# Matched Filter Decoding of Random Binary and Gaussian Codes in Broadband Gaussian Channel

Vitaly Abdrashitov, Muriel Médard and Dana Moshkovitz
Dep. of Electrical Eng. & Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA
Email: {vit,medard,dmoshkov}@mit.edu

***Abstract*—In this paper we consider the additive white Gaussian noise channel with an average input power constraint in the power-limited regime. A well-known result in information theory states that the capacity of this channel can be achieved by random Gaussian coding with analog quadrature amplitude modulation (QAM). In practical applications, however, discrete binary channel codes with digital modulation are most often employed. We analyze the matched filter decoding error probability in random binary and Gaussian coding setups in the wide bandwidth regime, and show that the performance in the two cases is surprisingly similar without explicit adaptation of the codeword construction to the modulation. The result also holds for the multiple access and the broadcast Gaussian channels, when signal-to-noise ratio is low. Moreover, the two modulations can be even mixed together in a single codeword resulting in a hybrid modulation with asymptotically close decoding behavior. In this sense the matched filter decoder demonstrates the performance that is largely insensitive to the choice of binary versus Gaussian modulation.**

## I. Introduction

The additive white Gaussian noise (AWGN) channel model has been thoroughly studied in information theory. In the discrete-time case the capacity of AWGN channel is given by $C = \frac{1}{2}\ln\left(1+\frac{P}{N_0}\right) = \frac{1}{2}\ln(1+\gamma)$ (nats) with an average input power constraint $P$, a noise power $N_0$, and the resulting signal-to-noise ratio (SNR) $\gamma = \frac{P}{N_0}$. This capacity can be achieved by the random coding with codeword symbols drawn from the zero-mean Gaussian distribution $\mathcal{N}(0, P)$ [1].

Gaussian inputs are difficult to implement by digital equipment, and are rarely used in real communications. Practical codebooks are binary. The usage of binary symbols over the AWGN channel does not allow to achieve the channel capacity. However, the gap to capacity is practically negligible for $\gamma < 0$ dB. The maximal achievable rate of a coding scheme with random binary inputs $\{+\sqrt{P}, -\sqrt{P}\}$ and the Gaussian channel is described in [2]; in the low-SNR regime ($\gamma \to 0$) it is approximated as

$$R_B(\gamma) = \frac{\gamma}{2} + O(\gamma^{\frac{3}{2}}). \quad (1)$$

As SNR goes to zero, the $R_B$ asymptotically approaches the channel capacity, i.e. the maximal achievable rate of a coding scheme with the Gaussian input distribution, since $C = \frac{1}{2}\ln(1+\gamma) \overset{\gamma\to 0}{\approx} \frac{\gamma}{2}$.

Similar results hold for the continuous-time AWGN channel with a (two-sided) noise spectral density $\frac{N_0}{2}$ and a passband bandwidth $W$. Its capacity is given by $C_{AWGN} = W\ln(1+\gamma) \overset{\gamma\to 0}{\approx} W\gamma$, with an average input power constraint $P$, where the SNR is $\gamma = \frac{P}{WN_0}$. This capacity is achieved by the random Gaussian $\mathcal{N}(0, \frac{P}{2W})$ coding with analog quadrature amplitude modulation and a continuous constellation. The random binary coding with codeword symbols $\{+\sqrt{\frac{P}{2W}}, -\sqrt{\frac{P}{2W}}\}$ and binary phase-shift keying (BPSK) modulation approaches the rate $C_{AWGN}$ tightly as the SNR goes to zero (per (1)). This is also true for high-order phase-shift keying (PSK) or quantized QAM modulations.

Note that the fact that two modulations have the same capacity in the low-SNR regime does not imply that they have similar code and decoder constructions. For instance, impulsive frequency-shift keying with low duty cycle is also capacity-approaching for the wide-bandwidth AWGN channel [3], but it employs a specific modulation and its error performance is different from that of random coding.

In our work we consider random coding and matched filter decoding. Matched filter is a maximum likelihood (ML) decoder, so it achieves the best (lowest) decoding error probability. The decoder outputs the codeword which has the maximal correlation with the channel output, i.e. the maximal sum of ML metrics over the entire channel output vector. Each metric is the product of a (demodulated) channel output symbol and the corresponding symbol of all codewords in the codebook. Therefore, this decoder uses metrics on per *codeword symbol* basis. The intuition behind codeword symbol metrics refers us to the technique of bit-interleaved coded modulation [4], which demonstrates optimal performance with decoding separated from demodulation. This behavior is quite unusual in the traditional coding theory, where decoding and demodulation should typically be performed together.

