Increased Capacity per Unit-Cost by Oversampling

Tobias Koch  
University of Cambridge  
Cambridge CB2 1PZ, UK  
Email: tobi.koch@eng.cam.ac.uk

Amos Lapidoth  
ETH Zurich  
8092 Zurich, Switzerland  
Email: lapidoth@isi.ee.ethz.ch

Abstract

It is demonstrated that doubling the sampling rate recovers some of the loss in capacity incurred on the bandlimited Gaussian channel with a one-bit output quantizer.

1 Introduction

We study the capacity of the continuous-time, bandlimited, additive white Gaussian noise (AWGN) channel with one-bit output quantization. Our focus is on the capacity at low transmit powers, i.e., on the capacity per unit-cost, which is defined as the slope of the capacity-vs-input-power curve at zero. We show that increasing the sampling rate reduces the loss in capacity per unit-cost caused by the quantization.

The capacity of the continuous-time AWGN channel without output quantization was studied by Shannon [1]. He showed that if the channel input is bandlimited to $W$ Hz and satisfies the average-power constraint $P$, and if the additive Gaussian noise is of double-sided power spectral density $N_0/2$, then the capacity (in nats per second) is given by (see also [2])

$$C(P) = W \log\left(1 + \frac{P}{WN_0}\right)$$  \hspace{1cm} (1)

where $\log(\cdot)$ denotes the natural logarithm function. This capacity can be achieved by transmitting

$$X(t) = \sum_{\ell=-\infty}^{\infty} X_\ell \text{sinc}(2W t - \ell), \hspace{0.5cm} t \in \mathbb{R}$$  \hspace{1cm} (2)

(2

where $\mathbb{R}$ denotes the set of real numbers), and by sampling the output $Y(\cdot)$ at Nyquist rate $2W$. Here $\{X_\ell, \ell \in \mathbb{Z}\}$ (where $\mathbb{Z}$ denotes the set of integers) is a sequence of independent and identically distributed (IID) Gaussian random variables of zero mean and variance $P$, and $t \mapsto \text{sinc}(t)$ denotes the sinc-function, i.e.,

$$\text{sinc}(t) = \begin{cases} 1, & t = 0 \\ \frac{\sin(\pi t)}{\pi t}, & t \neq 0. \end{cases}$$

The above (capacity-achieving) transmission scheme reduces the continuous-time channel to a discrete-time AWGN channel with inputs $\{X_\ell, \ell \in \mathbb{Z}\}$ and outputs $\{Y(\ell/(2W)), \ell \in \mathbb{Z}\}$. Yet, it is often required that the channel inputs and outputs be not only discrete in time, but also take on a discrete value, i.e., take value in a finite set rather than in $\mathbb{R}$. This is, for example, the case if the transmitter and receiver use digital signal processing techniques. To ensure that the channel inputs are discrete-valued, we can simply restrict ourselves to finite input alphabets. This restriction is not critical for small input powers $P$. Indeed, it is well-known that binary inputs achieve the capacity per unit-cost of the AWGN channel [1]. To ensure that the channel
outputs are discrete-valued, we have to employ a quantizer (analog-to-digital converter), which approximates the continuous-valued output by a finite number of bits.

The capacity (in nats per channel use) of the discrete-time AWGN channel with binary symmetric output quantization—where the quantizer produces 1 for a nonnegative output and −1 for a negative output—is given by

$$\log 2 - H_b\left(Q\left(\sqrt{P/\sigma^2}\right)\right)$$

where $\sigma^2$ denotes the variance of the additive noise, $H_b(\cdot)$ the binary entropy function, and $Q(\cdot)$ the $Q$-function; see [3, (3.4.18)], [4, p. 107]; [5, Thm. 2]. To the best of our knowledge, there exists no closed-form expression for the capacity of the discrete-time AWGN channel with nonbinary output quantization. However, numerical results are, for example, given in [5]. Furthermore, there exist analytical results concerning the capacity per-unit cost. For example, it was demonstrated that if a binary symmetric quantizer is employed, then the capacity per unit-cost equals $\frac{1}{\pi^2}$ [3, (3.4.20)]. It was further demonstrated that for an octal quantizer with uniform quantization the capacity per unit-cost is not less than 0.475$\frac{1}{\pi^2}$ [3, (3.4.21)]. Thus, at low transmit power, employing a binary quantizer causes a loss of a factor of $2/\pi$ relative to the capacity per unit-cost $\frac{1}{\pi^2}$ for unquantized decoding [1]. In contrast, by quantizing the output with 3 bits, a capacity per unit-cost can be achieved that is close to the capacity per unit-cost for unquantized decoding. (Note that the capacity of discrete-time channels is measured in nats per channel use, whereas the capacity of continuous-time channels is measured in nats per second. Since with a continuous-time signal of bandwidth $W$ Hz we can approximately transmit $2W$ samples per second, we have that one nat per channel use corresponds to $2W$ nats per second. By further noting that lowpass filtering and sampling Gaussian noise of double-sided power spectral density $N_0/2$ yields Gaussian noise-samples of variance $WN_0$, it follows that the capacity per unit-cost of the continuous-time channel corresponds to $2W$ times the capacity per unit-cost of the discrete-time channel with $\sigma^2$ replaced by $WN_0$.)

The above results suggest that, in order to reduce the loss in capacity per unit-cost caused by the quantization, one needs to increase the quantizer’s resolution. However, while this clearly holds for the discrete-time channel, this does not necessarily hold for the underlying continuous-time channel. Indeed, in contrast to the unquantized channel output, the quantized output is not bandlimited, and it is therefore prima facie not clear, whether sampling the quantized output at Nyquist rate is optimal. One might thus increase the capacity of the continuous-time channel by oversampling the quantized output, i.e., by sampling the quantized output at rates higher than the Nyquist rate.

When there is no additive noise, it was shown by Gilbert [6] and by Shamai [7] that oversampling indeed increases the capacity. In this paper, we demonstrate that oversampling also increases the capacity when the noise power is large relative to the transmit power. In particular, we show that, for binary symmetric output quantization, sampling the quantized output at twice the Nyquist rate yields a capacity per unit-cost that is not less than $0.75\frac{1}{N_0}$, which is strictly larger than the capacity per unit-cost $\frac{2}{\pi^2}N_0 \approx 0.64\frac{1}{N_0}$ that can be achieved by sampling the quantized output at Nyquist rate.

The rest of this paper is organized as follows. Section 2 describes the channel model. Section 3 defines channel capacity and capacity per unit-cost and presents the main result. Section 4 provides the proofs of the main result. Section 5 concludes the paper with a discussion of our results.

## 2 Channel Model

We consider the communication channel depicted in Figure 1 whose input $x(\cdot)$ is bandlimited to $W$ Hz and satisfies the average-power constraint $P$. The channel output $Y(kT_s)$ at integer multiples of the sampling interval $T_s > 0$ is

$$Y(kT_s) = \text{sgn}\left\{\left[(x + Z) * \text{LPF}_W\right](kT_s)\right\}, \quad k \in \mathbb{Z}$$

where $\text{sgn}\{\cdot\}$ denotes the sign function; $(a * b)(t)$ the convolution between $a(\cdot)$ and $b(\cdot)$ at time $t$; and $\text{LPF}_W(\cdot)$ is the impulse response of the ideal unit-gain lowpass filter of cutoff frequency $W$.\[2\]
Figure 1: System model.

The hard-limiter is a binary symmetric quantizer that produces 1 for a nonnegative input and −1 for a negative input. We assume that \( \{ Z(t), t \in \mathbb{R} \} \) is white Gaussian noise of double-sided power spectral density \( N_0 / 2 \).

Without loss of optimality, we restrict ourselves to signals \( x(\cdot) \) of the form

\[
x(t) = \frac{1}{\sqrt{2W}} \sum_{\ell = -\infty}^{\infty} x_\ell g \left( t - \frac{\ell}{2W} \right), \quad t \in \mathbb{R}
\]

where \( g(\cdot) \) is some unit-energy waveform that is bandlimited to \( W \) Hz. Indeed, by the Sampling Theorem [8, Thm. 8.4.5], any signal \( x(\cdot) \) that is bandlimited to \( W \) Hz can be written as (5) with

\[
x_\ell = x \left( \frac{\ell}{2W} \right), \quad \ell \in \mathbb{Z} \quad \text{and} \quad g(t) = \sqrt{2W} \text{sinc}(2Wt), \quad t \in \mathbb{R}.
\]

3 Channel Capacity and Capacity per Unit-Cost

We define the capacity (in nats per second) as

\[
C_{T_s}(P) \triangleq \lim_{n \to \infty} \sup_{n} \frac{2W}{n} I(X^n_1; Y^n_1)
\]

where the supremum is over all unit-energy waveforms \( g(\cdot) \) that are bandlimited to \( W \) Hz and over all joint distributions on \( (X_1, X_2, \ldots, X_n) \) satisfying \( \frac{1}{n} \sum_{k=1}^{n} E[X_k^2] \leq P \). Here \( \lim \) denotes the limit inferior; \( A_m^n \) is used to denote the sequence \( A_m, A_{m+1}, \ldots, A_n \); and

\[
Y_k \triangleq \left( Y \left( \left\lfloor \frac{2k - 1}{4W T_s} \right\rfloor T_s \right), Y \left( \left\lfloor \frac{2k - 1}{4W T_s} \right\rfloor T_s + T_s \right), \ldots, Y \left( \left\lfloor \frac{2k + 1}{4W T_s} \right\rfloor T_s \right) \right)
\]

(with \( \left\lfloor \cdot \right\rfloor \) and \( \left\lceil \cdot \right\rceil \) denoting the ceiling and the floor function). A more general definition of channel capacity for continuous-time channels can be found in [2, Sec. 8.1]. For the above channel (4), the capacity \( C_{T_s}(P) \) defined by (6) is a lower bound on the capacity defined in [2, Sec. 8.1]. The two capacities coincide, for example, for the continuous-time AWGN channel (without output quantization).

That oversampling can increase the capacity of the above channel has been demonstrated in the noiseless case, i.e., when \( N_0 = 0 \). In particular, Gilbert [6] showed that, for a Gaussian input \( X(\cdot) \), sampling the output at twice the Nyquist rate yields an information rate of \( 2.14W \) bits per second, which is strictly larger than the \( 2W \) bits per second that can be achieved by sampling the output at Nyquist rate. Shamai [7] further showed, inter alia, that by sampling the output at \( \eta \)-times the Nyquist rate, rates of \( 2W \log(1 + \eta) \) nats per second are achievable by transmitting a bandlimited process that possesses a single real zero within each Nyquist interval. In the absence of noise it is thus possible to trade amplitude resolution versus time resolution.

In this paper, we focus on the case where the variance of the additive noise is large relative to the transmit power. In particular, we study the capacity per unit-cost, defined as

\[
\hat{C}_{T_s}(0) \triangleq \lim_{P \to 0} \frac{C_{T_s}(P)}{P}
\]
(where \( \lim \) denotes the \textit{limit superior}). By the Data Processing Inequality [9, Thm. 2.8.1] it follows that quantizing the output does not increase the capacity. This implies that the capacity per unit-cost is upper bounded by the capacity per unit cost of the continuous-time AWGN channel (without output quantization)

\[
\dot{C}_T(0) \leq \lim_{P \to \infty} \frac{W \log \left( 1 + \frac{P}{WN_0} \right)}{P} = \frac{1}{N_0}.
\]

(8)

For the case where the output is sampled at Nyquist rate \( 1/T_s = 2W \), it was shown that the capacity per unit-cost is given by [3, (3.4.20)]

\[
\dot{C}_{1/2}(0) = \frac{2}{\pi} \frac{1}{N_0} \approx 0.637 \frac{1}{N_0}.
\]

(9)

Thus, when we sample the output at Nyquist rate, hard-limiting causes a loss of a factor of \( 2/\pi \). This loss can be reduced by sampling the output at twice the Nyquist rate:

\begin{equation}
\dot{C}_{1/4}(0) \geq \frac{2}{\pi} \frac{1}{N_0}
\end{equation}

\[
\approx 0.747 \frac{1}{N_0}
\]

(10)

where

\[
\rho = \frac{2}{\pi}, \quad g_0 = \frac{1 + \frac{2}{\pi} \lambda}{\sqrt{\frac{2}{\pi} \lambda^2 + \frac{1}{\pi} \lambda + 1}} \bigg|_{\lambda=1.4}, \quad \text{and} \quad g_1 = \frac{\frac{2}{\pi} + \frac{1}{\pi} \lambda}{\sqrt{\frac{2}{\pi} \lambda^2 + \frac{1}{\pi} \lambda + 1}} \bigg|_{\lambda=1.4}.
\]

Proof. See Section 4.2.

The main ingredients in the proof of Theorem 1 are expansions of the complementary cumulative distribution function (CCDF) of bivariate and trivariate Gaussian vectors around the orthant probability. We present these expansions in the following two propositions.

**Proposition 2.** Let \( (x,y) \mapsto \phi_{0,K}(x,y) \) denote the probability density function (PDF) of the bivariate, zero-mean, Gaussian vector of covariance matrix \( K = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \) for \( |\rho| < 1 \). Then, for every \( A \geq 0 \), \( \alpha \in \mathbb{R} \) and \( \beta \in \mathbb{R} \),

\[
\int_{-\alpha A}^{\infty} \int_{-\beta A}^{\infty} \phi_{0,K}(x,y) \, dy \, dx = \frac{1}{4} + \frac{1}{2\pi} \arcsin(\rho) + \frac{\alpha + \beta}{2\sqrt{2\pi}} A + \Delta(A, \alpha, \beta)
\]

(11)

where

\[
|\Delta(A, \alpha, \beta)| \leq A^2 \eta(A, \alpha, \beta)
\]

and where \( \eta(A, \alpha, \beta) = \eta(A, |\alpha|, |\beta|) \) is monotonically increasing in \( (A, |\alpha|, |\beta|) \) and is bounded for every finite \( A, \alpha, \) and \( \beta \).

