

Article

Geometry-Based Bounds on the Capacity of Peak-Limited and Band-Limited Signals over the Additive White Gaussian Noise Channel at a High SNR

Michael Peleg *  and Shlomo Shamai 

Faculty of Electrical and Computer Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel

* Correspondence: peleg.michael@gmail.com

Abstract

We present a new computable geometry-based upper bound on the capacity of peak-power-limited and band-limited signal over the Additive White Gaussian Noise Channel. The peak limit applies at continuous time. The bound is a function of the volume and shape of the transmitted signal set, namely the body, in the space of Nyquist-rate samples, comprising all of the points the transmitted signal can reach. At a high SNR, the bound is tight, better than previously known upper bounds and, together with a known lower bound, provides the capacity at an asymptotically high SNR. We found, using a numerical evaluation, the high-SNR capacity of signals with the structure used in Cyclic Prefix assisted Frequency Domain Equalization (CP-FDE) and OFDM for sequence length of up to 100 Nyquist intervals, and we present a conjecture that this result is correct for any sequence length and does not depend on the CPA-FDE structure. This paper extends the methodology developed in previous works. The penalty in power efficiency at a high SNR due to the peak power constraint relative to an average power constraint is about 7.5 dB in the low-pass case and about 5.4 dB in the band-pass case.

Keywords: peak power; amplitude; capacity; AWGN; band limited; entropy; upper bound



Academic Editor: Lei Liu

Received: 15 October 2025

Revised: 19 November 2025

Accepted: 19 November 2025

Published: 24 November 2025

Citation: Peleg, M.; Shamai, S. Geometry-Based Bounds on the Capacity of Peak-Limited and Band-Limited Signals over the Additive White Gaussian Noise Channel at a High SNR. *Entropy* **2025**, *27*, 1192. <https://doi.org/10.3390/e27121192>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

We investigate the Peak-Power Limited and Band Limited (PPBL) Additive White Gaussian Noise (AWGN) channel in which the signal is band-limited and its instantaneous power cannot exceed the power P . Equivalently, the instantaneous amplitude cannot exceed $\sqrt{P}$. The peak power (peak amplitude) limit applies continuously, not only at the Nyquist-rate sampling times. This model is relevant to many systems in which the peak power is limited by the power amplifier at the transmitter. The model fits real system requirements even better since the introduction of Digital Pre-Distortion (DPD), e.g., [1,2], which linearizes the power amplifier up to its maximal transmit power, thus ensuring high and predictable performance with peak-power-limited signals. Clearly, the capacity limits of this channel are of major practical interest, e.g., the discussion of advances in new physical layer technologies in [3], the optimization in Section 3.6 of [2], and discussion in [4]. The importance of limiting the peak power is also reflected in many works analyzing and reducing the Peak to Average Power Ratio (PAPR), e.g., [5–8]. The impact of the peak power limit is classical in many communications settings since the beginning of the wireless communication era [9], and it is relevant to a variety of practical communication models,

as, for example, fading channels and the like [4], and the optimization of constellations under the peak and average limits over the discrete-time channel [10]. The new results on the PPBL channel capacity may benefit recent and modern research on coding for this channel, e.g., [11], which investigates PAPR reduction for single carrier communication by transmit filter impulse-response optimization and [12], which induces a peak limit on 16-QAM Single-Carrier modulation by trellis shaping and exploits the shaping trellis as a component code of a turbo-coded scheme.