Based on the idea of ML symbol metrics, we show that in the case of random binary and Gaussian coding in the low-SNR regime, the error performance is essentially independent of the modulation in use, specifically, for analog Gaussian distributed QAM and BPSK. We also demonstrate that random

This material is based upon work supported by VT iDirect, Inc. under award No. 020903-001

binary and Gaussian codewords can be mixed together into a hybrid code as on Figure 1 with asymptotically negligible effect on performance.

The rest of the paper is organized as follows. In Section II, we consider a discrete-time channel, introduce pairwise error decoding probability, and use the Berry-Esseen theorem to estimate the difference between such probabilities in the cases of random binary and Gaussian coding. In Section III we use these results to give an upper bound in a bandlimited Gaussian channel to the difference of the average probabilities of decoding error between the two coding schemes, BPSK and analog QAM, respectively. We show this difference vanishes as the bandwidth grows sufficiently large. In Section IV we extend our results to the codes which mix both the binary and the Gaussian symbols in their codewords. In addition, we generalize our bound to multiple user and degraded broadcast channels.

## II. Discrete-time Channel — two codewords

Let us consider an additive discrete-time Gaussian channel $Y^n = X^n + Z^n$ with a Gaussian noise $z = (z_1, \dots, z_n)$, $z_i$'s are iid and drawn from $\mathcal{N}(0; N)$. A codebook $\mathcal{C}$ contains codewords $c_m \in \mathcal{C}$ of a length $n$. They are drawn randomly and independently from the binary Uniform $\left(\{+\sqrt{p}, -\sqrt{p}\}^n\right)$ or the Gaussian $\mathcal{N}(0; p)^n$ distribution in the binary or the Gaussian coding cases, respectively.

Suppose that a codeword $A = (A_1, \dots, A_n)$ has been sent. Upon receiving a noisy input $Y = A + Z$ the matched filter decoder outputs a codeword $B$ for which the channel output $Y$ is the most likely result. The decoder uses ML symbol metrics $\lambda_{mi} \triangleq Y_i c_{mi}$, where $c_{mi}$ is the $i$-th symbol of the codeword $c_m$. The decoder decision on the input $Y$ is given by $g(Y) = \arg\max_{c\in\mathcal{C}} \sum_{i=1}^n \lambda_{mi} = \arg\max_{c\in\mathcal{C}} c \cdot Y$ (here and further vectors are multiplied in the inner product sense). Thus the $g(Y)$'s correlation with the $Y$ is the largest over the entire codebook. The average pairwise probability of error, i.e. the probability that $A + Z$ will be decoded to a specific codeword $B \neq A$ is:

$$\begin{aligned}\epsilon(A,B) &= \Pr[B(A+Z) \geq A(A+Z)] \\ &= \Pr[\sum_{i=1}^n B_i(A_i+Z_i) \geq \sum_{i=1}^n A_i(A_i+Z_i)] \\ &= \Pr[\sum (B_i - A_i)(A_i+Z_i) \geq 0] \\ &= \Pr[\sum \left((B_i-A_i)(A_i+Z_i)+\mathbb{E}[A_1^2]\right) \geq n\mathbb{E}[A_1^2]] \\ &= \Pr[\frac{\sum U_i}{\sqrt{n}} \geq p\sqrt{n}].\end{aligned} \tag{2}$$

Here, $A, B$ are two different codewords in the codebook, and $U_i \triangleq (B_i - A_i)(A_i + Z_i) + p$. Since all codeword symbols in all the codewords are chosen independently from the input distribution, the random variables $\{A_1, \dots, A_n, B_1, \dots, B_n, Z_1, \dots, Z_n\}$ are mutually independent. Clearly the $U_i$'s are also independent for different values of $i$, and are identically distributed random variables with mean $\mathbb{E}U_i = \mathbb{E}[B_1A_1 + B_1Z_1 - A_1^2 - A_1Z_1 + p] = 0$. Let