\(^1\)The orthant probability is the probability that all components of a random vector have the same sign.
Proof. See Section 4.1.1. \(\square\)

**Proposition 3.** Let \((x, y, z) \mapsto \phi_{0,K}(x, y, z)\) denote the PDF of the trivariate, zero-mean, Gaussian vector of covariance matrix

\[
K = \begin{pmatrix}
1 & \varrho_{12} & \varrho_{13} \\
\varrho_{12} & 1 & \varrho_{23} \\
\varrho_{13} & \varrho_{23} & 1
\end{pmatrix}
\]

for \(|\varrho_{12}| < 1, |\varrho_{13}| < 1, |\varrho_{23}| < 1\) satisfying \(\det(K) > 0\) (where \(\det(K)\) denotes the determinant of the matrix \(K\)). Then, for every \(A \geq 0, \alpha \in \mathbb{R}, \beta \in \mathbb{R},\) and \(\gamma \in \mathbb{R},\)

\[
\int_{-\infty}^{\infty} \int_{-A}^{\infty} \phi_{0,K}(x, y, z) \, dz \, dx = 1 + \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin(\varrho_{12}) + \arcsin(\varrho_{13}) + \arcsin(\varrho_{23}) \right)
\]

\[
+ \frac{A}{\sqrt{2\pi}} \left[ \frac{\alpha + \beta + \gamma}{4} + \frac{\alpha}{2\pi} \arcsin \left( \frac{\varrho_{23} - \varrho_{12}\varrho_{13}}{\sqrt{(1 - \varrho_{12}^2)(1 - \varrho_{13}^2)}} \right) \right]
\]

\[
+ \frac{\beta}{2\pi} \arcsin \left( \frac{\varrho_{13} - \varrho_{12}\varrho_{23}}{\sqrt{(1 - \varrho_{12}^2)(1 - \varrho_{23}^2)}} \right) + \frac{\gamma}{2\pi} \arcsin \left( \frac{\varrho_{12} - \varrho_{13}\varrho_{23}}{\sqrt{(1 - \varrho_{13}^2)(1 - \varrho_{23}^2)}} \right)
\]

\[
+ \Delta(A, \alpha, \beta, \gamma)
\]

where

\[
|\Delta(A, \alpha, \beta, \gamma)| \leq A^2 \eta(A, \alpha, \beta, \gamma)
\]

and where \(\eta(A, \alpha, \beta, \gamma) = \eta(A, |\alpha|, |\beta|, |\gamma|)\) is monotonically increasing in \((A, |\alpha|, |\beta|, |\gamma|)\) and is bounded for every finite \(A, \alpha, \beta,\) and \(\gamma.\)

Proof. See Section 4.1.2. \(\square\)

### 4 Proofs

The proof of Theorem 1 is based on the expansions of the CCDF that were presented in Propositions 2 and 3. We derive these expansions in Section 4.1. The proof of Theorem 1 is given in Section 4.2.

#### 4.1 Complementary Cumulative Distribution Functions

We shall first evaluate the CCDF for the bivariate case. The result will then be used to solve the trivariate case.

##### 4.1.1 Bivariate Case

In order to prove Proposition 2, we express the CCDF as

\[
\int_{-A}^{\infty} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx
\]

\[
= \int_{0}^{\infty} \int_{0}^{\infty} \phi_{0,K}(x, y) \, dy \, dx + \int_{-A}^{\infty} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx + \int_{0}^{\infty} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx
\]

\[
= \frac{1}{4} + \frac{1}{2\pi} \arcsin(\varrho) + \int_{-A}^{\infty} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx + \int_{0}^{\infty} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx
\]

where the last step follows from the expression for the orthant probability of a bivariate, zero-mean Gaussian vector with correlation coefficient \(\varrho\) [10], see also [11]. We proceed by evaluating the integrals on the right-hand side (RHS) of (13). To evaluate the first integral, we express the
joint PDF \((x, y) \mapsto \phi_{0,K}(x, y)\) as the product of a marginal PDF \(f(x)\) and a conditional PDF \(f(y|x)\), i.e.,
\[
\phi_{0,K}(x, y) = \phi_{0,1}(x)\phi_{\varphi x, -\varphi^2}(y), \quad (x \in \mathbb{R}, y \in \mathbb{R})
\]
where \(x \mapsto \phi_{\mu,\sigma^2}(x)\) denotes the PDF of a Gaussian random variable of mean \(\mu\) and variance \(\sigma^2\). We thus obtain for the first integral on the RHS of (13)
\[
\int_{-\alpha A}^{0} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx = \int_{-\alpha A}^{0} \phi_{0,1}(x) \int_{-\beta A}^{\infty} \phi_{\varphi x, -\varphi^2}(y) \, dy \, dx
\]
\[
= \int_{-\alpha A}^{0} \phi_{0,1}(x) \left[ 1 - Q \left( \frac{\beta A + \rho x}{\sqrt{1 - \varphi^2}} \right) \right] \, dx
\]
(14)
where \(Q(\cdot)\) denotes the Q-function
\[
Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp \left( -\frac{\xi^2}{2} \right) \, d\xi, \quad x \in \mathbb{R}.
\]
Expressing \(x \mapsto Q(x)\) as a Taylor series around zero yields [12, (3.54)]
\[
\int_{-\alpha A}^{0} \int_{-\beta A}^{\infty} \phi_{0,K}(x, y) \, dy \, dx = \int_{-\alpha A}^{0} \phi_{0,1}(x) \left[ \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \frac{\varphi A}{\sqrt{1 - \varphi^2}} \right] \, dx
\]
\[
+ \frac{1}{2} \int_{-\alpha A}^{0} \phi_{0,1}(x) \int_{-\beta A}^{\infty} \phi_{\varphi x, -\varphi^2}(y) \, dy \, dx
\]
(15)
where
\[
\Delta_1(A, \alpha, \beta) \triangleq \frac{\beta A}{\sqrt{2\pi(1 - \varphi^2)}} \int_{-\alpha A}^{0} \phi_{0,1}(x) \, dx + \frac{\varphi}{\sqrt{2\pi(1 - \varphi^2)}} \int_{-\alpha A}^{0} \phi_{0,1}(x) x \, dx
\]
\[
- \int_{-\alpha A}^{0} \phi_{0,1}(x) \delta \left( \frac{\beta A + \rho x}{\sqrt{1 - \varphi^2}} \right) \, dx - \frac{1}{2} \delta(\alpha A)
\]
(16)
and where \(\delta(\cdot)\) denotes the remainder term of the Taylor series expansion of \(Q(\cdot)\), which satisfies [13, Sec. 0.317]
\[
|\delta(x)| \leq \frac{|x|^\beta}{6} \frac{1}{\sqrt{2\pi}}, \quad |x| \leq 1.
\]
(17)
Expressing again \(x \mapsto Q(x)\) as a Taylor series around zero, the first two terms on the RHS of (15) are evaluated as
\[
\frac{1}{2} \int_{-\alpha A}^{0} \phi_{0,1}(x) \, dx + \frac{1}{2} \delta(\alpha A) = \frac{1}{2} \left( \frac{1}{2} - Q(\alpha A) \right) + \frac{1}{2} \delta(\alpha A)
\]
\[
= \frac{\alpha}{2} \frac{A}{\sqrt{2\pi}}.
\]
(18)
It is shown in Appendix A.1 that \(\Delta_1(A, \alpha, \beta)\) satisfies
\[
|\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta)
\]
where \(\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)\) is monotonically increasing in \((A, |\alpha|, |\beta|)\) and is bounded for every finite \(A, \alpha, \beta\).
Along the same lines, we evaluate the second integral on the RHS of (13) as
\[
\int_{0}^{\infty} \int_{-\beta A}^{0} \phi_{0,K}(x, y) \, dy \, dx = \int_{-\beta A}^{0} \phi_{0,1}(y) \int_{0}^{\infty} \phi_{\varphi y, -\varphi^2}(x) \, dx \, dy
\]
\[
= \int_{-\beta A}^{0} \phi_{0,1}(y) \left[ 1 - Q \left( \frac{\varphi y}{\sqrt{1 - \varphi^2}} \right) \right] \, dy
\]
In order to prove Proposition 3, we express the CCDF as

\[ 4.1.2 \text{ Trivariate Case} \]

\[ \eta \text{ both monotonically increasing in } (A, \alpha, \beta) \]

By the Triangle Inequality [8, Sec. 2.4], we have

\[ \text{Proposition 2 follows now from (21) and (22) by noting that if } \]

\[ \text{where } \]

\[ \Delta_2(A, \alpha, \beta) \triangleq \frac{\theta}{\sqrt{2\pi(1-\phi^2)}} \int_{-\beta}^{0} \phi_{0,1}(y) \eta(y) \frac{dy}{A} - \int_{-\beta}^{0} \phi_{0,1}(y) \delta \left( \frac{\phi}{\sqrt{1-\phi^2}} \right) \frac{dy}{A} - \frac{1}{2} \delta(\beta A). \]  

(20)

It is shown in Appendix A.1 that \( \Delta_2(A, \alpha, \beta) \) satisfies

\[ |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]

where \( \eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|) \) is monotonically increasing in \((A, |\alpha|, |\beta|)\) and is bounded for every finite \(A, \alpha, \) and \(\beta.\)

Combining (13), (15), (18), and (19) yields

\[ \int_{-\alpha}^{\infty} \int_{-\beta}^{\infty} \phi_{0,k}(x, y, z) \frac{dy}{A} \frac{dx}{A} = \frac{1}{4} + \frac{1}{2\pi} \arcsin(\phi) + \frac{\alpha + \beta}{2 \sqrt{2\pi}} + \Delta(A, \alpha, \beta) \]

(21)

where

\[ \Delta(A, \alpha, \beta) \triangleq \Delta_1(A, \alpha, \beta) + \Delta_2(A, \alpha, \beta). \]

By the Triangle Inequality [8, Sec. 2.4], we have

\[ |\Delta(A, \alpha, \beta)| \leq |\Delta_1(A, \alpha, \beta)| + |\Delta_2(A, \alpha, \beta)| \]

\[ \leq A^2 (\eta_1(A, \alpha, \beta) + \eta_2(A, \alpha, \beta)). \]

(22)

Proposition 2 follows now from (21) and (22) by noting that if \( \eta_1(A, \alpha, \beta) \) and \( \eta_2(A, \alpha, \beta) \) are both monotonically increasing in \((A, |\alpha|, |\beta|)\), then so is

\[ \eta(A, \alpha, \beta) \triangleq \eta_1(A, \alpha, \beta) + \eta_2(A, \alpha, \beta); \]

and if \( \eta_1(A, \alpha, \beta) \) and \( \eta_2(A, \alpha, \beta) \) are bounded, then so is \( \eta(A, \alpha, \beta). \)

4.1.2 Trivariate Case

In order to prove Proposition 3, we express the CCDF as

\[ \int_{-\alpha}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} \]

\[ = \int_{0}^{\infty} \int_{0}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} + \int_{-\alpha}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} \]

\[ + \int_{0}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} + \int_{-\alpha}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} \]

\[ = \frac{1}{8} + \frac{1}{4\pi} (\arcsin(g_{12}) + \arcsin(g_{13}) + \arcsin(g_{23})) + \int_{-\alpha}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} \]

\[ + \int_{0}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} + \int_{-\alpha}^{\infty} \int_{-\beta}^{\infty} \int_{-\gamma}^{\infty} \phi_{0,k}(x, y, z) \frac{dz}{A} \frac{dy}{A} \frac{dx}{A} \]

(23)

where the last step follows from the expression for the orthant probability of a trivariate, zero-mean Gaussian vector of covariance matrix [14] (see also [11])

\[ K = \begin{pmatrix}
1 & g_{12} & g_{13} \\
g_{12} & 1 & g_{23} \\
g_{13} & g_{23} & 1
\end{pmatrix}. \]
If \( A = 0 \), then the integrals on the RHS of (23) are zero and Proposition 3 follows directly from (23). To prove Proposition 3 for \( A > 0 \), we continue by evaluating the integrals on the RHS of (23) separately. To evaluate the first integral, we express the joint PDF \( (x, y, z) \mapsto \phi_{0,K}(x, y, z) \) as the product of a marginal PDF \( f(x) \) and a conditional PDF \( f(y, z|x) \)

\[
\phi_{0,K}(x, y, z) = \phi_{0,1}(x)\phi_{\mu(x),K}(y, z), \quad (x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R})
\]

where \( (y, z) \mapsto \phi_{\mu,K}(\cdot)(y, z) \) denotes the PDF of a bivariate Gaussian vector of mean \( \mu \) and covariance matrix \( K \), and where

\[
\mu(x) = \begin{pmatrix} \theta_{12}x \\ \theta_{13}x \end{pmatrix} \quad \text{and} \quad K(x) = \begin{pmatrix} 1 - \theta_{12}^2 & \theta_{23} - \theta_{12}\theta_{13} \\ \theta_{23} - \theta_{12}\theta_{13} & 1 - \theta_{13}^2 \end{pmatrix}.
\]