With the exact capacity of the classical Average-Power-Limited and Band-Limited (APBL) channel found by Shannon [9] and used widely for decades, the continuous-time PPBL channel capacity has only been studied sparsely, yielding lower and upper bounds on capacity with a wide gap in between. We believe the reason for this is the difficulty of analyzing this channel, as suggested already in [9]. Shannon analyzed this channel and presented lower and upper bounds on capacity in the low-pass case already in [9]. His lower bound was based on an achievability scheme with a low-pass signaling impulse response; his upper bound, valid at an asymptotically high SNR, releases the peak limit to be applicable to Nyquist-rate samples only and then upper-bounds the entropy of the transmitted signals. The work by [13] used an improved impulse response to slightly improve the lower bound of [9] and provided an improved upper bound valid at all SNRs by releasing the peak limit to be applicable to Nyquist-rate samples only while applying the capacities of the corresponding sampled discrete-time channels available in [14,15]. It also extended the bounds to the more practical band-pass case. The capacity of the discrete-time peak-limited channel and analytical bounds on it are further investigated in [16,17]. Smith [14] derived the exact capacity-achieving distribution of signals over the discrete-time peak-limited channel and proved that it comprises discrete amplitudes. The capacity of this channel upper-bounds that of our continuously peak-limited channel, and both channels are peak-limited; still, the question whether the optimal signal in our case is similar to that in [14] remains unanswered. Recently, we presented a computable lower bound [18] on the capacity of the time-continuous PPBL channel at any Signal to Noise Ratio (SNR) applicable to the low-pass and the band-pass cases. The bound there was evaluated numerically over the Cyclic Prefix-assisted Frequency Domain Equalization (CP-FDE) or Orthogonal Frequency Division Multiplexing (OFDM) signaling. The CP-FDE signals are not strictly band-limited because they are limited in time; however, they are practically band-limited in the sense of zero inter-channel interference between users if the rules for cyclic prefix are adhered to, thus enabling the assignment of adjacent users to channels with no frequency gaps in between. This is applied, for example, in the multiuser uplink of the Long-Term Evolution (LTE) mobile communications system using Single-Carrier FDMA (SC-FDMA) [19].

The problem investigated here is related to communication over the Constrained Gaussian Channel (CGC) [20] in which a wideband peak-limited signal is fed into a transmit filter in the transmitter. As shown in [18], the capacity of the CGC is an upper bound of the capacity of the PPBL channel. The work by [20] uses the results of [21] on the Power Spectral Density (PSD) of unit processes, which are the inputs to the CGC channel, to derive the upper-bound capacity of the CGC channel, which is also valid for the PPBL channel. In [22,23], the approach is specified to the PPBL channel, gaining additional insights. The upper bound on the CGC channel was tightened in [23]. However, the CGC-based approach cannot provide an upper bound on capacity tight at high SNRs, as explained in Section 3.3.2 below. The review by [24] presents and categorizes a wide range of modulation schemes with different types of peak limits, including the CGC and the PPBL models.

Contributions: We derive a computable upper bound on the capacity of peak-limited CP-FDE and OFDM-type signals over the AWGN band-limited channel based on the volume and the shape of the signal set in the space of Nyquist-rate samples, comprising all of the points the transmitted signal can reach. This technique extends the geometric analysis of the transmitted signal set in [18] to analyze the geometric region into which the additive noise extends the transmit signal set to become the received signal set. The new upper bound is tighter at high SNRs than the known upper bounds. The existence of such a bound was mentioned in [18]; here, the bound is proved and its gap to the known lower bound is evaluated. The new upper bound approaches the lower bound at an asymptotically high SNR, thus providing capacity at an asymptotically high SNR. We also verify the lower bound in [18] by evaluating it using a more straightforward numerical method which enables us to control the standard deviation of the results and to present confidence intervals, unlike the more powerful but more complex method used in [18].

The methodology used in the current paper evolved relatively to that in the previous paper [18] as follows: In [18], we developed a lower bound on the capacity. The bound is a function of the volume of the transmitted signal set and uses the relation between maximal entropy and the volume of the signal set [9]. Since the signal set is N -dimensional, there was a need to develop a numerical technique to estimate the volume in the N -dimensional space, including a complex investigation of its accuracy. In the current paper, we developed an upper bound on the capacity to complement the lower bound [18]. To do so, we had to evaluate the entropy of the noisy received signal. Unlike the transmit signal [18], the received noisy signal is not confined to a bounded volume, and its entropy cannot be bounded by a function of the volume of the transmit signal set alone. Therefore, a new method bounding the contribution to the entropy of sections of space, each section extending over an unbounded volume, occupied by the received signal set, was developed. We utilized the volume-estimation core from [18] as an element of the new method.