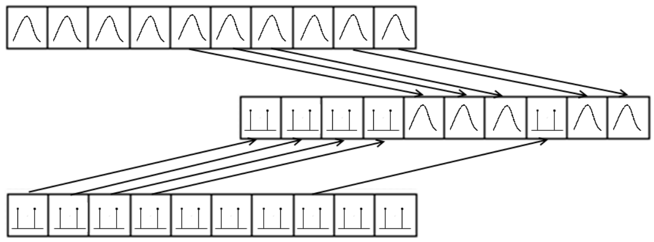

Fig. 1. Mixing of binary and Gaussian codewords

$\mu_k = \mathbb{E}[A_1^k] = \mathbb{E}[B_1^k]$ be the $k$-th moments of codeword symbols. In the binary case $\mu_{2k} = p^k$, in the Gaussian case $\mu_{2k} = (2k-1)!!p^k$. All odd moments are zero. For the $U_i$'s we can calculate their variance

$$\begin{aligned}\sigma^2 &= Var[U_i] \\ &= \mathbb{E}\left[(B_1A_1 + B_1Z_1 - A_1^2 - A_1Z_1 + p)^2\right] \\ &= 2\mathbb{E}[Z_1^2]p + \mu_4,\end{aligned} \tag{3}$$

and upper-bound the absolute third moment via the Cauchy-Schwarz inequality

$$\begin{aligned}\rho &= \mathbb{E}[|U_1|^3] \\ &\leq \sqrt{\mathbb{E}[U_1^4]\mathbb{E}[U_1^2]} \\ &= \big(\mathbb{E}[Z_1^4](6p^2 + 2\mu_4) + \mathbb{E}[Z_1^2](-24p^3 + 30p\mu_4 + 6\mu_6) \\ &\quad + 3p^4 - 6p^2\mu_4 + \mu_4^2 + 2p\mu_6 + \mu_8\big)^{1/2} (2\mathbb{E}[Z_1^2]p + \mu_4)^{1/2}.\end{aligned} \tag{4}$$

In the calculation of the variance and the third moment bound all terms $A_1^i B_1^j Z_1^k$ with odd powers of $A_1$, $B_1$, or $Z_1$ are zero mean, and do not affect the result. Note, that by (3), (4) for $p \ll \mathbb{E}[Z_1^2] = N$ (the low-SNR regime) we have

$$\begin{aligned}&\sigma_b^2 \approx \sigma_g^2 \approx 2Np \\ &\rho_b \leq 4\sqrt{3}(Np)^{3/2}, \rho_g \leq 6\sqrt{2}(Np)^{3/2}.\end{aligned} \tag{5}$$

The indices 'b' and 'g' of $\sigma, \rho$ denote the binary and the Gaussian cases, respectively.

We shall upper bound the difference between the pairwise probabilities (2) in the binary and the Gaussian cases using the Berry-Esseen theorem [5]. The theorem is a quantitative version of the central limit theorem: it gives the maximal error of approximation of a scaled sum of iid random variables $X_i$ by the normal distribution. If $\mathbb{E}[X_i] = 0, \mathbb{E}[X_i^2] > 0, \mathbb{E}[|X_i^3|] < \infty$, then for any real $t$

$$\left|\Pr[\frac{\sum_{i=1}^n X_i}{\sqrt{n}} \leq t] - \Phi_{\mathbb{E}[X_1^2]}(t)\right| \leq \frac{C\mathbb{E}[|X_i^3|]}{\mathbb{E}[X_i^2]^{3/2}\sqrt{n}}, \tag{6}$$

where $\Phi_{\sigma^2}(t)$ is the cumulative distribution function (CDF) of the zero-mean normal distribution with variance $\sigma^2$, and $C < 0.5$ [6].

Per the Berry-Esseen theorem we have

$$\left|\Pr[\frac{\sum U_i}{\sqrt{n}} \leq t] - \Phi_{\sigma^2}(t)\right| \leq \frac{C\rho}{\sigma^3\sqrt{n}}. \tag{7}$$

Let us compare the average pairwise probabilities of error (2) for the Gaussian and the binary cases. Since the pairwise probabilities of error do not depend on specific codewords $A, B$, we denote the result just by $\Delta\epsilon_{pw}$, where the subscript stands for "pairwise":