We thus have

\[
\begin{align*}
\int_{-\alpha A}^{0} \int_{-\beta A}^{0} \int_{-\gamma A}^{\infty} \phi_{0,K}(x, y, z) \, dz \, dy \, dx \\
= \int_{-\alpha A}^{0} \phi_{0,1}(x) \int_{-\beta A}^{0} \int_{-\gamma A}^{\infty} \phi_{\mu(x),K}(y, z) \, dz \, dy \, dx \\
= \int_{-\alpha A}^{0} \phi_{0,1}(x) \int_{-\beta A}^{\infty} \int_{-\gamma A}^{\infty} \phi_{0,K'}(y', z') \, dy' \, dz' \, dx \\
= \int_{-\alpha A}^{0} \phi_{0,1}(x) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{23} - \theta_{12}\theta_{13}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] \, dx \\
+ \int_{-\alpha A}^{0} \phi_{0,1}(x) \frac{1}{2\sqrt{2\pi}} \frac{\beta A + \theta_{12}x}{\sqrt{1 - \theta_{12}^2}} + \frac{\gamma A + \theta_{13}x}{\sqrt{1 - \theta_{13}^2}} \, dx \\
+ \int_{-\alpha A}^{0} \phi_{0,1}(x) \Delta \left( A, \frac{\beta + \theta_{12}x}{\sqrt{1 - \theta_{12}^2}}, \frac{\gamma + \theta_{13}x}{\sqrt{1 - \theta_{13}^2}} \right) \, dx \\
= \left( \int_{-\alpha A}^{0} \phi_{0,1}(x) \, dx + \Delta(\alpha A) \right) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{23} - \theta_{12}\theta_{13}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] + \Delta_1(\alpha, \beta, \gamma) \quad (24)
\end{align*}
\]

where \( \Delta(\alpha, \beta, \gamma) \) and \( \delta(\cdot) \) are as in the previous section, and where

\[
K'(x) = \begin{pmatrix} 1 - \theta_{12}^2 & \theta_{23} - \theta_{12}\theta_{13} \\ \theta_{23} - \theta_{12}\theta_{13} & 1 - \theta_{13}^2 \end{pmatrix}
\]

and

\[
\Delta_1(\alpha, \beta, \gamma) = \delta(\alpha A) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{23} - \theta_{12}\theta_{13}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] \\
+ \int_{-\alpha A}^{0} \phi_{0,1}(x) \frac{1}{2\sqrt{2\pi}} \frac{\beta A + \theta_{12}x}{\sqrt{1 - \theta_{12}^2}} + \frac{\gamma A + \theta_{13}x}{\sqrt{1 - \theta_{13}^2}} \, dx \\
+ \int_{-\alpha A}^{0} \phi_{0,1}(x) \Delta \left( A, \frac{\beta + \theta_{12}x}{\sqrt{1 - \theta_{12}^2}}, \frac{\gamma + \theta_{13}x}{\sqrt{1 - \theta_{13}^2}} \right) \, dx. \quad (25)
\]

Here the second step follows by substituting

\[
y' = \frac{y - \theta_{12}x}{\sqrt{1 - \theta_{12}^2}} \quad \text{and} \quad z' = \frac{z - \theta_{13}x}{\sqrt{1 - \theta_{13}^2}}
\]

and the third step follows from Proposition 2.
By applying (18), we obtain for the first term on the RHS of (24)

\[
\left( \int_{-\alpha A}^{0} \phi_{0,1}(x) \, dx + \delta(\alpha A) \right) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{23} - \theta_{12} \theta_{13}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] = \frac{A}{\sqrt{2\pi}} \left[ \frac{\alpha}{4} + \frac{\alpha}{2\pi} \arcsin \left( \frac{\theta_{23} - \theta_{12} \theta_{13}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right].
\]  

(26)

It is shown in Appendix A.2 that \( \Delta_1(A, \alpha, \beta, \gamma) \) satisfies

\[
|\Delta_1(A, \alpha, \beta, \gamma)| \leq A^2 \eta_1(A, \alpha, \beta, \gamma)
\]

where \( \eta_1(A, \alpha, \beta, \gamma) = \eta_1(A, |\alpha|, |\beta|, |\gamma|) \) is monotonically increasing in \((A, |\alpha|, |\beta|, |\gamma|)\) and is bounded for every finite \( A, \alpha, \beta, \) and \( \gamma \).

To evaluate the second integral on the RHS of (23), we express the joint PDF \((x, y, z) \mapsto \phi_{0,k}(x, y, z)\) as

\[
\phi_{0,k}(x, y, z) = \phi_{0,1}(y) \phi_{\mu(y),K(y)}(x, z), \quad (x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R})
\]

where

\[
\mu(y) = \left( \begin{array}{c} \theta_{12} y \\ \theta_{23} y \end{array} \right) \quad \text{and} \quad K(y) = \left( \begin{array}{ccc} 1 - \theta_{12}^2 & \theta_{13} - \theta_{12} \theta_{23} & 1 - \theta_{23}^2 \\ \theta_{13} - \theta_{12} \theta_{23} & 1 - \theta_{23}^2 & \theta_{12}^2 \\ \theta_{12}^2 & \theta_{13} - \theta_{12} \theta_{23} & 1 - \theta_{12}^2 \end{array} \right).
\]

We thus have

\[
\int_{0}^{\infty} \int_{-\gamma A}^{0} \int_{-\gamma A}^{0} \phi_{0,k}(x, y, z) \, dz \, dy \, dx

= \int_{0}^{\infty} \phi_{0,1}(y) \int_{0}^{\infty} \int_{-\gamma A}^{0} \phi_{\mu(y),K(y)}(x, z) \, dz \, dx \, dy

= \int_{0}^{\infty} \phi_{0,1}(y) \int_{-\gamma A}^{0} \int_{-\gamma A}^{0} \phi_{0,k}(x', z') \, dx' \, dz' \, dy

= \int_{0}^{\infty} \phi_{0,1}(y) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{13} - \theta_{12} \theta_{23}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] \, dy

+ \int_{0}^{\infty} \phi_{0,1}(y) \frac{1}{2\sqrt{2\pi}} \left( \frac{\theta_{12} y}{\sqrt{1 - \theta_{12}^2}} + \frac{\gamma A + \theta_{23} y}{\sqrt{1 - \theta_{23}^2}} \right) \, dy

+ \int_{0}^{\infty} \phi_{0,1}(y) \Delta \left( A, \frac{y_{12} y/A}{\sqrt{1 - \theta_{12}^2}}, \frac{y_{13} y/A}{\sqrt{1 - \theta_{13}^2}} \right) \, dy

= \left( \int_{0}^{\infty} \phi_{0,1}(y) \, dy + \delta(\beta A) \right) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{13} - \theta_{12} \theta_{23}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] + \Delta_2(A, \alpha, \beta, \gamma)

= \frac{A}{\sqrt{2\pi}} \left[ \frac{\beta}{4} + \frac{\beta}{2\pi} \arcsin \left( \frac{\theta_{13} - \theta_{12} \theta_{23}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right] + \Delta_2(A, \alpha, \beta, \gamma)
\]

(27)

where

\[
K'(y) = \left( \begin{array}{c} 1 \\ \frac{\theta_{13} - \theta_{12} \theta_{23}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \\ 1 \end{array} \right)
\]

and

\[
\Delta_2(A, \alpha, \beta, \gamma) \triangleq -\delta(\beta A) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{13} - \theta_{12} \theta_{23}}{\sqrt{(1 - \theta_{12}^2)(1 - \theta_{13}^2)}} \right) \right]

+ \int_{-\alpha A}^{0} \phi_{0,1}(y) \frac{1}{2\sqrt{2\pi}} \left( \frac{\theta_{12} y}{\sqrt{1 - \theta_{12}^2}} + \frac{\gamma A + \theta_{23} y}{\sqrt{1 - \theta_{23}^2}} \right) \, dy

+ \int_{-\alpha A}^{0} \phi_{0,1}(y) \Delta \left( A, \frac{y_{12} y/A}{\sqrt{1 - \theta_{12}^2}}, \frac{y_{13} y/A}{\sqrt{1 - \theta_{13}^2}} \right) \, dy.
\]

(28)
Here the second step follows by substituting
\[ x' = \frac{x - \vartheta_{12} y}{\sqrt{1 - \vartheta_{12}^2}} \quad \text{and} \quad z' = \frac{z - \vartheta_{23} y}{\sqrt{1 - \vartheta_{23}^2}}, \]
the third step follows from Proposition 2; and the last step follows from (18). It is shown in Appendix A.2 that \( \Delta_2(A, \alpha, \beta, \gamma) \) satisfies
\[
|\Delta_2(A, \alpha, \beta, \gamma)| \leq A^2 \eta_2(A, \alpha, \beta, \gamma)
\]
where \( \eta_2(A, \alpha, \beta, \gamma) = \eta_2(A, |\alpha|, |\beta|, |\gamma|) \) is monotonically increasing in \( (A, |\alpha|, |\beta|, |\gamma|) \) and is bounded for every finite \( A, \alpha, \beta, \) and \( \gamma \).

To evaluate the third integral on the RHS of (23), we express the joint PDF \((x, y, z) \mapsto \phi_{0,K}(x, y, z)\) as
\[
\phi_{0,K}(x, y, z) = \phi_{0,1}(z) \phi_{\mu(z), K(z)}(x, y), \quad (x \in \mathbb{R}, y \in \mathbb{R}, z \in \mathbb{R})
\]
where
\[
\mu(z) = \left( \frac{\vartheta_{13} z}{\vartheta_{23} z} \right) \quad \text{and} \quad K(z) = \left( \begin{array}{ccc}
1 - \vartheta_{13}^2 & \vartheta_{12} - \vartheta_{13} \vartheta_{23} \\
\vartheta_{12} - \vartheta_{13} \vartheta_{23} & 1
\end{array} \right).
\]
We have
\[
\int_0^\infty \int_0^\infty \int_{-\gamma_A}^0 \phi_{0,K}(x, y, z) \, dz \, dy \, dx
\]
\[
= \int_{-\gamma_A}^0 \phi_{0,1}(z) \int_0^\infty \int_0^\infty \phi_{\mu(z), K(z)}(x, y) \, dy \, dx \, dz
\]
\[
= \int_{-\gamma_A}^0 \phi_{0,1}(z) \int_{-\vartheta_{13} z}^{\vartheta_{13} z} \int_{-\vartheta_{23} z}^{\vartheta_{23} z} \phi_{0,K'(z)}(x', y') \, dy' \, dx' \, dz
\]
\[
= \int_{-\gamma_A}^0 \phi_{0,1}(z) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\vartheta_{12} - \vartheta_{13} \vartheta_{23}}{\sqrt{(1 - \vartheta_{13}^2)(1 - \vartheta_{23}^2)}} \right) \right] \, dz
\]
\[
+ \int_{-\gamma_A}^0 \phi_{0,1}(z) \frac{1}{2\sqrt{2\pi}} \left( \frac{\vartheta_{13} z}{\sqrt{1 - \vartheta_{13}^2}} + \frac{\vartheta_{23} z}{\sqrt{1 - \vartheta_{23}^2}} \right) \, dz
\]
\[
+ \int_{-\gamma_A}^0 \phi_{0,1}(z) \Delta \left( A, \sqrt{\vartheta_{13}^2 / \sqrt{1 - \vartheta_{13}^2}}, \sqrt{\vartheta_{23}^2 / \sqrt{1 - \vartheta_{23}^2}} \right) \, dz
\]
\[
= \left( \int_{-\gamma_A}^0 \phi_{0,1}(z) \, dz + \delta(\gamma A) \right) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\vartheta_{12} - \vartheta_{13} \vartheta_{23}}{\sqrt{(1 - \vartheta_{13}^2)(1 - \vartheta_{23}^2)}} \right) \right] + \Delta_3(A, \alpha, \beta, \gamma)
\]
\[
= \frac{A}{\sqrt{2\pi}} \left[ \frac{\gamma}{4} + \frac{\gamma}{2\pi} \arcsin \left( \frac{\vartheta_{12} - \vartheta_{13} \vartheta_{23}}{\sqrt{(1 - \vartheta_{13}^2)(1 - \vartheta_{23}^2)}} \right) \right] + \Delta_3(A, \alpha, \beta, \gamma)
\]
where
\[
K'(z) = \left( \begin{array}{c}
\frac{1}{\sqrt{(1 - \vartheta_{13}^2)(1 - \vartheta_{23}^2)}}
\frac{\vartheta_{12} - \vartheta_{13} \vartheta_{23}}{\sqrt{(1 - \vartheta_{13}^2)(1 - \vartheta_{23}^2)}}
\end{array} \right)
\]
and
\[
\Delta_3(A, \alpha, \beta, \gamma) \equiv -\delta(\gamma A) \left[ \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\vartheta_{12} - \vartheta_{13} \vartheta_{23}}{\sqrt{(1 - \vartheta_{13}^2)(1 - \vartheta_{23}^2)}} \right) \right]
\]
\[
+ \int_{-\gamma_A}^0 \phi_{0,1}(z) \frac{1}{2\sqrt{2\pi}} \left( \frac{\vartheta_{13} z}{\sqrt{1 - \vartheta_{13}^2}} + \frac{\vartheta_{23} z}{\sqrt{1 - \vartheta_{23}^2}} \right) \, dz
\]
\[
+ \int_{-\gamma_A}^0 \phi_{0,1}(z) \Delta \left( A, \sqrt{\vartheta_{13}^2 / \sqrt{1 - \vartheta_{13}^2}}, \sqrt{\vartheta_{23}^2 / \sqrt{1 - \vartheta_{23}^2}} \right) \, dz.
\]
Here the second step follows by substituting
\[ x' = \frac{x - \vartheta_{13} z}{\sqrt{1 - \vartheta_{13}^2}} \quad \text{and} \quad y' = \frac{y - \vartheta_{23} z}{\sqrt{1 - \vartheta_{23}^2}}, \]
the third step follows from Proposition 2; and the last step follows from (18). It is shown in Appendix A.2 that \( \Delta_3(A, \alpha, \beta, \gamma) \) satisfies
\[ |\Delta_3(A, \alpha, \beta, \gamma)| \leq A^2 \eta_3(A, \alpha, \beta, \gamma) \]
where \( \eta_3(A, \alpha, \beta, \gamma) = \eta_3(A, |\alpha|, |\beta|, |\gamma|) \) is monotonically increasing in \( (A, |\alpha|, |\beta|, |\gamma|) \) and is bounded for every finite \( A, \alpha, \beta, \) and \( \gamma \).