Notation: Bold italic *letters* denote vectors. Log is the natural logarithm, unless stated otherwise. Differential entropy is denoted by h ; E denotes the statistical expectation. The N -dimensional vector space of real variables is denoted R^N . The Probability Density Function (PDF) of x is $p_x(x)$ or $p(x)$. Convex bodies are denoted by calligraphic capital letters such as $\mathcal{X}$. The unity-radius N -ball centered at 0 is $\mathcal{B}$. The addition of convex bodies is performed by the Minkowski addition and is denoted by the usual $+$.

2. System Model

We begin with real-valued signals representing a low-pass channel. The system is presented in Figure 1. The encoder produces a real-valued low-pass signal $x(t)$ in the frequency band $|f| < B$. The signal is constrained to be peak-limited, that is, $|x(t)| \leq \sqrt{P}$. The signal passes an AWGN channel and is decoded. The channel output $y(t)$ is

$$y(t) = x(t) + n(t) \quad (1)$$

where $n(t)$ is a white Gaussian noise with power spectral density N_0 ($0.5N_0$ two-sided) and in-band power $\sigma_n^2 = N_0B$. The Nyquist interval is $T = 0.5/B$. The channel and/or the decoder may include a brick-wall low-pass filter with bandwidth B ; this does not influence the capacity, since $x(t)$ is already band-limited. We seek the bounds on the capacity, which is the maximal Mutual Information (MUI) denoted $I(x;y)$ per Nyquist interval. The Nyquist-rate-sampled $x(t)$ is denoted by the vector $\mathbf{x} = (x_1 \dots x_k \dots x_N)$ of length N , $x_k = x(kT)$. The vectors $\mathbf{y}$ and $\mathbf{n}$ denote the sampled received signal and noise, respectively. The sets of signals $\mathbf{x}$ and $\mathbf{y}$ are denoted $\mathcal{X}$ and $\mathcal{Y}$, respectively.

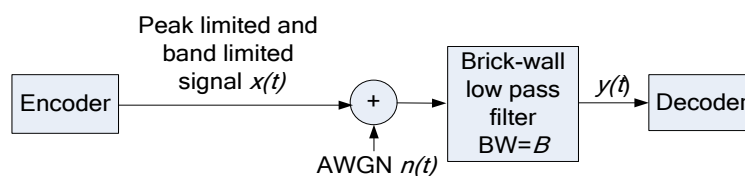


Figure 1. System model of communication over the PPBL channel, low-pass case.

The signal to noise ratio (SNR) is defined as $\rho = \frac{P}{BN_0}$, as in the classical APBL channel. The capacity in bits per Nyquist interval of an APBL channel is the following famous [9] equation:

$$C_a = 0.5 \log_2 \left(\frac{P}{N_0 B} + 1 \right) \quad (2)$$

To render the numerical analysis feasible, N needed to be finite. To this end, we modelled $x(t)$ as cyclic with a period of NT and then limited to a duration of NT . This is identical to the CP-FDE and OFDM signaling formats before adding the cyclic prefix, so our analysis is applicable to CP-FDE and OFDM. In this setting, $x(t)$ can be expressed in terms of its Nyquist-rate samples, e.g., [25], as

$$x(t) = \sum_{i=0}^{N-1} x_i \cdot \varphi(t - iT); \quad 0 \leq t < NT \quad (3)$$

where $\varphi(t)$ is the periodic sinc function (4), also known as the Dirichlet kernel, instead of the usual sinc function (see, e.g., [25]):

$$\varphi(t) = \frac{\sin(\pi t/T)}{N \cdot \sin[\pi t/(N \cdot T)]} \quad (4)$$

The expression is valid for odd N and has a counterpart for even N (see, e.g., [25]). Note that $\varphi(t)$ is periodic, with a period of NT . The functions $\varphi(t - iT)$, $i = 0 \dots N - 1$ over $0 < t < NT$ form an orthogonal basis of our band-limited periodic signals, as can be shown easily using Equation (11) in [25].

The definition of capacity, which requires long sequences, is applicable if each error-correction block comprises many N -samples long CP-FDE signal segments (3), with each segment different and generated jointly by the error-correcting encoder. In the case of CP-FDE and OFDM signaling, each signal segment (3) of duration NT is to be prepended by its cyclic prefix and concatenated with the other segments to form the error-correction block. Thus, our model directly applies to CP-FDE and OFDM.