$$
\begin{aligned}
\Delta\epsilon_{pw} &\triangleq |\epsilon_g(A,B) - \epsilon_b(A,B)| && \text{(8a)}\\
&= \left|\Pr[\frac{\sum U_i^g}{\sqrt{n}} \leq t] - \Pr[\frac{\sum U_i^b}{\sqrt{n}} \leq t]\right| && \text{(8b)}\\
&= \Bigg|\Pr[\frac{\sum U_i^g}{\sqrt{n}} \leq t] - \Phi_{\sigma_g^2}(t) + \Phi_{\sigma_g^2}(t) && \text{(8c)}\\
&\quad - \Phi_{\sigma_b^2}(t) + \Phi_{\sigma_b^2}(t) - \Pr[\frac{\sum U_i^b}{\sqrt{n}} \leq t]\Bigg| && \text{(8d)}\\
&\leq \left|\Pr[\frac{\sum U_i^g}{\sqrt{n}} \leq t] - \Phi_{\sigma_g^2}(t)\right| && \text{(8e)}\\
&\quad + \left|\Phi_{\sigma_g^2}(t) - \Phi_{\sigma_b^2}(t)\right| && \text{(8f)}\\
&\quad + \left|\Phi_{\sigma_b^2}(t) - \Pr[\frac{\sum U_i^b}{\sqrt{n}} \leq t]\right| && \text{(8g)}\\
&\leq \frac{C\rho_g}{\sigma_g^3\sqrt{n}} + \Phi_{\sigma_b^2}(p\sqrt{n}) - \Phi_{\sigma_g^2}(p\sqrt{n}) + \frac{C\rho_b}{\sigma_b^3\sqrt{n}}. && \text{(8h)}
\end{aligned}
$$

Here the terms (8e) and (8g) are bounded by the Berry-Esseen theorem (7), $t$ is substituted by $t{=}p\sqrt{n}{>}0$. Also note that $\Phi_{\sigma_b^2}(t) > \Phi_{\sigma_g^2}(t)$ when $t > 0$, and we can expand (8f).

The bound (8h) shows that the average difference between the pairwise error probabilities can be made arbitrarily small by increasing the codelength $n$.

## III. Continuous-time Channel — Many codewords

Now we consider a continuous-time band-limited channel with additive white Gaussian noise with a flat spectral density $N_0/2$ and a bandwidth $W$ (Hz). Suppose we are given an average input power constraint $P$. We are in the wide bandwidth regime ($\gamma = \frac{P}{WN_0} \to 0$), and we fix an information rate $R < 2WR_B(\gamma)$ (bits/s), and a message duration $T$ (s). After that we generate a random codebook $\mathcal{C}$ with $2^{RT}$ codewords. Each codeword has $n = 2WT$ components with the variance $p = P/2W$. The binary codewords are modulated with BPSK, and the Gaussian codewords with analog QAM. After modulation, passage through the noisy channel, and reception the receiver sees each codeword symbol perturbed with Gaussian noise with the variance $N = N_0/2$.

We shall use the union bound to estimate the difference between the average probabilities of incorrect ML decoding in the binary and the Gaussian cases, given that a codeword $c$ has been sent:

$$
\begin{aligned}
\Delta\epsilon(c) &\triangleq |\epsilon_g(c) - \epsilon_b(c)|\\
&= |\sum_{i:c_i\neq c} \epsilon_g(c, c_i) - \sum_{i:c_i\neq c} \epsilon_b(c, c_i)|\\
&\leq \sum_{i:c_i\neq c} \Delta\epsilon_{pw} \leq (2^{RT} - 1)\Delta\epsilon_{pw}.
\end{aligned} \tag{9}
$$

In order to make the union bound small, we fix the codebook size $2^{RT}$, and make the $\Delta\epsilon_{pw}$ small by letting the codelength $n = \frac{2PT}{N_0\gamma}$ grow by decreasing the SNR and keeping the message duration $T$ constant. Thus, we express the $\Delta\epsilon_{pw}$ in (9) in terms of the message duration $T$ and the SNR $\gamma$:

$$
\begin{aligned}
\Delta\epsilon_{pw} \leq \frac{C\rho_g}{\sigma_g^3\sqrt{\frac{2PT}{N_0\gamma}}} &+ \frac{C\rho_b}{\sigma_b^3\sqrt{\frac{2PT}{N_0\gamma}}}\\
&+ \Phi_{\sigma_b^2}(\sqrt{\frac{PT\gamma N_0}{2}}) - \Phi_{\sigma_g^2}(\sqrt{\frac{PT\gamma N_0}{2}}).
\end{aligned} \tag{10}
$$

Note that the $p$ is expressed as $\frac{\gamma N_0}{2}$.