Combining (23)–(29) yields
\[
\int_{-a}^{\infty} \int_{-b}^{\infty} \int_{-c}^{\infty} \phi_{0, K}(x, y, z) \, dz \, dy \, dx \\
= \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin(\vartheta_{12}) + \arcsin(\vartheta_{13}) + \arcsin(\vartheta_{23}) \right) \\
+ \frac{A}{2\pi} \left[ \frac{\alpha + \beta + \gamma}{4} + \frac{\alpha}{2\pi} \arcsin \left( \frac{\vartheta_{23} - \vartheta_{12} \vartheta_{13}}{\sqrt{(1 - \vartheta_{12}^2)(1 - \vartheta_{13}^2)}} \right) \\
+ \frac{\beta}{2\pi} \arcsin \left( \frac{\vartheta_{13} - \vartheta_{12} \vartheta_{23}}{\sqrt{(1 - \vartheta_{12}^2)(1 - \vartheta_{23}^2)}} \right) + \frac{\gamma}{2\pi} \arcsin \left( \frac{\vartheta_{12} - \vartheta_{13} \vartheta_{23}}{\sqrt{(1 - \vartheta_{12}^2)(1 - \vartheta_{23}^2)}} \right) \right] \]
\[ + \Delta(A, \alpha, \beta, \gamma) \quad (31) \]
where
\[ \Delta(A, \alpha, \beta, \gamma) \triangleq \Delta_1(A, \alpha, \beta, \gamma) + \Delta_2(A, \alpha, \beta, \gamma) + \Delta_3(A, \alpha, \beta, \gamma). \]

By the Triangle Inequality, we have
\[ |\Delta(A, \alpha, \beta, \gamma)| \leq |\Delta_1(A, \alpha, \beta, \gamma)| + |\Delta_2(A, \alpha, \beta, \gamma)| + |\Delta_3(A, \alpha, \beta, \gamma)| \]
\[ \leq A^2 (\eta_1(A, \alpha, \beta, \gamma) + \eta_2(A, \alpha, \beta, \gamma) + \eta_3(A, \alpha, \beta, \gamma)). \]  
(32)

Proposition 3 follows now from (31) and (32) by noting that if \( \eta_1(A, \alpha, \beta, \gamma), \eta_2(A, \alpha, \beta, \gamma), \) and \( \eta_3(A, \alpha, \beta, \gamma) \) are monotonically increasing in \( (A, |\alpha|, |\beta|, |\gamma|) \), then so is
\[ \eta(A, \alpha, \beta, \gamma) \triangleq \eta_1(A, \alpha, \beta, \gamma) + \eta_2(A, \alpha, \beta, \gamma) + \eta_3(A, \alpha, \beta, \gamma); \]
and if \( \eta_1(A, \alpha, \beta, \gamma), \eta_2(A, \alpha, \beta, \gamma), \) and \( \eta_3(A, \alpha, \beta, \gamma) \) are bounded, then so is \( \eta(A, \alpha, \beta, \gamma) \).

4.2 Proof of Theorem 1

To prove Theorem 1, we derive a lower bound on \( C_+(P) \) and compute its ratio to \( P \) in the limit as \( P \) tends to zero. To this end, we evaluate \( (2W)/n I(X^k; Y^k) \) for \( \{X_k, k \in \mathbb{Z}\} \) being a sequence of IID, binary random variables with
\[ X_k = \begin{cases} \sqrt{P}, & \text{with probability } \frac{1}{2} \\ -\sqrt{P}, & \text{with probability } \frac{1}{2} \end{cases} \]

We shall restrict ourselves to waveforms \( g(\cdot) \) that satisfy
\[ \sum_{\ell \neq 0} \left| g \left( \frac{\ell - 1/2}{2W} \right) \right| < \infty \]  
(33)
\[ \sum_{\ell \neq 0} \left| g \left( \frac{\ell}{2W} \right) \right| < \infty \]  
(34)
\[ \sum_{\ell \neq 0} \left| g \left( \frac{\ell + 1/2}{2W} \right) \right| < \infty. \]  
(35)
By the chain rule for mutual information [9, Thm. 2.5.2]

$$\frac{2W}{n} I(X^n_1; Y^n_1) = \frac{2W}{n} \sum_{k=1}^{n} I(X_k; Y_k \mid X^{k-1}_1)$$

$$\geq \frac{2W}{n} I(X_k; Y_k)$$

$$= \frac{2W}{n} \sum_{k=1}^{n} I(X_k; Y_{k-\frac{1}{2}}, Y_k, Y_{k+\frac{1}{2}})$$

$$= 2W I(X_1; Y_1, Y^-_2)$$

where we define

$$Y_\tau \triangleq Y \left( \frac{\tau}{2W} \right), \quad \tau \in \mathbb{R}.$$

Here the second step follows because reducing observations cannot increase mutual information and because \{X_k, k \in \mathbb{Z}\} is IID; the third step follows from the definition of \(Y_k\); and the last step follows because the joint law of \((X_k, Y_{k-\frac{1}{2}}, Y_k, Y_{k+\frac{1}{2}})\) does not depend on \(k\).

### 4.2.1 The Joint Law of \((X_1, Y^-_2, Y_1, Y^+_2)\)

In order to evaluate \(I(X_1; Y^-_2, Y_1, Y^+_2)\), we shall compute the conditional probability of \((Y^-_2, Y_1, Y^+_2)\), conditioned on \(X_1\). To this end, we first compute the conditional probability of \((Y^-_2, Y_1, Y^+_2)\), conditioned on \(X_{-\infty}^\infty\), and then average over \(X_{-\infty}^\infty\).

To compute \(\Pr(Y^-_2 = 1, Y_1 = 1, Y^+_2 = 1 \mid X_{-\infty}^\infty = x_{-\infty}^\infty)\), we first note that by (4)

$$Y_\tau = 1 \iff \frac{1}{\sqrt{2W}} \sum_{\ell=-\infty}^{\infty} x_\ell g \left( \frac{\tau - \ell}{2W} \right) + Z_\tau \geq 0, \quad \tau \in \mathbb{R} \quad (37)$$

where we define

$$Z_\tau \triangleq (Z \ast \LPF_W) \left( \frac{\tau}{2W} \right), \quad \tau \in \mathbb{R}.$$

Let

$$\alpha(x_{-\infty}^\infty) \triangleq \frac{1}{\sqrt{\Pr(\mathbb{W}^0) \circ \mathbb{W}^0)} \sum_{\ell=-\infty}^{\infty} x_\ell g \left( \frac{1/2 - \ell}{2W} \right)$$

$$\beta(x_{-\infty}^\infty) \triangleq \frac{1}{\sqrt{\Pr(\mathbb{W}^0) \circ \mathbb{W}^0)} \sum_{\ell=-\infty}^{\infty} x_\ell g \left( \frac{1 - \ell}{2W} \right)$$

and

$$\gamma(x_{-\infty}^\infty) \triangleq \frac{1}{\sqrt{\Pr(\mathbb{W}^0) \circ \mathbb{W}^0)} \sum_{\ell=-\infty}^{\infty} x_\ell g \left( \frac{3/2 - \ell}{2W} \right).$$

It follows from (37) that

$$\Pr(Y^-_2 = 1, Y_1 = 1, Y^+_2 = 1 \mid X_{-\infty}^\infty = x_{-\infty}^\infty)$$

$$= \Pr \left( Z_\tau^+ \geq -\alpha(x_{-\infty}^\infty) \sqrt{\Pr(\mathbb{W}^0)}, Z_1 \geq -\beta(x_{-\infty}^\infty) \sqrt{\Pr(\mathbb{W}^0)}, Z_\tau^- \geq -\gamma(x_{-\infty}^\infty) \sqrt{\Pr(\mathbb{W}^0)} \right)$$

$$= \Pr \left( \frac{1}{\sqrt{\Pr(\mathbb{W}^0)}}, \frac{1}{\sqrt{\Pr(\mathbb{W}^0)}}, \frac{1}{\sqrt{\Pr(\mathbb{W}^0)}} \right)$$

$$= \frac{1}{8} + \frac{1}{2\pi} \arcsin(\rho) + \sqrt{\frac{2\pi}{\rho}} \left[ \alpha(x_{-\infty}^\infty) + \beta(x_{-\infty}^\infty) + \gamma(x_{-\infty}^\infty) \right. $$

$$\left. + \frac{\alpha(x_{-\infty}^\infty) + \gamma(x_{-\infty}^\infty)}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1-\rho^2}} \right) - \frac{\beta(x_{-\infty}^\infty)}{2\pi} \arcsin \left( \frac{\rho^2}{1-\rho^2} \right) \right]$$

$$+ \Delta \left( \sqrt{\Pr}, \alpha(x_{-\infty}^\infty), \beta(x_{-\infty}^\infty), \gamma(x_{-\infty}^\infty) \right) \quad (38)$$
where
\[ K = \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}, \quad \rho = \frac{2}{\pi}. \]

Here the third step follows by noting that \( \{Z_{\tau}, \tau \in \mathbb{R}\} \) is a zero-mean Gaussian process of autocovariance function \( \tau \mapsto WN_0 \text{sinc}(\tau); \) and the last step follows from Proposition 3.

Since \( \{X_k, k \in \mathbb{Z}\} \) is IID and of zero mean, we have
\[
\alpha_0 \triangleq \mathbb{E} \left[ \alpha(X_{\infty}^1) \bigg| X_1 = \sqrt{P} \right] = \frac{1}{\sqrt{2W}(WN_0)} g \left( -\frac{1}{4W} \right) + \frac{1}{\sqrt{2P}(2W)(WN_0)} \sum_{\ell \neq 1} \mathbb{E}[X_\ell] g \left( \frac{1/2 - \ell}{2W} \right) = \frac{1}{\sqrt{2W}(WN_0)} g \left( -\frac{1}{4W} \right)
\]

and
\[
\beta_0 \triangleq \mathbb{E} \left[ \beta(X_{\infty}^2) \bigg| X_1 = \sqrt{P} \right] = \frac{1}{\sqrt{2W}(WN_0)} g \left( 0 \right),
\]

\[
\gamma_0 \triangleq \mathbb{E} \left[ \gamma(X_{\infty}^3) \bigg| X_1 = \sqrt{P} \right] = \frac{1}{\sqrt{2W}(WN_0)} g \left( \frac{1}{4W} \right).
\]

By setting \( X_1 = \sqrt{P} \) and averaging over \( (X_{\infty}^0, X_{\infty}^2) \), we thus obtain
\[
\Pr(Y_\omega = 1, Y_\omega = 1, Y_{\omega} = 1 \bigg| X_1 = \sqrt{P}) = \frac{1}{\delta} + \frac{1}{2\pi} \arcsin(\rho) + \mathbb{E} \left[ \sqrt{\frac{P}{2\pi}} \left[ \alpha(X_{\infty}^1) + \beta(X_{\infty}^2) + \gamma(X_{\infty}^3) \right] + \frac{\alpha(X_{\infty}^1) + \gamma(X_{\infty}^3)}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \beta(X_{\infty}^2) \frac{\arcsin \left( \frac{\rho^2}{1 - \rho^2} \right)}{2\pi} \right] \bigg| X_1 = \sqrt{P}
\]
\[
+ \mathbb{E} \left[ \Delta \left( \sqrt{\frac{P}{2\pi}}, \alpha(X_{\infty}^1), \beta(X_{\infty}^2), \gamma(X_{\infty}^3) \right) \bigg| X_1 = \sqrt{P} \right]
\]
\[
= \frac{1}{\delta} + \frac{1}{2\pi} \arcsin(\rho) + \mathbb{E} \left[ \sqrt{\frac{P}{2\pi}} \left[ \frac{\alpha_0 + \beta_0 + \gamma_0}{4} + \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \beta_0 \frac{\arcsin \left( \frac{\rho^2}{1 - \rho^2} \right)}{2\pi} \right] \right. \bigg| X_1 = \sqrt{P}
\]
\[
+ \mathbb{E} \left[ \Delta \left( \sqrt{\frac{P}{2\pi}}, \alpha(X_{\infty}^1), \beta(X_{\infty}^2), \gamma(X_{\infty}^3) \right) \bigg| X_1 = \sqrt{P} \right].
\]