If N is increased, the cyclic prefixes become negligible and the signal with large and finite N is similar to a standard low-pass signal by the similarity of the sinc and periodic sinc (4) impulse responses, as follows: The standard low-pass signal $x(t)$ is determined at time t by its Nyquist-rate samples x_i weighted by the sinc impulse response. We shall partition those samples into two groups: The first group comprises the N -neighboring samples x_i centered around t , such that $|iT - t| < NT/2$. Each such sample has one parallel sample in the finite N case (3), weighted by the periodic sinc impulse response (4). Comparison of the two impulse responses shows that they are similar at the relevant times; the differences are smaller than approximately $\frac{\pi-2}{\pi N}$ relative to the peak of the impulse response, which is 0.0071 for $N = 51$ and smaller for larger N . The second group comprises all the other more distant samples. These have no parallel in the finite N case; however, the sinc impulse response relevant to those samples is low, with a maxima of about $2 / (\pi N)$ relative to the peak of the impulse response; thus, their influence decreases with growing N .

3. Bounds on Capacity

3.1. On the Shape of the Signal Set

The transmit signal set $\mathcal{X}$ of vectors x of Nyquist-rate samples has the following attributes, illustrated in Figure 2: The set $\mathcal{X}$ is convex, that is, if signals x_1 and x_2 are in the signal set, so is $ax_1 + (1-a)x_2$ for any real positive scalar $a < 1$ (see [18] for proof and discussion). The signal region is symmetric, that is, if x is permissible, then $-x$ is too. The set $\mathcal{X}$ is time-invariant, that is, invariant with respect to a time-shift in an integer number of Nyquist-rate samples; the shift becomes circular in the case of the CP-FDE channel model. The set $\mathcal{X}$ is confined to be in an N -cube, defined by $|x_i| \leq \sqrt{P}$ by the peak limit when applied to the samples. The set $\mathcal{X}$ touches the N -cube at the center of each of its faces at the points $x = \dots 0, 0, 0 \mp \sqrt{P}, 0, 0, \dots$. Thus, all the faces of the hypercube are supporting hyperplanes of $\mathcal{X}$. There are many points of $\mathcal{X}$ on each such supporting hyperplane, since there are many different band-limited signals attaining maximal absolute value at any $t = t_m$. The set $\mathcal{X}$ does not reach most of the N -cube vertexes (corners); however, it does touch the special vertexes $x = \mp \sqrt{P} [1, 1, \dots 1]$, which produce a constant $x(t) = \pm \sqrt{P}$. The distance from the origin to any vertex of the N -cube is larger by a factor of $\sqrt{N}$ relative to the distance from the origin to the centers of the N -cube faces. Using (3) and (4) and shifting the Nyquist-rate sampling times by a fraction of the sampling interval from $kT, k = 1 \dots N$, to $kT + \tau$ transforms x from one set of orthogonal coordinates $\varphi(t - iT)$ to another $\varphi(t - \tau - iT)$ and creates an additional hypercube, confining $\mathcal{X}$ and having the same properties of the faces.

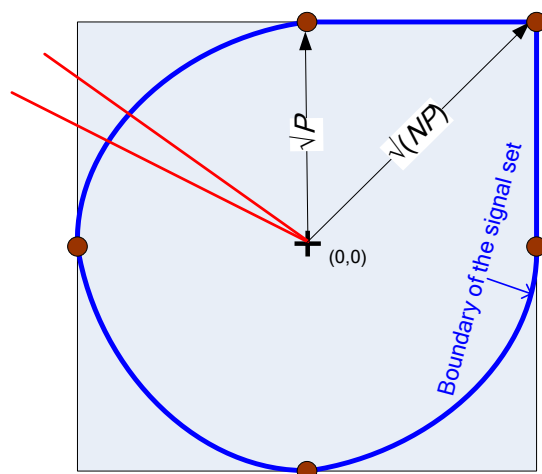


Figure 2. Illustration of the shape of the signal set. The square represents the N -cube. The red lines represent the cone used in Section 3.3.2.