After that we use the expressions (3), (4) for the $\sigma, \rho$, and expand our bound in Taylor series at $\gamma = 0$:

$$
\begin{aligned}
\Delta\epsilon_{pw} &\leq (\frac{3}{\sqrt{2}} + \sqrt{3})C\sqrt{\frac{\gamma}{S}} + \frac{\gamma}{2\sqrt{2\pi}}\sqrt{S}e^{-S/2}\\
&\quad + O(\sqrt{\frac{\gamma^3}{S}}) + O(\gamma^2\sqrt{S}e^{-S/2})\\
&\overset{\gamma<1}{\leq} (\frac{3}{\sqrt{2}} + \sqrt{3})C\sqrt{\frac{\gamma}{S}} + \frac{\gamma}{2\sqrt{2\pi}}\sqrt{S}e^{-S/2}\\
&\quad + \sqrt{\frac{\gamma^3}{S}} + \frac{\gamma^2}{2}\sqrt{S}e^{-S/2},
\end{aligned} \tag{11}
$$

where $S = \frac{PT}{N_0} = TC_{AWGN}(W \to \infty)$ is the maximal possible amount of information in nats per codeword in the infinite bandwidth regime. Here we again make use of the low-SNR assumption.

Next, we use the union bound to estimate the difference between the average decoding error probabilities over the entire codebook. These average probabilities do not depend on the specific transmitted codeword, and the parameter $c$ is omitted. We also upper-bound the numerical constants, and use $\sqrt{S}e^{-S/2} < 1, \forall S \geq 0$:

$$
\begin{aligned}
\Delta\epsilon &\leq (2^{RT} - 1)\Delta\epsilon_{pw} && \text{(12a)}\\
&\leq (2^{RT} - 1)\left(2\sqrt{\frac{\gamma}{S}} + \frac{\gamma}{5} + \sqrt{\frac{\gamma^3}{S}} + \frac{\gamma^2}{2}\right) && \text{(12b)}\\
&= (2^{RT} - 1)\left(\frac{2}{\sqrt{WT}} + \frac{S}{5WT} + \frac{S}{(WT)^{3/2}} + \frac{S^2}{(WT)^2}\right) && \text{(12c)}\\
&\leq (2^{RT} - 1)\frac{2 + 2\gamma}{\sqrt{WT}} \triangleq \delta. && \text{(12d)}
\end{aligned}
$$

The final bound $\delta = \delta(T, W, \gamma, R)$ is given in terms of the codeword duration, the bandwidth, the SNR and the transmission rate. Increasing the bandwidth $W$ (thus decreasing the SNR) will make the difference between the average error probabilities in the binary and the Gaussian cases arbitrarily close to zero for the fixed rate $R$ and the duration $T$:

$$
\lim_{W\to\infty} \Delta\epsilon \leq \lim_{W\to\infty} \delta = 0. \tag{13}
$$

Therefore, the ML decoder performance on random binary codes approaches that of random Gaussian codes asymptotically.

Numerical values of the bound $\delta$ for different values of the rate, the duration $T$, and the corresponding codelength $n = 2WT$, are shown on Figure 2. The plot shows the boundary between the practical region and the unrealistic region of

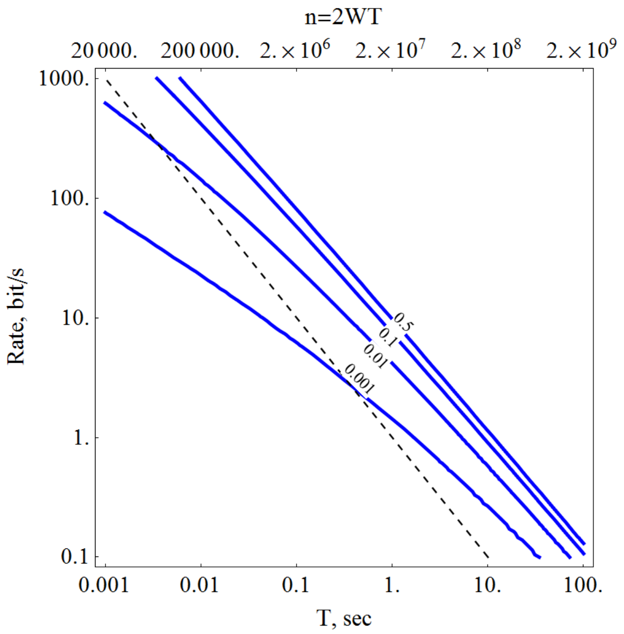


Fig. 2. Upper-bound on $\Delta\epsilon$ for various codeword durations and bit rates; $W = 10^7$ Hz, $\gamma = -40$–$0$ dB; the dashed line corresponds to codebooks with exactly $|\mathcal{C}| = 2^{RT} = 2$ codewords, the area below this line is the impractical region of less than 2 codewords.