We next show that
\[
\mathbb{E} \left[ \Delta \left( \sqrt{\frac{P}{2\pi}}, \alpha(X_{\infty}^1), \beta(X_{\infty}^2), \gamma(X_{\infty}^3) \right) \bigg| X_1 = \sqrt{P} \right] = o(\sqrt{P})
\]
where \( x \mapsto o(x) \) satisfies \( \lim_{x \downarrow 0} o(x)/x = 0 \). To this end, we use the Triangle Inequality to upper bound
\[
|\alpha(x_{\infty}^1)| \leq \frac{1}{\sqrt{2P}(2W)(WN_0)} \sum_{\ell = -\infty}^{\infty} |x_{\ell}| \left| g \left( \frac{1/2 - \ell}{2W} \right) \right| = \frac{1}{\sqrt{2W}(WN_0)} \sum_{\ell = -\infty}^{\infty} \left| g \left( \frac{1/2 - \ell}{2W} \right) \right| \triangleq o_{\max}
\]
which, by (33), is finite. Here the second step follows because, for our choice of \( \{ X_k, k \in \mathbb{Z} \} \), we have with probability one \( |X_k| = \sqrt{P} \). Similarly, we have

\[
|\beta(x^{\infty}_k)| \leq \frac{1}{\sqrt{(2W)(W N_0)}} \sum_{\ell = -\infty}^{\infty} |g \left( \frac{1 - \ell}{2W} \right)|
\]

\[\triangleq \beta_{\max} < \infty \]

and

\[
|\gamma(x^{\infty}_k)| \leq \frac{1}{\sqrt{(2W)(W N_0)}} \sum_{\ell = -\infty}^{\infty} |g \left( \frac{3/2 - \ell}{2W} \right)|
\]

\[\triangleq \gamma_{\max} < \infty.\]

We thus obtain

\[
\mathbb{E} \left[ \Delta \left( \sqrt{P}, \alpha(X^{\infty}_k), \beta(X^{\infty}_k), \gamma(X^{\infty}_k) \right) \mid X_1 = \sqrt{P} \right]
\]

\[
\leq \mathbb{E} \left[ \left\| \Delta \left( \sqrt{P}, \alpha(X^{\infty}_k), \beta(X^{\infty}_k), \gamma(X^{\infty}_k) \right) \right\| \mid X_1 = \sqrt{P} \right]
\]

\[
\leq \mathbb{P} \mathbb{E} \left[ \eta \left( \sqrt{P}, \alpha(X^{\infty}_k), \beta(X^{\infty}_k), \gamma(X^{\infty}_k) \right) \mid X_1 = \sqrt{P} \right]
\]

\[
\leq \mathbb{P} \eta \left( \sqrt{P}, \alpha_{\max}, \beta_{\max}, \gamma_{\max} \right)
\]

\[= o(\sqrt{P}) \tag{44}\]

where the first step follows from the Triangle Inequality; the second step follows because \( \Delta(A, \alpha, \beta, \gamma) \leq K^2 \eta(A, \alpha, \beta, \gamma) \); the third step follows because \( \eta(A, \alpha, \beta, \gamma) = \eta(A, |\alpha|, |\beta|, |\gamma|) \) is monotonically increasing in \((A, |\alpha|, |\beta|, |\gamma|)\); and the last step follows because \( \eta(A, \alpha, \beta, \gamma) \) is bounded for finite \( A, \alpha, \beta, \) and \( \gamma \). This proves (43).

Combining (42) and (43) yields

\[
\Pr(Y_2 = 1, Y_1 = 1, Y_2 = 1 \mid X_1 = \sqrt{P})
\]

\[= \frac{1}{8} + \frac{1}{2\pi} \arcsin(\rho)
\]

\[+ \sqrt{\frac{P}{2\pi}} \left( \frac{\alpha_0 + \beta_0 + \gamma_0}{4} + \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right) + o(\sqrt{P}) \tag{45}\]

By averaging (38) over \( X^{\infty}_k \), and by noting that

\[\mathbb{E} \left[ \alpha(X^{\infty}_k) \right] = \mathbb{E} \left[ \beta(X^{\infty}_k) \right] = \mathbb{E} \left[ \gamma(X^{\infty}_k) \right] = 0 \tag{46}\]

and

\[\mathbb{E} \left[ \Delta \left( \sqrt{P}, \alpha(X^{\infty}_k), \beta(X^{\infty}_k), \gamma(X^{\infty}_k) \right) \right] = o(\sqrt{P}) \]

we obtain for the unconditional probability

\[
\Pr(Y_2 = 1, Y_1 = 1, Y_2 = 1) = \frac{1}{8} + \frac{1}{2\pi} \arcsin(\rho) + o(\sqrt{P}) \tag{47}\]

It thus follows from Bayes’ law that

\[
\Pr(X_1 = \sqrt{P} \mid Y_2 = 1, Y_1 = 1, Y_2 = 1)
\]

\[= \frac{1}{2} + \sqrt{\frac{P}{2\pi}} \left( \frac{\alpha_0 + \beta_0 + \gamma_0}{4} + \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right) + o(\sqrt{P}) \]

\[= \frac{1}{2} + \sqrt{\frac{P}{2\pi}} \left( \frac{\alpha_0 + \beta_0 + \gamma_0}{4} + \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right) + o(\sqrt{P}) \tag{48}\]
Thus, if \( \frac{\partial^2 o(x)}{\xi^2} = \frac{\partial o(x)}{\xi} \) for any \( \xi \neq 0 \) and \( \theta \in \mathbb{R} \). Note that the first two terms on the RHS of (48) depend only on \( X_{-1}, X_0, X_2, X_3, \ldots \) and via their means. Thus, if \( \{X_k, k \in \mathbb{Z}\} \) is of zero mean, then intersymbol interference affects only the \( o(\sqrt{P}) \)-term, which does not influence our lower bound on the capacity per unit-cost.

The probability \( \Pr(Y_2 = 1, Y_1 = 1, Y_3 = -1 \mid X_{-\infty} = x_{-\infty}) \) can be computed in a similar way. We have

\[
\Pr(Y_2 = 1, Y_1 = 1, Y_3 = -1 \mid X_{-\infty} = x_{-\infty}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{0,k}(x, y, z) \, dz \, dy \, dx
\]

where \( K \) is the same as in (38), and where

\[
G = \left( \begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right), \quad \rho = \frac{2}{\pi}
\]

Here the second step follows because

\[
\int_{-\infty}^{\infty} \phi_{0,k}(x, y, z) \, dz = \phi_{0,g}(x, y), \quad (x \in \mathbb{R}, y \in \mathbb{R})
\]

and the third step follows from Propositions 2 and 3. By setting \( X_1 = \sqrt{P} \) and averaging over \( (X_{-\infty}, X_{2}) \), we obtain

\[
\Pr(Y_2 = 1, Y_1 = 1, Y_3 = -1 \mid X_1 = \sqrt{P}) = \frac{1}{8} + \sqrt{\frac{P}{2\pi}} \left[ \frac{\alpha(x_{-\infty}) + \beta(x_{-\infty}) - \gamma(x_{-\infty})}{4} \right]
\]

\[
- \frac{\alpha(x_{-\infty}) + \gamma(x_{-\infty})}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta(x_{-\infty})}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right)
\]

\[
+ \Delta \left( \sqrt{P}, \alpha(x_{-\infty}), \beta(x_{-\infty}) \right) - \Delta \left( \sqrt{P}, \alpha(x_{-\infty}), \beta(x_{-\infty}), \gamma(x_{-\infty}) \right)
\]

(49)

where \( \alpha_0, \beta_0, \) and \( \gamma_0 \) are defined in (39)–(41). It was shown above that the last term on the RHS of (50) decays faster than \( \sqrt{P} \) (43). Repeating the steps in (44), it can also be shown that

\[
\mathbb{E} \left[ \Delta \left( \sqrt{P}, \alpha(x_{-\infty}), \beta(x_{-\infty}) \right) \mid X_1 = \sqrt{P} \right] = o(\sqrt{P}).
\]
By noting that \( o(\sqrt{P}) - o(\sqrt{P}) = o(\sqrt{P}) \), we thus obtain

\[
\Pr(Y_\alpha = 1, Y_1 = 1, Y_2 = -1 \mid X_1 = \sqrt{P}) = \frac{1}{8} \\
+ \sqrt{\frac{P}{2\pi}} \left[ \frac{\alpha_0 + \beta_0 - \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin\left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin\left( \frac{\rho^2}{1 - \rho^2} \right) \right] + o(\sqrt{P}).
\]

It follows from (46) that the unconditional probability is given by

\[
\Pr(Y_\alpha = 1, Y_1 = 1, Y_2 = -1) = \frac{1}{8} + o(\sqrt{P}).
\]

By Bayes’ law, we thus have

\[
\Pr(X_1 = \sqrt{P} \mid Y_\alpha = 1, Y_1 = 1, Y_2 = -1) = \frac{1}{2} \\
+ \sqrt{\frac{P}{2\pi}} \left[ \frac{\alpha_0 - \beta_0 + \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin\left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin\left( \frac{\rho^2}{1 - \rho^2} \right) \right] + o(\sqrt{P}).
\]

(51)

The other conditional probabilities can be computed along the same lines. The probabilities corresponding to \((Y_\alpha = 1, Y_1 = -1, Y_2 = 1)\) and \((Y_\alpha = -1, Y_1 = 1, Y_2 = 1)\) are given by

\[
\Pr(X_1 = \sqrt{P} \mid Y_\alpha = 1, Y_1 = -1, Y_2 = 1) \\
= \frac{1}{2} + \sqrt{\frac{P}{2\pi}} \left[ \frac{\alpha_0 - \beta_0 + \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin\left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin\left( \frac{\rho^2}{1 - \rho^2} \right) \right] + o(\sqrt{P}).
\]

(54)

and

\[
\Pr(X_1 = \sqrt{P} \mid Y_\alpha = -1, Y_1 = 1, Y_2 = 1) = \frac{1}{2} \\
+ 4\sqrt{\frac{P}{2\pi}} \left[ \frac{-\alpha_0 + \beta_0 + \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin\left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin\left( \frac{\rho^2}{1 - \rho^2} \right) \right] + o(\sqrt{P}).
\]

(55)

The remaining probabilities can be computed from (48), (53), (54), and (55) by noting that, due to the symmetry of \( \phi_{0,K}(\cdot) \),

\[
\Pr(X_1 = \sqrt{P} \mid Y_\alpha = y_\alpha, Y_1 = y_1, Y_2 = y_2) = \Pr(X_1 = -\sqrt{P} \mid Y_\alpha = y_\alpha, Y_1 = -y_1, Y_2 = -y_2).
\]

4.2.2 Evaluating \( I(X_1; Y_\alpha, Y_1, Y_2) \)

Let \( \varphi(y_\alpha, y_1, y_2) \triangleq \Pr(X_1 = \sqrt{P} \mid Y_\alpha = y_\alpha, Y_1 = y_1, Y_2 = y_2) \).

We express the mutual information \( I(X_1; Y_\alpha, Y_1, Y_2) \) as

\[
I(X_1; Y_\alpha, Y_1, Y_2) = H(X_1) - H(X_1 \mid Y_\alpha, Y_1, Y_2) \\
= \log 2 - \mathbb{E}\left[ H_b\left( \varphi(Y_\alpha, Y_1, Y_2) \right) \right]
\]

(56)

where \( H_b(p) \triangleq -p \log p - (1 - p) \log(1 - p), 0 \leq p \leq 1 \) (with \( 0 \log 0 \triangleq 0 \)) denotes the binary entropy function. To evaluate the binary entropy function, we express \( H_b(\cdot) \) as a Taylor series expansion around \( \frac{1}{2} \), i.e.,

\[
H_b\left( \frac{1}{2} + \xi \right) = \log 2 - 2\xi^2 + o(\xi^2), \quad |\xi| \leq \frac{1}{2}
\]

(57)
Substituting $\xi = \varphi(y_{1y}, y_{2y}) - 1/2$ and averaging over $(Y_1, Y_2, Y_2)$ yields thus

$$I(X_1; Y_2, Y_1, Y_2) = \log 2 - E \left[ \log 2 - 2 \left( \varphi(Y_2, Y_1, Y_2) - \frac{1}{2} \right)^2 + o \left( \left( \varphi(Y_2, Y_1, Y_2) - \frac{1}{2} \right)^2 \right) \right]$$

$$= 2E \left[ \left( \varphi(Y_2, Y_1, Y_2) - \frac{1}{2} \right)^2 \right] + E \left[ o \left( \left( \varphi(Y_2, Y_1, Y_2) - \frac{1}{2} \right)^2 \right) \right]$$

$$= 2E \left[ \left( \varphi(Y_2, Y_1, Y_2) - \frac{1}{2} \right)^2 \right] + o(P)$$

(58)

where the last step follows because, by (48), (53), (54), and (55), we have for every $(y_{1y}, y_{2y})$

$$\left( \varphi(y_{1y}, y_{2y}) - \frac{1}{2} \right)^2 = O(P)$$

where $x \mapsto O(x)$ satisfies $\lim_{x \to 0} |O(x)/x| < \infty$, so

$$o \left( \left( \varphi(y_{1y}, y_{2y}) - \frac{1}{2} \right)^2 \right) = o(P)$$

which implies that

$$E \left[ o \left( \left( \varphi(Y_2, Y_1, Y_2) - \frac{1}{2} \right)^2 \right) \right] = o(P)$$

since the expectation is given by the sum of eight terms that all decay faster than $P$.

By applying (48), (53), (54), and (55) to (58), we obtain

$$I(X_1; Y_2, Y_1, Y_2) = \frac{P}{\pi} \left[ \frac{(\alpha_0 + \beta_0 + \gamma_0) + \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{\beta_0}{\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right)}{\frac{1}{2} + \frac{1}{\pi} \arcsin(\rho)} \right]$$

$$+ 4 \left( \frac{\alpha_0 + \beta_0 - \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2$$

$$+ \left( \frac{\alpha_0 - \beta_0 + \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)$$

$$\left( 1 - \frac{1}{\pi} \arcsin(\rho) \right)$$

(59)

$$+ 4 \left( -\frac{\alpha_0 + \beta_0 + \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2$$

where $\alpha_0$, $\beta_0$, and $\gamma_0$ are defined in (39)–(41).