3.2. Volume-Based Lower Bound

The geometrical lower bound on the capacity C , based on the Entropy Power Inequality (EPI), was presented in [18]. We summarize the relevant expressions below.

$$C \geq 0.5 \log_2 \left(\frac{\gamma \cdot P}{N_0 B} + 1 \right) = 0.5 \log_2 (\gamma \cdot \rho + 1) \quad (5)$$

The power factor $\gamma \leq 1$ is

$$\gamma = \frac{V_x^{\frac{2}{N}}}{2\pi e} \quad (6)$$

where V_x is the volume of the signal set $\mathcal{X}$ evaluated for signals with power $P = 1$.

An updated numerical method with an accuracy analysis added is presented next. The accuracy analysis is independent of the shape of $\mathcal{X}$, rendering it vastly simpler than

its counterpart in [18]. As stated above, all the vectors x of samples comprising $\mathcal{X}$ are in an N -cube. To evaluate V_x , the components of x are generated at random u.i.i.d in $-\sqrt{P} \leq x_i \leq \sqrt{P}$. This covers the volume of the hypercube uniformly and is repeated N_t times. Each x is tested if being in $\mathcal{X}$, that is, not violating the peak limit in the transmission at any time t . Then V_x is the probability p_c of x being inside $\mathcal{X}$ multiplied by the volume $(2\sqrt{P})^N$ of the hypercube, yielding

$$V_x = p_c \cdot (2\sqrt{P})^N \quad (7)$$

and, from (6), the power factor, evaluated without loss of generality at $P = 1$, is

$$\gamma = \frac{2p_c^{2/N}}{\pi e} \quad (8)$$

next, from (5)

$$C \geq 0.5 \log \left(\frac{P}{\sigma_n^2} \cdot \frac{2p_c^{2/N}}{\pi e} + 1 \right), \quad (9)$$

thus, the power factor is a product of $\frac{2}{\pi e}$ present in the asymptotic high SNR upper bound [9] based on the discrete-time channel and of the new term $p_c^{2/N}$, accounting for the peak constraint applied also to all times between the Nyquist-rate samples. This is a somewhat different method of evaluating the power factor γ relative to the method used in [18], so it is a partial verification. Furthermore, using this new and simpler method, the accuracy of the numerical evaluation can be controlled better, showing that the results in Figure 3 are accurate, as follows: The value of p_c is evaluated by generating $x(t)$ at random for N_t times. The estimated p_c is a binomial random variable, scaled by $1/N_t$, of mean p_c and standard deviation of

$$\sigma_p = \sqrt{p_c(1 - p_c)/N_t} \quad (10)$$

To plot a confidence interval of the lower bound (9), we can plot the bound along with lines offsetting p_c by $\pm 4\sigma_p$. This is shown in Figure 3 using an oversampling factor $N_k = 30$.

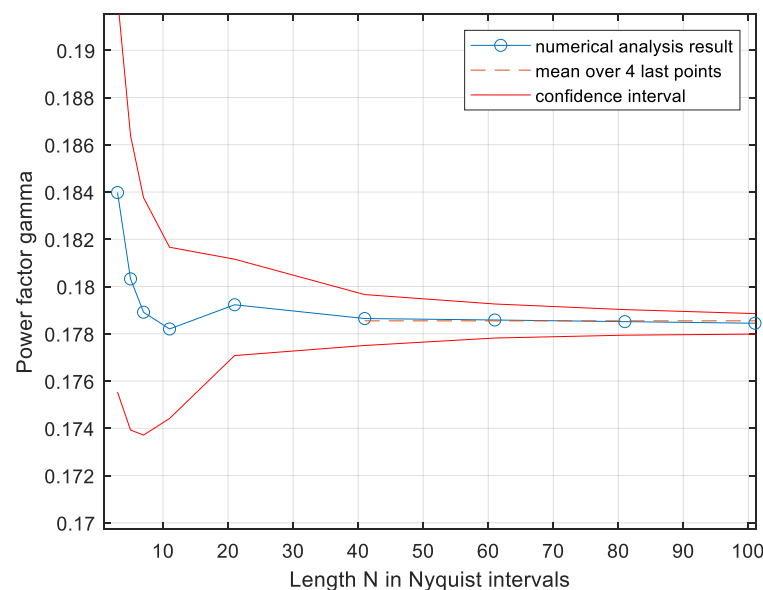


Figure 3. Lower bound on the power factor as a function of the sequence length. The red lines are the confidence interval $\pm 4\sigma_p$. The number of simulations N_t was increased with increasing N to yield about 1000 hits inside the signal set.