codebooks with fewer than 2 codewords. For a bandwidth $10^7$ Hz the difference between the error probabilities can be made as low as $10^{-3}$ for bit rates below 3 bit/s, although small codebooks with few very long codewords are needed for such $\Delta\epsilon$. In this setup the bound is virtually the same for the SNR in a range $[-40, 0]$ dB.

## IV. Generalizations

The approach used in the proof suggests that neither the binary alphabet, nor the uniformity of the input distribution are essential for our results. The same method of bounding the difference between the decoding error probabilities can also be applied to different channel setups. Further generalizations will be pointed out in the conclusion.

In this section we extend our results to the hybrid binary-Gaussian random coding with mixed modulation (BPSK and continuous QAM). We also demonstrate that the results derived in the previous section are valid for multiuser environments. In particular we generalize the bound (12d) to multiple access and the degraded broadcast Gaussian channels.

### *A. Hybrid code*

The result can be extended to mixtures of the binary and the Gaussian random coding in which $n_b < n$ codeword symbols are drawn from the binary distribution, and the remaining $n_g = n - n_b$ symbols — from the Gaussian distribution (see Figure 1). The specific places of the binary and the Gaussian symbols in a codeword are not important, as long as they are the same for all codewords. The difference $\Delta\epsilon_{pw}^{h-b}$ between the pairwise error probability of this hybrid coding and that of the pure binary coding can be upper-bounded in a way similar to (8):

$$\begin{aligned}\Delta\epsilon_{pw}^{h-b} &= \left|\Pr[\sum_{i=1}^{n} U_i^h \le pn] - \Pr[\sum_{i=1}^{n} U_i^b \le pn]\right| \\ &= \left|\Pr[\frac{\sum_{i=1}^{n_g} U_i^g}{\sqrt{n_g}} \le \frac{pn}{\sqrt{n_g}}] - \Pr[\frac{\sum_{i=1}^{n_g} U_i^b}{\sqrt{n_g}} \le \frac{pn}{\sqrt{n_g}}]\right| \\ &\le \frac{C\rho_g}{\sigma_g^3\sqrt{n_g}} + \Phi_{\sigma_b^2}(\frac{pn}{\sqrt{n_g}}) - \Phi_{\sigma_g^2}(\frac{pn}{\sqrt{n_g}}) + \frac{C\rho_b}{\sigma_b^3\sqrt{n_g}}.\end{aligned} \tag{14}$$

If $n_g = \alpha_g n$ for some constant $\alpha_g$, then, by following the arguments used for (10)–(12d), we result in

$$\begin{aligned}\Delta\epsilon^{h-b} &= (2^{RT}-1)\Delta\epsilon_{pw}^{h-b} \\ &\le (2^{RT}-1)\frac{2+2\gamma}{\sqrt{WT\alpha_g}} = \delta/\sqrt{\alpha_g},\end{aligned} \tag{15}$$

where $\delta$ is the upper bound (12d). The difference $\Delta\epsilon^{h-g}$ between the hybrid and the Gaussian codes has the same expression with $\alpha_b = n/n_b$ instead of the $\alpha_g$. On the other hand $\Delta\epsilon^{h-b} \le \Delta\epsilon^{h-g} + \Delta\epsilon$ by the triangular inequality. Therefore,

$$\begin{aligned}\Delta\epsilon^{h-b} &\le (2^{RT}-1)\min\{(\Delta\epsilon^{h-g}+\Delta\epsilon), \delta/\sqrt{\alpha_g}\} \\ &\le \delta\min\{\alpha_b^{-1/2}+1, \alpha_g^{-1/2}\} \le 2.15\delta.\end{aligned} \tag{16}$$

Similarly, $\Delta\epsilon^{h-g} \le 2.15\delta$. The similar results can also be obtained be applying a more general version of Berry-Esseen theorem for non-identically distributed summands ([5]) to the mixture input distribution.