Combining (59) with (36) and (6), and computing the ratio to $P$ in the limit as $P$ tends to zero, yields the lower bound on the capacity per unit-cost

$$\hat{C}_{\pi \sigma} (0) \geq \frac{2W}{\pi} \left[ \frac{(\alpha_0 + \beta_0 + \gamma_0) + \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{\beta_0}{\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right)}{\frac{1}{2} + \frac{1}{\pi} \arcsin(\rho)} \right]$$

$$+ 4 \left( \frac{\alpha_0 + \beta_0 - \gamma_0}{4} - \frac{\alpha_0 + \gamma_0}{2\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\beta_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2$$
functions whose energy is less than or equal to 1 is convex, since

\[
\frac{(a_0 - \beta_0 + \gamma_0)}{4} - \frac{a_0 + \gamma_0}{2\pi} \arcsin\left(\frac{\rho}{\sqrt{1 - \rho^2}}\right) + \frac{\beta_0}{2\pi} \arcsin\left(\frac{\rho^2}{1 - \rho^2}\right)
\]

\[
+ 4 \left(-a_0 + \beta_0 + \gamma_0 \right) - \frac{a_0 + \gamma_0}{2\pi} \arcsin\left(\frac{\rho}{\sqrt{1 - \rho^2}}\right) + \frac{\beta_0}{2\pi} \arcsin\left(\frac{\rho^2}{1 - \rho^2}\right)
\]

\[
\triangleq \hat{R}(0).
\]

Note that this lower bound holds for all unit-energy waveforms \(g(.)\) that are bandlimited to \(W\) Hz and that satisfy (33)–(35). In the following section we evaluate the RHS of (60) for a specific choice of \(g(.)\).

### 4.2.3 Choosing a Waveform

Any choice of \(g(.)\) satisfying the above conditions yields a lower bound on \(\hat{C}_{\frac{1}{W}}(0)\). Here we shall find the waveform that gives rise to the largest lower bound. Thus, we wish to maximize the RHS of (60) over all unit-energy waveforms \(g(.)\) that are bandlimited to \(W\) Hz and that satisfy (33)–(35), namely,

\[
\sum_{\ell \neq 0} \left| g\left(\frac{\ell - 1/2}{2W}\right) \right| < \infty
\]

\[
\sum_{\ell \neq 0} \left| g\left(\frac{\ell}{2W}\right) \right| < \infty
\]

\[
\sum_{\ell \neq 0} \left| g\left(\frac{\ell + 1/2}{2W}\right) \right| < \infty.
\]

To further facilitate the optimization problem, we choose \(g(.)\) to be a symmetric function, which implies that \(\alpha_0 = \gamma_0\). (Numerical computation suggests that this choice is indeed optimal.) In this case, the rate per unit-cost \(\hat{R}(0)\) is of the form

\[
a\alpha_0^2 + 2b\alpha_0\beta_0 + c\beta_0^2
\]

(where \(a > 0\), \(c > 0\) and \(ac - b^2 > 0\), so \(\hat{R}(0)\) is a convex function of \((\alpha_0, \beta_0)\), and all pairs \((\alpha_0, \beta_0)\) that give rise to the same rate per unit-cost lie on an ellipse. Let \(G\) denote the set of all waveforms satisfying the above conditions, and let \(B\) denote the set of all pairs \((\alpha_0, \beta_0)\) that arise from these waveforms. Since \(\hat{R}(0)\) is convex in \((\alpha_0, \beta_0)\), it follows that the supremum of \(\hat{R}(0)\) over \(B\) lies on the boundary of \(B\), which we shall determine next.

While the set of all bandlimited functions that satisfy (33)–(35) is convex, the set of all unit-energy functions is not (see [15, Sec. 2] for a definition of a convex set). Indeed, if for example the functions \(g_1(.)\) and \(g_2(.)\) are of unit energy and satisfy \(g_1(t) = -g_2(t), \ t \in \mathbb{R}\), then

\[
t \mapsto \xi g_1(t) + (1 - \xi)g_2(t), \quad \xi \in [0, 1]
\]

is of energy \((2\xi - 1)^2\), which is not equal to 1 unless \(\xi = 1\) or \(\xi = 0\). Nevertheless, the set of all functions whose energy is less than or equal to 1 is convex, since

\[
\left\| \xi g_1 + (1 - \xi)g_2 \right\|^2 \leq \left( \xi \left\| g_1 \right\| + (1 - \xi) \left\| g_2 \right\| \right)^2 \leq 1, \quad \left( \left\| g_1 \right\|^2 \leq 1, \left\| g_2 \right\|^2 \leq 1, \xi \in [0, 1] \right)
\]

where \(\left\| g \right\| \triangleq \sqrt{\int (g(t))^2 \, dt}\), and where the inequality follows from the Triangle Inequality. Let \(G' \supseteq G\) denote the set of all waveforms that are bandlimited to \(W\) Hz, that satisfy (33)–(35), and whose energy is less than or equal to 1. Furthermore, let \(B' \supseteq B\) denote the set of all pairs \((\alpha_0, \beta_0)\) that arise from these waveforms. Since convexity is preserved under intersection [15, Thm. 2.1], it follows that the set \(G'\) (and hence also \(B'\)) is convex. We further have that

\[
\sup_{g'} \hat{R}(0) = \sup_{g'} \hat{R}(0) = \sup_{\hat{R}(0)}
\]

18
where the first step follows because every rate per unit-cost \( \dot{R}(0) \) that corresponds to a waveform satisfying \( \| g \| < 1 \) can be increased by normalizing the waveform; and where the second step follows because \( \dot{R}(0) \) depends on \( g'(\cdot) \) only via \( (\alpha_0, \beta_0) \) and because \( g' \) determines \( \mathcal{B}' \). The above optimization problem can thus be expressed as the maximization of a convex function over a convex set.

Let \( \mathcal{B}' \) denote the closure of \( \mathcal{B}' \) (i.e., let \( \mathcal{B}' \) be the smallest closed set containing \( \mathcal{B}' \)) and let \( \lambda \mapsto \zeta(\lambda) \) be defined as

\[
\zeta(\lambda) \triangleq \sup_{(\alpha_0, \beta_0) \in \mathcal{B}'} |\lambda \alpha_0 + \beta_0|, \quad \lambda \in \mathbb{R}.
\]  

(61)

Since \( \mathcal{B}' \) is a closed convex set, it follows that the boundary of \( \mathcal{B}' \) is given by the points \( (\alpha_0, \beta_0) \in \mathcal{B}' \) that achieve the supremum in (61) [15, Sec. 13]. Note that the boundary of \( \mathcal{B}' \) is symmetric with respect to the origin, since \( (\alpha_0, \beta_0) \) and \( (-\alpha_0, -\beta_0) \) yield the same value for \( |\lambda \alpha_0 + \beta_0|, \lambda \in \mathbb{R} \).

For every \( \lambda \in \mathbb{R} \), the supremum in (61) is achieved by the pair \((\alpha_0, \beta_0)\) that corresponds to \( g(\cdot) \) being of Fourier Transform

\[
\hat{g}(f) = \frac{1}{\sqrt{2\pi \mathcal{W}}} \frac{1 + \lambda \cos \left( \frac{\pi f}{2\mathcal{W}} \right)}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{2} \lambda + 1}} I\{|f| \leq \mathcal{W}\}, \quad f \in \mathbb{R}
\]  

(62)

where \( I\{|\cdot|\} \) denotes the indicator function, i.e., \( I\{|\text{statement}\} \) is 1 if the statement is true and 0 otherwise. This yields

\[
\alpha_0 = \frac{1}{\sqrt{\mathcal{W} N_0}} \frac{\frac{3}{2} + \frac{1}{2} \lambda}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{2} \lambda + 1}}, \quad \lambda \in \mathbb{R}
\]  

(63)

and

\[
\beta_0 = \frac{1}{\sqrt{\mathcal{W} N_0}} \frac{1 + \frac{3}{2} \lambda}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{2} \lambda + 1}}, \quad \lambda \in \mathbb{R}.
\]  

(64)

Note that the waveform given by (62) is of unit energy and is bandlimited to \( \mathcal{W} \) Hz, but it does not satisfy (33)–(35). Nevertheless, it is shown in Appendix B that there exist pairs \( (\alpha_0, \beta_0) \in \mathcal{B}' \) that are arbitrarily close to (63) and (64). The pairs \( (\alpha_0, \beta_0) \) given by (63) and (64) are thus in the closure of \( \mathcal{B}' \) and the boundary of \( \mathcal{B}' \) is hence parameterized by (63) and (64). (Notice that (63) and (64) describe only one half of the boundary of \( \mathcal{B}' \). The other half is given by \( (-\alpha_0, -\beta_0) \).)

The above problem is illustrated in Figure 2 for \( \mathcal{W} = N_0 = 1 \). The outer curve (solid line) is the contour line corresponding to \( \dot{R}(0) \approx 0.747 \). The inner curve (dot-dashed line) depicts the boundary of \( \mathcal{B}' \) as parameterized by (63) and (64). The two curves touch at two points: one corresponds to \( \lambda = 1.4 \) and the other is the same point reflected through the origin. Since \( \dot{R}(0) \) is convex in \( (\alpha_0, \beta_0) \), it follows that any point that lies inside the depicted contour line (solid line) yields a rate per unit-cost that is smaller than 0.747. Thus, the two points on the boundary of \( \mathcal{B}' \) that touch the contour line yield \( \dot{R}(0) \approx 0.474 \), whereas the other points \( (\alpha_0, \beta_0) \in \mathcal{B}' \) yield a rate per unit-cost that is smaller. Therefore, we conclude that choosing \( (\alpha_0, \beta_0) \) according to (63) and (64) with \( \lambda = 1.4 \) maximizes \( \dot{R}(0) \).

In the following, we summarize the above arguments to compute the supremum of \( \dot{R}(0) \) over the set \( \mathcal{B}' \) for general \( \mathcal{W} \) and \( N_0 \). To this end, we first recall that \( \dot{R}(0) \) is convex in \( (\alpha_0, \beta_0) \) and it therefore suffices to maximize \( \dot{R}(0) \) over all boundary points of \( \mathcal{B}' \). We next use that every boundary point of \( \mathcal{B}' \) is given by the parametric equations (63) and (64). Therefore, by applying (63) and (64) to (60), the maximization of \( \dot{R}(0) \) over the set \( \mathcal{B}' \) can be expressed as the maximization over the real scalar \( \lambda \)

\[
\dot{C}_{\mathcal{W}}(0) \geq \sup_{\lambda \in \mathbb{R}} \frac{2}{\pi N_0} \left[ \frac{\tilde{\alpha}(\lambda) - \frac{\bar{\alpha}(\lambda)}{2} + \frac{\bar{\alpha}(\lambda)}{4} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{\bar{\alpha}(\lambda)}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right]^{\frac{1}{2}} \right]^{\frac{1}{2}} + \frac{1}{\pi} \arcsin(\rho)
\]  

(64)
Figure 2: The contour line corresponding to $\dot{R}(0) \approx 0.747$; the boundary of $B'$ as parameterized by (63) and (64); and the point $(\alpha_0, \beta_0)$ that corresponds to $\lambda = 1.4$.

\[
\begin{align*}
+ 8 & \left( \frac{\tilde{\beta}(\lambda)}{4} - \frac{\tilde{\alpha}(\lambda)}{\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\tilde{\beta}(\lambda)}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2 \\
& + \left( \frac{\tilde{\alpha}(\lambda) - \tilde{\beta}(\lambda)}{4} - \frac{\tilde{\beta}(\lambda)}{\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{\tilde{\beta}(\lambda)}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2 \\
& \left[ \frac{1}{4} + \frac{1}{\pi} \arcsin(\rho) \right]
\end{align*}
\]  

(65)

where

\[\tilde{\alpha}(\lambda) \triangleq \frac{2 + \frac{1}{\pi} \lambda}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{4} \lambda + 1}}, \quad \lambda \in \mathbb{R}\]

and

\[\tilde{\beta}(\lambda) \triangleq \frac{1 + \frac{2}{\pi} \lambda}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{4} \lambda + 1}}, \quad \lambda \in \mathbb{R}.
\]

Numerical computation shows that the supremum on the RHS of (65) is achieved for $\lambda = 1.4$, which is consistent with our interpretation of Figure 2. The capacity per unit-cost $\dot{C}_{\frac{1}{N_0}}(0)$ is thus lower bounded

\[
\dot{C}(0) \geq \left[ \frac{g_0 + g_1 + g_2}{4} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) - \frac{g_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right]^2
\]

\[
+ 8 \left( \frac{g_0}{4} - \frac{g_1}{\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{g_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2
\]

\[
+ \left( \frac{g_1}{2} - \frac{g_0}{4} - \frac{g_1}{\pi} \arcsin \left( \frac{\rho}{\sqrt{1 - \rho^2}} \right) + \frac{g_0}{2\pi} \arcsin \left( \frac{\rho^2}{1 - \rho^2} \right) \right)^2 \\
\left[ \frac{1}{4} + \frac{1}{\pi} \arcsin(\rho) \right]
\]  

\approx 0.747 \frac{1}{N_0}
\]  

(66)
are bounded for every finite $A$ where $\Delta \eta_2$.

We show that, for $F$.

Fruitful discussions with Michèle Wigger are gratefully acknowledged.

Acknowledgment

This proves Theorem 1.