Thus, the power factor is 0.1785 and is very stable with respect to increasing N . This is the same value as the one derived in Figure 12 of [18] by a different numerical method. The stability with rising N , the narrow confidence interval of less than 1% of the power factor itself, and the similarity, at sufficiently large N , of the standard low-pass signals to the CP-FDE signals used in the analysis (presented at the end of the system model section) lead to the following conjecture:

Conjecture 1. *The result $\gamma = 0.1785$ is also valid for infinite N and for general low-pass signals that are not necessarily a CP-FDE type.*

3.3. Geometry-Based Upper Bound

3.3.1. Analysis

The capacity per sample is

$$C = \frac{1}{N} [h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x})] = \frac{1}{N} [h(\mathbf{y}) - h(\mathbf{n})] = \frac{1}{N} h(\mathbf{y}) - 0.5 \log(2\pi e \sigma_n^2) \quad (11)$$

maximized over the PDF of the transmit signals $\mathbf{x}$. An interesting progress on a geometry-based upper bound over a discrete-time MIMO channel was achieved in [26]. One of its results, section *Asymptotic gap* there on the tightness of the EPI-based lower bound, applies to our problem (explanation: Equation (7) of [26] is the parallel of our (1) in its vector form with a finite N provided $\mathbf{H} = \mathbf{I}_{N \times N}$, and our problem conforms to the convexity constraint on the signal set $\mathcal{X}$ as required in [26]). Thus, at an asymptotically high SNR, the lower bound (5), based on EPI, is tight. However, for the continuous-time channel here, we could not derive the intrinsic volumes of $\mathcal{X}$ required in the Steiner formula as performed in discrete-time in [26], so we are not able to derive quantitative results based on those, e.g., how high the SNR needs to be for the upper bound to be tight. To achieve this, we used a different approach. We proceeded in two stages: First, we established the following Theorem 2. Then we developed a numerical procedure upper-bounding the volume V_w in (12) and upper bounding the capacity (11) using (12).

Theorem 2. *The entropy of the received signal is upper bounded as follows*

$$h(\mathbf{y}) \leq \log(V_w) \quad (12)$$

where V_w is volume of $\mathcal{W}$ and $\mathcal{W}$ is the Minkowski sum

$$\mathcal{W} = \mathcal{X} + R_f \cdot \mathcal{B}, \quad (13)$$

$\mathcal{B}$ denotes an N -ball of unit radius and $R_f = \sigma_n \cdot \sqrt{2e} \left[\Gamma\left(\frac{N}{2} + 1\right) \right]^{\frac{1}{N}}$, which is approximately $R_f \approx \sigma_n \cdot \sqrt{N}$.

Remark: The Minkowski sums, similar to (13), of signal sets and noise sets appear in previous works, e.g., [26] and the references within, usually with infinite N , leading to the “ball-hardening” of the noise. In our case, with a finite N , the received signal cannot be confined to a finite volume (see Figure 4) and the object $\mathcal{B}$ in (13) is not a noise set, but rather a body, illustrated in Figure 5, the addition of which to $\mathcal{X}$ ensures the bound (12); see the following proof:

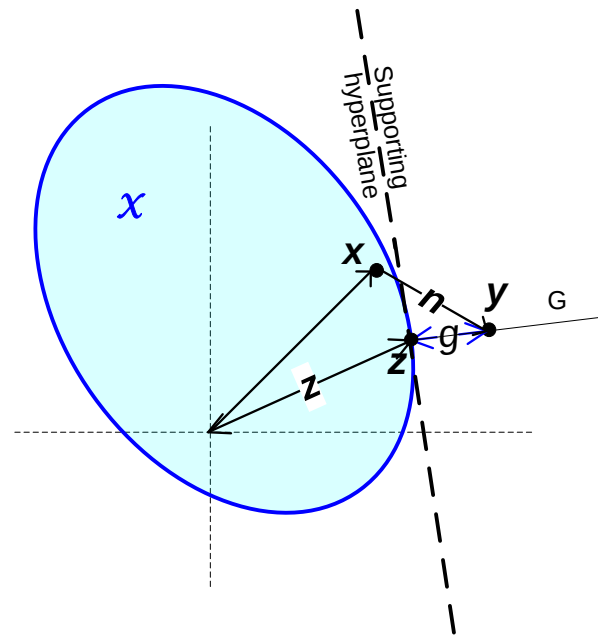


Figure 4. The supporting hyperplane, the transmitted signal x in the set $\mathcal{X}$, the noise n , the received signal y and the line G .

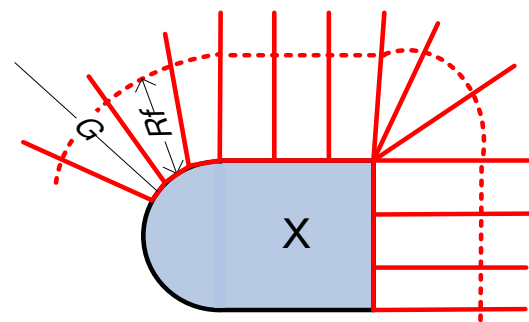


Figure 5. The gray signal set $\mathcal{X}$ and the red sections surrounding it. The dashed line illustrates the boundaries of the entropy-equivalent sections.

Proof of Theorem 2. We start with the attributes of the PDF $p_y(y)$ of the received signal at the region outside of $\mathcal{X}$. For a pair x, y the corresponding noise is $n(x, y) = y - x$ with N -dimensional normal distribution and variance per dimension of σ^2 . Then

$$p(y) = \int_{\text{all } x} p(x)p(y|x)dx = \int_{\text{all } x} p(x)p[n(x, y)]dx \quad (14)$$

Now, let us introduce a supporting (tangent) hyperplane touching the convex body $\mathcal{X}$ at a point z on it (see Figure 4). Line G , starting at z , is perpendicular to the supporting hyperplane. Consider a point y on G .

Note that the supporting hyperplane is $N-1$ dimensional and the distance g from it to the point y is one-dimensional. Note that, for any z on the surface of $\mathcal{X}$, there is at least one supporting hyperplane containing it and the associated perpendicular line marked G in Figure 4. By the normal distribution of n ,

$$p[n(y, x)] = C_1(x) e^{-0.5((g+C_2(x))/\sigma_n)^2}$$

for some pair of positive scalars $C_1(x)$, $C_2(x)$. Then, following (14), the slowest decaying $p_y(y)$ with respect to g , used in the proof of Proposition 3 below, is

$$p_y^s(\mathbf{y}) = C_4 \cdot e^{-0.5(g/\sigma_n)^2} \quad (15)$$

for some constant C_4 . The slowest decaying $p_y^s(\mathbf{y})$ denotes that, for any $p_y(\mathbf{y})$, the ratio $p_y(\mathbf{y})/p_y^s(\mathbf{y})$ is non-increasing in g . Denote $p_y^s(g) = C_4 \cdot e^{-0.5(g/\sigma_n)^2}$.

Next, we partition the volume of the received signals $\mathbf{y}$ outside of $\mathcal{X}$ into sections constructed as follows:

1. The base of each section is a small area on the $N-1$ -dimensional surface of $\mathcal{X}$ around some point denoted $\mathbf{z}$. or a small $M-1$ —dimensional body, $M < N$, around $\mathbf{z}$ in the surface of $\mathcal{X}$ or the point itself.
2. The section extends from the base outward along a line G perpendicular to the supporting hyperplane at $\mathbf{z}$. The section is narrow enough to render $p_y(\mathbf{y})$ a function $p_y(g)$ of g only where g is the distance of $\mathbf{y}$ in the section from the supporting hyperplane. There may be several supporting hyperplanes at any $\mathbf{z}$, e.g., at the special vertexes presented in Section 3.1 above.
3. The intersection between the section and a hyperplane parallel to the supporting hyperplane and at a distance g from it has an $(N-1)$ volume $S(g)$. The volume $S(g)$ is a non-decreasing function of g .