### *B. Gaussian Multiple Access Channel*

Consider $M$ transmitters (users) with average power constraints $P_m$. A wideband multiple access channel with a Gaussian noise spectral density $N_0/2$ and a bandwidth $W$ (Hz) is given by $Y = \sum_{m=1}^{M} X(m) + Z$. We assume that the total SNR is small $\frac{\sum_{m=1}^{M} P_m}{WN_0} \ll 1$. Each user is associated with a random codebook $\mathcal{C}_m$ of a size $2^{nR_m}$. If the transmission rates $R_m$ of the users are in the MAC capacity region

$$\sum_{m\in S} R_m < W\log(1+\frac{\sum_{m\in S} P_m}{WN_0}) \quad \forall S\subseteq\{1,\dots,M\}, S\neq\emptyset, \tag{17}$$

then the receiver can decode the messages of all the users correctly with high probability by decoding the messages successively with interference cancellation [8], [9].

In the low-SNR regime there is even no need for interference cancellation, we just treat the signals from all the users but one as an extra noise $\sum_{k\neq m} X(k)$. The corresponding shrinking of the capacity region is at most $W(\log(1 + \frac{P_m}{WN_0+\sum_{k\neq m} P_k}) - \log(1 + \frac{P_m}{WN_0})) \approx -\frac{P_m\sum_{k=1}^{M} P_k}{WN_0^2}$ and is negligible with respect to the capacity in the low-SNR regime.

As before, we start with analyzing the average pairwise error probabilities in the binary and Gaussian cases, and then proceed to the case of many codewords. Let $A = (A^m)_{m=1}^{M}$, $B = (B^m)_{m=1}^{M}$ be two sets of independent and identically distributed codewords for all the users, $\mathbb{E}[|A_i^m|^2] = \mathbb{E}[|B_i^m|^2] =$

$\frac{P_m}{2W}$. Then, the probability of decoding into the $B$ when the $A$ was sent is bounded by

$$\epsilon_{pw}^{MAC}(A,B) \leq \sum_m \Pr[B^m(\sum_{k=1}^{M} A^k{+}Z){\geq}A^m(\sum_{k=1}^{M} A^k{+}Z)]. \tag{18}$$

Each summand on the right side can be expressed by (2) with $U_i^m = (B_i^m - A_i^m)\cdot(\sum_{k=1}^{M} A_i^k + Z_i) + \frac{P_m}{2W}$. The low-SNR approximations (5) still hold for $\mathbb{E}[|U_i^m|^2]$ and $\mathbb{E}[|U_i^m|^3]$, given $p = \frac{P_m}{2W}$. Next, we use (8h) to bound the difference between the average pairwise error probabilities in the binary and the Gaussian cases for every term in the summation (18). For the case of many codewords with a duration $T$ and a codelength $n = 2WT$ the difference between the average probabilities of decoding error per (12d) is bounded by

$$\Delta\epsilon^{MAC} \leq \sum_{m=1}^{M}(2^{R_mT}-1)\frac{2+\frac{2P_m}{WN_0}}{\sqrt{WT}}. \tag{19}$$

### C. Gaussian Degraded Broadcast Channel

Consider a transmitter $X$ with an average power constraint $P$, and $M$ receivers (users) $\{Y_m\}_{m=1}^M$ with Gaussian noise spectral densities $\{N_m/2\}_{m=1}^M$, $N_1 < N_2 < \cdots < N_m$. A wideband degraded broadcast channel with a bandwidth $W$ is given by $Y(m) = X + Z(m)$ for $1{\leq}m{\leq}M$. In the low-SNR regime $\frac{P}{WN_1} \ll 1$. Each user is associated with a random codebook $\mathcal{C}_m$ of a size $2^{nR_m}$. If the transmission rates $R_m$ between the transmitter and the users are in the capacity region

$$R_m < W\log(1{+}\frac{\alpha_m P}{WN_m + \sum_{k=1}^{m-1}\alpha_k P}) \quad \forall m \in \{1,\ldots,M\}$$
$$\text{for some } (\alpha_1,\ldots,\alpha_M), \text{such that } \alpha_k{\geq}0, \sum_{k=1}^{M}\alpha_k{=}1, \tag{20}$$

then all the receivers can decode their messages correctly with high probability, given the transmitter allocates $\alpha_m$ fraction of its power to the $m$-th user. Each user decodes the messages of less capable users first, then performs interference cancellation and decodes its own message [10], [11].