5 Summary and Conclusion

We demonstrated that doubling the sampling rate recovers some of the loss in capacity per unit-cost incurred on the bandlimited Gaussian channel with a one-bit output quantizer. Indeed, when the channel output is sampled at Nyquist rate $2W$, it is well-known that the capacity per unit-cost is given by $2^{2W} \approx 0.64 2^{2W}$, which is a factor of $2$ smaller than the capacity per unit-cost of the same channel but without output quantizer. We showed that, by sampling the output at twice the Nyquist rate, a capacity per unit-cost not less than $0.75 2^{2W}$ can be achieved. This can be viewed as a very-noisy counterpart of the work by Gilbert [6] and by Shamai [7], which demonstrated that oversampling increases the capacity of the above channel when there is no additive noise.

The conclusions that can be drawn from this result are twofold. Firstly, we demonstrated that in order to reduce the loss in capacity per unit-cost caused by the quantization, one can either increase the quantization resolution or the sampling rate. Thus, it is possible to trade amplitude resolution (quantization) versus time resolution (sampling rate). Secondly, we observe that while sampling the output at Nyquist rate is optimal for the AWGN channel (without output quantization), this does not hold when the output is quantized. Thus, a communication scheme that is optimal in the sense that it achieves the capacity need not be optimal anymore if the channel output is processed by a noninvertible operation (such as quantization).

Acknowledgment

Fruitful discussions with Michèle Wigger are gratefully acknowledged.

A Appendix to Section 4.1

A.1 Bivariate Case

We show that, for $A \geq 0$, $\alpha \in \mathbb{R}$, and $\beta \in \mathbb{R}$, we have

$$|\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta)$$

where $\Delta_1(A, \alpha, \beta)$ and $\Delta_2(A, \alpha, \beta)$ are defined in (16) and (20), and where $\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)$ and $\eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|)$ are monotonically increasing in $(A, |\alpha|, |\beta|)$ and are bounded for every finite $A$, $\alpha$, and $\beta$. We have

$$|\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta)$$

where

$$g_1 = \frac{2}{\pi} + \frac{1}{4} \lambda \left| \frac{1}{\frac{1}{2} \lambda^2 + \frac{3}{2} \lambda + 1} \right|_{\lambda=1.4}$$

and

$$g_0 = \frac{2}{\pi} + \frac{1}{4} \lambda \left| \frac{1}{\frac{1}{2} \lambda^2 + \frac{3}{2} \lambda + 1} \right|_{\lambda=1.4}$$

This proves Theorem 1.

\[ |\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]

where $\Delta_1(A, \alpha, \beta)$ and $\Delta_2(A, \alpha, \beta)$ are defined in (16) and (20), and where $\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)$ and $\eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|)$ are monotonically increasing in $(A, |\alpha|, |\beta|)$ and are bounded for every finite $A$, $\alpha$, and $\beta$. We have

\[ |\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]

where $\Delta_1(A, \alpha, \beta)$ and $\Delta_2(A, \alpha, \beta)$ are defined in (16) and (20), and where $\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)$ and $\eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|)$ are monotonically increasing in $(A, |\alpha|, |\beta|)$ and are bounded for every finite $A$, $\alpha$, and $\beta$. We have

\[ |\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]

where $\Delta_1(A, \alpha, \beta)$ and $\Delta_2(A, \alpha, \beta)$ are defined in (16) and (20), and where $\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)$ and $\eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|)$ are monotonically increasing in $(A, |\alpha|, |\beta|)$ and are bounded for every finite $A$, $\alpha$, and $\beta$. We have

\[ |\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]

where $\Delta_1(A, \alpha, \beta)$ and $\Delta_2(A, \alpha, \beta)$ are defined in (16) and (20), and where $\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)$ and $\eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|)$ are monotonically increasing in $(A, |\alpha|, |\beta|)$ and are bounded for every finite $A$, $\alpha$, and $\beta$. We have

\[ |\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]

where $\Delta_1(A, \alpha, \beta)$ and $\Delta_2(A, \alpha, \beta)$ are defined in (16) and (20), and where $\eta_1(A, \alpha, \beta) = \eta_1(A, |\alpha|, |\beta|)$ and $\eta_2(A, \alpha, \beta) = \eta_2(A, |\alpha|, |\beta|)$ are monotonically increasing in $(A, |\alpha|, |\beta|)$ and are bounded for every finite $A$, $\alpha$, and $\beta$. We have

\[ |\Delta_1(A, \alpha, \beta)| \leq A^2 \eta_1(A, \alpha, \beta) \quad \text{and} \quad |\Delta_2(A, \alpha, \beta)| \leq A^2 \eta_2(A, \alpha, \beta) \]
where

\[ \eta_1(A, \alpha, \beta) \triangleq \frac{|\alpha||\beta|}{\sqrt{2\pi(1-\varrho^2)}} + \frac{|\varrho|^2}{\sqrt{2\pi(1-\varrho^2)}} + \frac{|\alpha|(|\beta| + |\varrho|)|\alpha|^3}{6\sqrt{2\pi}(1-\varrho^2)^{3/2}} A^2 + \frac{|\alpha|^3}{12\sqrt{2\pi}} A. \]

Here the first step follows from the Triangle Inequality; the second step follows by upper bounding \( \phi_{0,1}(x) \leq 1, x \in \mathbb{R} \) and from (17); and the third step follows because, over the range of integration, \( x \leq |\alpha| A \) and \( |\beta A + \varrho x| \leq (|\beta| + |\varrho| |\alpha|) A \).

Along the same lines, we obtain

\[
|\Delta_2(A, \alpha, \beta)| \\
\leq \frac{|\varrho|}{\sqrt{2\pi(1-\varrho^2)}} \int_{-\beta A}^0 \phi_{0,1}(y)|y|dy + \int_{-\beta A}^0 \phi_{0,1}(y) \bigg| \delta \left( \frac{\varrho y}{\sqrt{1-\varrho^2}} \right) \bigg| dy + \frac{1}{2} |\delta(\beta A)| \\
\leq \frac{|\varrho|}{\sqrt{2\pi(1-\varrho^2)}} \int_0^{|\beta|A} y dy + \int_0^{|\beta|A} \frac{(|\varrho|y)^3}{6\sqrt{2\pi}(1-\varrho^2)^{3/2}} dy + \frac{|\beta|^3}{12\sqrt{2\pi}} A^3
\]

where

\[
\eta_2(A, \alpha, \beta) \triangleq \frac{|\varrho|^2}{\sqrt{2\pi(1-\varrho^2)}} + \frac{|\varrho|^3\beta^4}{6\sqrt{2\pi}(1-\varrho^2)^{3/2}} A^2 + \frac{|\beta|^3}{12\sqrt{2\pi}} A.
\]

This proves the claim.

### A.2 Trivariate Case

We show that, for \( A > 0, \alpha \in \mathbb{R}, \beta \in \mathbb{R}, \) and \( \gamma \in \mathbb{R} \), we have

\[
|\Delta_1(A, \alpha, \beta, \gamma)| \leq A^2 \eta_1(A, \alpha, \beta, \gamma) \\
|\Delta_2(A, \alpha, \beta, \gamma)| \leq A^2 \eta_2(A, \alpha, \beta, \gamma)
\]

and

\[
|\Delta_3(A, \alpha, \beta, \gamma)| \leq A^2 \eta_3(A, \alpha, \beta, \gamma)
\]

where \( \Delta_1(A, \alpha, \beta, \gamma), \Delta_2(A, \alpha, \beta, \gamma), \) and \( \Delta_3(A, \alpha, \beta, \gamma) \) are defined in (25), (28), and (30), and where \( \eta_1(A, \alpha, \beta, \gamma) = \eta_1(A, |\alpha|, |\beta|, |\gamma|), \eta_2(A, \alpha, \beta, \gamma) = \eta_2(A, |\alpha|, |\beta|, |\gamma|), \) and \( \eta_3(A, \alpha, \beta, \gamma) = \eta_3(A, |\alpha|, |\beta|, |\gamma|) \) are monotonically increasing in \((A, |\alpha|, |\beta|, |\gamma|)\) and are bounded for every finite \( A, \alpha, \beta, \) and \( \gamma \). We have

\[
|\Delta_1(A, \alpha, \beta, \gamma)| \leq |\delta(\alpha A)| \bigg| \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{12} - \theta_1 \theta_{13}}{\sqrt{(1-\theta_{12})(1-\theta_{13})}} \right) \bigg| \\
+ \left| \int_{-\alpha A}^0 \phi_{0,1}(x) \frac{1}{2\sqrt{2\pi}} \frac{\beta A + \varrho_1 x}{\sqrt{1-\theta_{12}}} + \frac{\gamma A + \varrho_1 x}{\sqrt{1-\theta_{13}}} \right| dx \bigg| \\
+ \left| \int_{-\alpha A}^0 \phi_{0,1}(x) \Delta \left( A, \frac{\beta + \varrho_1 x}{\sqrt{1-\theta_{12}}} \right) + \frac{\gamma A + \varrho_1 x}{\sqrt{1-\theta_{13}}} \right| dx \bigg| \\
\leq \frac{1}{2} |\delta(\alpha A)| + \frac{1}{2\sqrt{2\pi}} \int_{-\alpha A}^0 \left| \frac{1}{2\sqrt{2\pi}} \frac{\beta A + \varrho_1 x}{\sqrt{1-\theta_{12}}} + \frac{\gamma A + \varrho_1 x}{\sqrt{1-\theta_{13}}} \right| dx \bigg| \\
+ \left| \Delta \left( A, \frac{\beta + \varrho_1 x}{\sqrt{1-\theta_{12}}} \right) \frac{\gamma A + \varrho_1 x}{\sqrt{1-\theta_{13}}} \right| dx \bigg| \\
\leq \frac{|\alpha|^3}{12\sqrt{2\pi}} A^3 + \frac{1}{2\sqrt{2\pi}} \int_0^{|\alpha|A} \left( \frac{|\beta|A + |\varrho_1| x}{\sqrt{1-\theta_{12}}} + \frac{|\gamma|A + |\varrho_1| x}{\sqrt{1-\theta_{13}}} \right) dx \bigg| \\
+ A^2 \int_0^{|\alpha|A} \eta \left( A, \frac{|\beta|A + |\varrho_1| x}{\sqrt{1-\theta_{12}}} + \frac{|\gamma|A + |\varrho_1| x}{\sqrt{1-\theta_{13}}} \right) dx.
\]
\[
\begin{align*}
\Delta_0 & \leq \frac{|\alpha|^3}{12\sqrt{2\pi}} A^3 + \frac{|\alpha|}{2\sqrt{2\pi}} \left( \frac{|\beta| + |\theta_{12}| |\alpha|}{\sqrt{1-\theta_{12}^2}} + \frac{|\gamma| + |\theta_{13}| |\alpha|}{\sqrt{1-\theta_{13}^2}} \right) A^2 \\
& + |\alpha| \eta \left( A, \frac{|\beta| + |\theta_{12}| |\alpha|}{\sqrt{1-\theta_{12}^2}}, \frac{|\gamma| + |\theta_{13}| |\alpha|}{\sqrt{1-\theta_{13}^2}} \right) A^3 \\
& = A^2 \eta_1(A, \alpha, \beta, \gamma) \quad (69)
\end{align*}
\]

where \( \eta(A, \alpha, \beta) \) is as in Proposition 2, and where
\[
\eta_1(A, \alpha, \beta, \gamma) \triangleq \frac{|\alpha|^3}{12\sqrt{2\pi}} A + \frac{|\alpha|}{2\sqrt{2\pi}} \left( \frac{|\beta| + |\theta_{12}| |\alpha|}{\sqrt{1-\theta_{12}^2}} + \frac{|\gamma| + |\theta_{13}| |\alpha|}{\sqrt{1-\theta_{13}^2}} \right) A^2 \\
+ |\alpha| \eta \left( A, \frac{|\beta| + |\theta_{12}| |\alpha|}{\sqrt{1-\theta_{12}^2}}, \frac{|\gamma| + |\theta_{13}| |\alpha|}{\sqrt{1-\theta_{13}^2}} \right) A.
\]

Here the first step follows from the Triangle Inequality; the second step follows by upper bounding \( \phi_{0,1}(x) \leq 1, \ x \in \mathbb{R} \) and \( \arcsin(x) \leq \pi/2, |x| \leq 1; \) the third step follows again from the Triangle Inequality, from the upper bound
\[
|\Delta(A, \alpha, \beta)| \leq A^2 \eta(A, \alpha, \beta)
\]
and from the monotonicity of \( \eta(A, \alpha, \beta) \) in \( (A, |\alpha|, |\beta|) \); and the fourth step follows because, over the range of integration, \( x \leq |\alpha| A \).

