

\[
\text{MMSE}^c(\text{snr}_2) \leq \frac{\beta_0}{1 + \beta_0 \text{snr}_0} \tag{14}
\]

for some positive \(\beta_1, \beta_0\) such that \(\beta_1 + \beta_0 \leq 1\) and \(\beta_1 < \beta_0\), is the following

\[
I(\text{snr}_2) = \frac{1}{2} \log \left(\frac{1 + \beta_1 \text{snr}_2 + \beta_0 \text{snr}_0 + 1}{1 + \beta_1 \text{snr}_1 + \beta_0 \text{snr}_0}\right)
\]

and is attainable when using the optimal three-layers Gaussian superposition codebook designed for \((\text{snr}_0, \text{snr}_1, \text{snr}_2)\) with rate-splitting coefficients \((\beta_0, \beta_1)\).
When $\beta_0 < \beta_1$ the first constraint can be removed and we return to the case of a single constraint given in Theorem 3.

**Proof:** It is simple to verify that the optimal Gaussian three-layers superposition codebook complies with the above MMSE constraints and attains the maximum rate. Thus, we need to derive a tight upper bound on the rate. Deriving the upper bound begins with the usage of Theorem 3 on $I(sn_1)$, when we take into consideration only the constraint on $\text{MMSE}^c(sn_0)$. This provides the following upper bound:

$$I(sn_1) \leq \frac{1}{2} \log \left( 1 + \beta_0 sn_1 \right) + \frac{1}{2} \log \left( \frac{1 + sn_0}{1 + \beta_1 sn_0} \right)$$  \hspace{1cm} (15)

On the other hand, using the I-MMSE approach we have:

$$I(sn_2) - I(sn_1) = \frac{1}{2} \int_{sn_1}^{sn_2} \text{MMSE}^c(\tau) d\tau \leq \frac{1}{2} \int_{sn_1}^{sn_2} \text{mmseo}_G(\tau) d\tau$$

where $\text{mmseo}_G(\tau)$ is the MMSE assuming a Gaussian random variable with variance $\beta_1$. This inequality is valid since according to the first constraint we have

$$\text{MMSE}^c(sn_1) \leq \frac{\beta_1}{1 + \beta_1 sn_1} \text{mmseo}_G(sn_1)$$  \hspace{1cm} (16)

thus, according to the single crossing point property

$$\text{MMSE}^c(\tau) \leq \text{mmseo}_G(\tau), \quad \forall \tau \geq sn_1$$  \hspace{1cm} (17)

leading to the inequality in equation (16). Putting together (15) and (16) we obtain the desired upper bound.

**IV. DISCUSSION AND CONCLUSIONS**

These results provide the engineering insight to the good performance of the HK superposition scheme on the two-user interference channel, as shown in [6]. From our results, that show that the HK superposition scheme, is optimal MMSE-wise, we can conclude that one cannot construct better codes of the type defined in [1], [2] that will beat HK, through the use of estimation. Note that, as mentioned in [1, section VI], the codes constructed there have an important complexity advantage over HK codes. Furthermore, keep in mind that the HK scheme is efficient in the two-user interference channel and a simple approach in the general $K$-user interference channel. Finally, we expect that the I-MMSE approach to shed further light on the capacity region of the two-user interference channel itself.

As mentioned in the introduction, Corollary 2 in [7] can also be derived directly from the I-MMSE formulation. Using our notation, we begin with the distance rate. Since, $0 \leq I(sn_2) \leq \frac{1}{2} \log \left( 1 + \alpha sn_2 \right)$ there exists an $\alpha^* \in [0,1]$ such that $R_d = \frac{1}{\pi} \log \left( 1 + \alpha^* sn_2 \right)$. Thus, the averaged MMSE of the code crosses the MMSE of a scalar Gaussian input with power $\alpha^*$ in the range $[0, sn_1]$. Now, using the I-MMSE,

$$I(sn_2) = \frac{1}{2} \log \left( 1 + \alpha^* sn_1 \right) + \int_{sn_1}^{sn_2} \text{MMSE}^c(\gamma) d\gamma$$

$$\leq \frac{1}{2} \log \left( 1 + \alpha^* sn_2 \right)$$  \hspace{1cm} (18)

where the last transition is due to the “single crossing point” property which ensures us that the MMSE of the scalar Gaussian input with power $\alpha^*$ will remain an upper bound on $\text{MMSE}^c(\gamma)$ in the range $[sn_1, \infty)$. Moreover, the above derivation does not indicate a superposition coding scheme, but rather a Gaussian code with reduced power of $\alpha^*$. Such a scheme, which attains the required minimum rate at $sn_2$, does not attain the minimum MMSE at $sn_1$. Extending to two mutual information disturbance constraint is trivial, as only the most constraining one determines the result. Thus, the different measurement of “disturbance” suggested here is conceptually different than the one suggested in [7].

Finally, the two MMSE constraints solution, presented here, can be directly extended to the $K$ MMSE constraint, as shown in [15].

**ACKNOWLEDGMENT**

This work has been supported by the Israel Science Foundation (ISF) and the CORNET Consortium. Ronit Bustin is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities.

**REFERENCES**