As we did with multiple access channel, we simplify decoding in the low-SNR regime by treating other users' messages as an extra noise. The capacity region shrinks to $R_m < W\log(1+\frac{\alpha_m P}{WN_m+P(1-\alpha_m)})$, its difference with (20) is negligible. Let $A = (A^m)_{m=1}^M$, $B = (B^m)_{m=1}^M$ be two sets of independent and identically distributed codewords for all the users, $\mathbb{E}[|A_i^m|^2] = \mathbb{E}[|B_i^m|^2] = \frac{\alpha_m P}{2W}$. Then, the probability of decoding into the $B$ when the $A$ was sent is bounded by

$$\epsilon_{pw}^{DBC}(A,B){\leq}\sum_m \Pr[(B^m{-}A^m)\left(\sum_{k=1}^{M} A^k{+}Z(m)\right) \geq 0]. \tag{21}$$

Each term on the right side can be expressed by (2) with $U_i^m = (B_i^m{-}A_i^m)(\sum_{k=1}^{M} A_i^k{+}Z_i(m)){+}\frac{\alpha_m P}{2W}$. The low-SNR approximations (5) still hold for $\mathbb{E}[|U_i^m|^2]$ and $\mathbb{E}[|U_i^m|^3]$, given $p{=}\frac{\alpha_m P}{2W}$. By using (8h) and (12d), we get a bound for the difference between the average probabilities of decoding error with the Gaussian/binary coding for the case of many codewords with a duration $T$ and a codelength $n{=}2WT$:

$$\Delta\epsilon^{DBC} \leq \sum_{m=1}^{M}(2^{R_mT}-1)\frac{2+\frac{2\alpha_m P}{WN_m}}{\sqrt{WT}}. \tag{22}$$

## V. Conclusion

In this paper we compared the performance of matched filter decoding of the binary and Gaussian coding in the AWGN channel with average input power constraint. We showed that in the low-SNR wide-bandwidth regime in capacity approaching setups with binary encoder and BPSK modulation the decoding error probability is asymptotically close to that of the random Gaussian coding scheme with continuous constellation with the same codelength. The result can be easily extended to multiuser channels. The binary input distribution achieves the lowest value of $\rho/\sigma^3$ in (8h), and is therefore the closest to the optimal Gaussian input distribution in terms of the bound (12a) on $\Delta\epsilon$. However, the technique we used for comparison of the input distributions also works for non-binary distributions (with more than 2 points of support), thus, our result is valid for higher order digital modulations (QPSK, 16-QAM, etc.). The performance of non-uniform input distributions can also be bounded with our method. They come up, for example, in the channels with a peak power constraint, whose optimal input distributions are not Gaussian, nor even continuous [7]. Finally, different codeword symbols do not need to be identically distributed. We analyzed a hybrid code which uses both Gaussian, and binary symbols in its codewords, and requires both discrete and continuous modulation at the same time. Such a code is comparable to the pure binary or the pure Gaussian codes in terms of the error performance.

## References


[1] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, jul–oct 1948.

[2] A. Nasif and G. Karystinos, "Binary transmissions over additive Gaussian noise: a closed-form expression for the channel capacity," *Proc. 2005 Conference on Information Sciences and Systems (CISS)*, 2005.

[3] R. Kennedy, *Fading dispersive communication channels*. Wiley-Interscience, 1969.

[4] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *Information Theory, IEEE Transactions on*, vol. 44, no. 3, pp. 927–946, 1998.

[5] A. Berry, "The accuracy of the Gaussian approximation to the sum of independent variates," *Transactions of The American Mathematical Society*, pp. 122–136, 1941.

[6] I. Shevtsova, "On the absolute constants in the Berry-Esseen type inequalities for identically distributed summands," *arXiv preprint arXiv:1111.6554*, 2011.

[7] J. Smith, "The information capacity of amplitude and variance-constrained scalar Gaussian channels," *Information and Control*, vol. 18, pp. 203–219, 1973.

[8] H. H.-J. Liao, "Multiple access channels," Ph.D. dissertation, HAWAII UNIV HONOLULU, 1972.

[9] R. Ahlswede, *Multi-way communication channels*, ser. Second International Symposium on Information Theory: Tsahkadsor, Armenia, USSR, Sept. 2 - 8, 1971. Akadémiai Kiadó, 1973, pp. 23–51.

[10] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *Information Theory, IEEE Transactions on*, vol. 19, no. 2, pp. 197–207, 1973.

[11] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Problemy Peredachi Informatsii*, vol. 10, no. 3, pp. 3–14, 1974.