Along the same lines, we obtain for \( \Delta_2(A, \alpha, \beta, \gamma) \)
\[
|\Delta_2(A, \alpha, \beta, \gamma)| \leq |\delta(\beta A)| \left| \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{13} - \theta_{12}\theta_{23}}{(1-\theta_{12}^2)(1-\theta_{23}^2)} \right) \right| \\
+ \int_{-\beta A}^0 \phi_{0,1}(y) \left| \frac{1}{2\sqrt{2\pi}} \frac{\theta_{12} y}{\sqrt{1-\theta_{12}^2}} + \frac{\gamma A + \theta_{23} y}{\sqrt{1-\theta_{23}^2}} \right| dy \\
+ \int_{-\beta A}^0 \phi_{0,1}(y) \left| \Delta \left( A, \frac{\theta_{12} y / A}{\sqrt{1-\theta_{12}^2}} + \frac{\gamma + \theta_{23} y / A}{\sqrt{1-\theta_{23}^2}} \right) \right| dy \\
\leq \frac{|\beta|^3}{12\sqrt{2\pi}} A^3 + \frac{|\beta|}{2\sqrt{2\pi}} \left( \frac{|\theta_{12}| |\beta|}{\sqrt{1-\theta_{12}^2}} + \frac{|\gamma| + |\theta_{23}| |\beta|}{\sqrt{1-\theta_{23}^2}} \right) A^2 \\
+ |\beta| \eta \left( A, \frac{|\theta_{12}| |\beta|}{\sqrt{1-\theta_{12}^2}}, \frac{|\gamma| + |\theta_{23}| |\beta|}{\sqrt{1-\theta_{23}^2}} \right) A^3 \\
= A^2 \eta_2(A, \alpha, \beta, \gamma) \quad (70)
\]

where
\[
\eta_2(A, \alpha, \beta, \gamma) \triangleq \frac{|\beta|^3}{12\sqrt{2\pi}} A + \frac{|\beta|}{2\sqrt{2\pi}} \left( \frac{|\theta_{12}| |\beta|}{\sqrt{1-\theta_{12}^2}} + \frac{|\gamma| + |\theta_{23}| |\beta|}{\sqrt{1-\theta_{23}^2}} \right) A \\
+ |\beta| \eta \left( A, \frac{|\theta_{12}| |\beta|}{\sqrt{1-\theta_{12}^2}}, \frac{|\gamma| + |\theta_{23}| |\beta|}{\sqrt{1-\theta_{23}^2}} \right) A.
\]

Finally, we obtain for \( \Delta_3(A, \alpha, \beta, \gamma) \)
\[
|\Delta_3(A, \alpha, \beta, \gamma)| \leq |\delta(\gamma A)| \left| \frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \frac{\theta_{12} - \theta_{13}\theta_{23}}{(1-\theta_{13}^2)(1-\theta_{23}^2)} \right) \right| \\
+ \int_{-\gamma A}^0 \phi_{0,1}(z) \left| \frac{1}{2\sqrt{2\pi}} \frac{\theta_{13} z}{\sqrt{1-\theta_{13}^2}} + \frac{\theta_{23} z}{\sqrt{1-\theta_{23}^2}} \right| dz \\
+ \int_{-\gamma A}^0 \phi_{0,1}(z) \Delta \left( A, \frac{\theta_{13} z / A}{\sqrt{1-\theta_{13}^2}} + \frac{\theta_{23} z / A}{\sqrt{1-\theta_{23}^2}} \right) dz
\]

23
that give rise to pairs \((\alpha, \beta)\) where

\[
W \leq A_2 \eta_3(A, \alpha, \beta, \gamma)
\]

This proves the claim.

### B Appendix to Section 4.2

We show that there exist bandlimited, unit-energy waveforms \(g(\cdot)\) that satisfy (33)–(35) and that give rise to pairs \((\alpha_0, \beta_0)\) that are arbitrarily close to

\[
\frac{1}{\sqrt{2N}} \frac{\frac{2}{T} + \frac{4}{\sqrt{\lambda^2 + 4\lambda + 1}}}{\sqrt{\frac{1}{2}} \lambda^2 + \frac{4}{\sqrt{\lambda^2 + 4\lambda + 1}}} = \lambda \in \mathbb{R}
\]

and

\[
\frac{1}{\sqrt{2N}} \frac{1 + \frac{2}{T}}{\sqrt{\frac{1}{2}} \lambda^2 + \frac{4}{\sqrt{\lambda^2 + 4\lambda + 1}}} = \lambda \in \mathbb{R}
\]

To this end, we first note that a sufficient condition for the waveform \(g(\cdot)\) to satisfy (33)–(35) is

\[
|g(t)| \leq K_1, \quad t \in \mathbb{R} \quad \text{and} \quad |g(t)| \leq \frac{K_2}{t^2}, \quad |t| > T
\]

for some nonnegative \(K_1\), \(K_2\), and some positive \(T\). Indeed, if (74) holds, then we have for every \(\tau \in \mathbb{R}\)

\[
\sum_{\ell \neq 0} \left| g \left( \frac{\ell + \tau}{2W} \right) \right| \leq \sum_{\ell = -\infty}^{\infty} \left| g \left( \frac{\ell + \tau}{2W} \right) \right|
\]

\[
= \sum_{|\ell + \tau| \leq T + \frac{1}{2W}} \left| g \left( \frac{\ell + \tau}{2W} \right) \right| + \sum_{|\ell + \tau| > T + \frac{1}{2W}} \left| g \left( \frac{\ell + \tau}{2W} \right) \right|
\]

\[
\leq 2 K_1 (2WT + 1) + \sum_{|\ell + \tau| > T + \frac{1}{2W}} \frac{K_2(2W)^2}{(\ell + \tau)^2}
\]

\[
\leq 2 K_1 (2WT + 1) + 2 \int_{2WT + 1}^{\infty} \frac{K_2(2W)^2}{(t - 1)^2} dt
\]

\[
= 2 K_1 (2WT + 1) + 2 K_2 \frac{2W}{T}
\]

\[
< \infty
\]

where the first step follows because \(|g(\tau/(2W))|\) is nonnegative; the third step follows from (74) and because there are not more than \(2WT + 1\) terms satisfying \(|(\ell + \tau)/(2W)| \leq T + 1/(2W)\); and the fourth step follows by upper bounding the sum by an integral.
The waveform $g(\cdot)$ corresponding to the Fourier Transform (62) is given by

$$g(t) = \sqrt{\frac{2W}{\lambda^2 + \frac{1}{2} \lambda + 1}} \left( \sin(2Wt) + \frac{\lambda}{2} \sin \left(2Wt - \frac{1}{2}\right) + \frac{\lambda}{2} \sin \left(2Wt + \frac{1}{2}\right) \right), \quad t \in \mathbb{R}. \quad (76)$$

We note that the sum of the last two terms on the RHS of (76) equals

$$\frac{\lambda}{2} \sin \left(2Wt - \frac{1}{2}\right) + \frac{\lambda}{2} \sin \left(2Wt + \frac{1}{2}\right)$$

which satisfies (74) and hence also (33)–(35). However, $t \mapsto \sin(2Wt)$ decays like $1/t$ and does neither satisfy (33) nor (35). We therefore replace $t \mapsto \sin(2Wt)$ in (76) by a raised-cosine pulse of bandwidth $W$ Hz and of roll-off factor $\xi$ to obtain the desired waveform

$$g_\xi(t) = \sqrt{\frac{2W}{\Psi(\xi)}} \left( \frac{1}{1 + \xi} \sin \left(\frac{2Wt}{1 + \xi}\right) \cos \left(\frac{\pi 2W \xi t}{\lambda_{\xi} t}\right) \right.$$

$$\left. + \frac{\lambda}{2} \sin \left(2Wt - \frac{1}{2}\right) + \frac{\lambda}{2} \sin \left(2Wt + \frac{1}{2}\right) \right), \quad t \in \mathbb{R} \quad (77)$$

where

$$\Psi(\xi) \triangleq 2W \int \left( \frac{1}{1 + \xi} \sin \left(\frac{2Wt}{1 + \xi}\right) \cos \left(\frac{\pi 2W \xi t}{\lambda_{\xi} t}\right) \right.$$

$$\left. + \frac{\lambda}{2} \sin \left(2Wt - \frac{1}{2}\right) + \frac{\lambda}{2} \sin \left(2Wt + \frac{1}{2}\right) \right)^2 \, dt.$$

For every $\xi > 0$, the raised-cosine pulse on the RHS of (77) satisfies (74) and hence also (33)–(35). It therefore follows from the Triangle Inequality that also the pulse $g_\xi(\cdot)$ satisfies (33)–(35), since for every $t \in \mathbb{R}$

$$|g_\xi(t)| \leq \sqrt{\frac{2W}{\Psi(\xi)}} \left| \frac{1}{1 + \xi} \sin \left(\frac{2Wt}{1 + \xi}\right) \cos \left(\frac{\pi 2W \xi t}{\lambda_{\xi} t}\right) \right.$$

$$\left. + \frac{\lambda}{2} \sin \left(2Wt - \frac{1}{2}\right) + \frac{\lambda}{2} \sin \left(2Wt + \frac{1}{2}\right) \right|.$$

Furthermore, since $g_\xi(\cdot)$ is of unit energy and bandlimited to $W$ Hz, it follows that $g_\xi(\cdot)$ is in $\mathcal{G}'$, and hence the pairs $(\alpha_0, \beta_0)$ that arise from $g_\xi(\cdot)$ are in $\mathcal{B}'$.

It remains to show that such $(\alpha_0, \beta_0)$ approach (72) and (73) as $\xi$ tends to zero. To this end, we show that for every $t \in \mathbb{R}$

$$\lim_{\xi \downarrow 0} g_\xi(t) = g(t). \quad (78)$$

This implies

$$\lim_{\xi \downarrow 0} \alpha_0 = \lim_{\xi \downarrow 0} \frac{1}{\sqrt{(2W)(\sqrt{\lambda_0} g_0)}} g_\xi \left( -\frac{1}{4W} \right) = \frac{1}{\sqrt{\lambda_0}} \frac{\frac{1}{2} + \frac{\lambda}{4}}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{4} \lambda + 1}}$$

and

$$\lim_{\xi \downarrow 0} \beta_0 = \lim_{\xi \downarrow 0} \frac{1}{\sqrt{(2W)(\sqrt{\lambda_0} g_0)}} g_\xi (0) = \frac{1}{\sqrt{\lambda_0}} \frac{1 + \frac{\lambda}{2}}{\sqrt{\frac{1}{2} \lambda^2 + \frac{1}{4} \lambda + 1}}.$$
which in turn proves the claim.

To prove (78), we note that $t \mapsto \sqrt{\Psi(\xi)}g_\xi(t)$—which we shall denote by $t \mapsto h_\xi(t)$—satisfies

$$
\lim_{\xi \downarrow 0} h_\xi(t) = \sqrt{\left(\frac{1}{2} \lambda^2 + \frac{4}{\pi} \lambda + 1\right)} g(t), \quad t \in \mathbb{R}
$$

(79)

where $g(\cdot)$ is the Inverse Fourier Transform of (62). Furthermore, since the Fourier Transform of $h_\xi(\cdot)$ satisfies $|\hat{h}_\xi(f)|^2 \leq (1/2 \lambda^2 + 4/\pi \lambda + 1) |\hat{g}(f)|^2$, $f \in \mathbb{R}$ (with $\hat{g}(\cdot)$ given by (62)), and since $\int (1/2 \lambda^2 + 4/\pi \lambda + 1) |\hat{g}(f)|^2 df < \infty$, it follows from the Dominated Convergence Theorem [16, Thm. 1.34] that

$$
\lim_{\xi \downarrow 0} \Psi(\xi) = \lim_{\xi \downarrow 0} \int (h_\xi(t))^2 dt = \lim_{\xi \downarrow 0} \int |\hat{h}_\xi(f)|^2 df = \int \lim_{\xi \downarrow 0} |\hat{h}_\xi(f)|^2 df = \frac{1}{2} \lambda^2 + \frac{4}{\pi} \lambda + 1
$$

(80)

where the second step follows from Parseval’s Theorem [8, Thm. 6.2.9], and where the last step follows from (79) and because $g(\cdot)$ is of unit energy. It thus follows that for every $t \in \mathbb{R}$

$$
\lim_{\xi \downarrow 0} g_\xi(t) = \lim_{\xi \downarrow 0} \frac{h_\xi(t)}{\sqrt{\Psi(\xi)}} = \frac{\lim_{\xi \downarrow 0} h_\xi(t)}{\sqrt{\lim_{\xi \downarrow 0} \Psi(\xi)}} = g(t)
$$

(81)

where the last step follows from (79) and (80). This proves (78) and hence also the claim.

References

[1] C. E. Shannon, “A mathematical theory of communication,” Bell System Technical Journal, vol. 27, pp. 379–423 and 623–656, July and October 1948.

[2] R. G. Gallager, Information Theory and Reliable Communication. John Wiley & Sons, 1968.

[3] A. J. Viterbi and J. K. Omura, Principles of Digital Communication and Coding. McGraw-Hill, 1979.

[4] R. J. McEliece, The Theory of Information and Coding, 2nd ed. Cambridge University Press, 2002.

[5] J. Singh, O. Dabeer, and U. Madhow, “Transceiver design with low-precision analog-to-digital conversion: An information theoretic perspective,” 2009, to appear in IEEE Transactions on Communications. [Online]. Available: http://arxiv.org/abs/0804.1172v1

[6] E. N. Gilbert, “Increased information rate by oversampling,” IEEE Transactions on Information Theory, vol. 39, pp. 1973–1976, November 1993.

[7] S. Shamai (Shitz), “Information rates by oversampling the sign of a bandlimited process,” IEEE Transactions on Information Theory, vol. 40, pp. 1230–1236, July 1994.

[8] A. Lapidoth, A Foundation in Digital Communication. Cambridge University Press, 2009.

[9] T. M. Cover and J. A. Thomas, Elements of Information Theory, 1st ed. John Wiley & Sons, 1991.

[10] W. F. Sheppard, “On the calculation of the double integral expressing normal correlation,” Transaction of the Cambridge Philosophical Society, vol. 19, pp. 23–66, 1900.

[11] S. S. Gupta, “Probability integrals of multivariate normal and multivariate t,” The Annals of Mathematical Statistics, vol. 34, pp. 792–828, 1963.

[12] S. Verdú, Multiuser Detection. Cambridge University Press, 1998.

[13] I. S. Gradsteyn and I. M. Ryzhik, Table of Integrals, Series, and Products, 6th ed., A. Jeffrey, Ed. Academic Press, San Diego, 2000.
[14] F. N. David, “A note on the evaluation of the multivariate normal integral,” *Biometrika*, vol. 40, pp. 458–459, 1953.

[15] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[16] W. Rudin, *Real and Complex Analysis*, 3rd ed. McGraw-Hill, 1987.