Figure 5 illustrates examples of such sections. A section, a base of which is on a flat part of the surface of $\mathcal{X}$, may be a cylinder with a constant $S(g)$. A section the base of which is the vertex of the N -cube will include cones (single-nappe infinitely extending cones, that is, corresponding to values of g from zero to infinity), with $S(g)$ proportional to

$$S(g) = C_3 \cdot g^{N-1}. \quad (16)$$

The growth of $S(g)$ in (16) is the fastest needed one, since such sections can cover an N -ball from a single center point $\mathbf{z}$; see also the relation to the Steiner formula below.

Each such section extends to infinity, that is, g extends from 0 to infinity. We want to calculate the entropy contributed to $h(\mathbf{y})$ by a single section, indexed i . This contribution to entropy is the N -dimensional integral

$$h_i = - \int_{\text{section } i} p_y(\mathbf{y}) \log(p_y(\mathbf{y})) d\mathbf{y}$$

Denote the probability of $\mathbf{y}$ to be in a section i as p_i . The volume of each section can be integrated in slices taken at each value of g , the volume of each slice being $dg \cdot S(g)$; the volume of the section is

$$V_i = \int_0^\infty S(g) dg \quad (17)$$

which is infinite, and

$$h_i = \int_0^\infty p_y(g) S(g) \log \frac{1}{p_y(g)} dg \quad (18)$$

and

$$p_i = \int_0^\infty p_y(g) S(g) dg \quad (19)$$

Note that, by (15), $p_y(g)$ is a decreasing function of g and $S(g)$ is an increasing one by the definition of the sections construction above. Let us consider a situation in which p_i is fixed. Then the following proposition holds:

Proposition 3. *The $p_y(g)$ which maximizes h_i for a given p_i and $S(g)$ is the slowest decreasing one $p_y^s(g)$ in (15).*

Proof. Let there be two different functions $p_y(g)$, $p_y^s(g)$, both upholding (19). Such two functions must have crossings due to (19) and due to identical p_i , and there can be only one such crossing due to $p_y^s(g)$ being the slowest decreasing. The function $p_y^s(g)$ can be obtained from $p_y(g)$ by incremental steps described below, each step moving some of the function mass from lower to higher region of g . This modification will enlarge h_i , since this operation moves probability mass from regions with lower values of $\log \frac{1}{p_y(g)}$ to regions with higher values (see (18)). Each step transforms $p_{y1}(g)$ into $p_{y2}(g)$ so that at the first step $p_{y1}(g) = p_y(g)$ and at the last step $p_{y2}(g) = p_y^s(g)$. At each step, $p_{y1}(g)$ is reduced in a narrow region dg of g around g_1 and increased around a narrow region of g around g_2 , such that $g_1 < g_2$, adjusting the increment to keep p_i in (19) constant while accounting for $S(g)$. Thus, to produce $p_{y2}(g)$, $p_{y1}(g_1)$ will be decremented by some dp and $p_{y1}(g_2)$ will be incremented by $dp \frac{S(g_1)}{S(g_2)}$. Define $p_{y3}(g)$ as similar to $p_{y2}(g)$, except that dp is increased so as to yield $p_{y3}(g_1) = p_{y3}(g_2)$. Clearly, h_i , associated with p_{y3} , will be at least as large as that associated with p_{y1} , since uniform distribution maximizes entropy. Since $p_{y2}(g)$ is a linear combination of $p_{y1}(g)$ and of $p_{y3}(g)$, the transition from p_{y1} to p_{y2} increases h_i in (18) by the concavity of the entropy function. $\square$

So, inserting (15) and (16) into (18) will yield the upper bound on the entropy in a section with $S(g)$ as given in (16). Consider evaluating (18) after such an insertion parametrized by $p_i = 1$ and by $S(g)$ being the surface of an N -ball of radius g . In this case, $p_y(g)$ is the PDF of the N -dimensional Gaussian noise; thus, h_i will be equal to the corresponding entropy, which is

$$h_n = \frac{N}{2} \log(2\pi e \sigma_n^2).$$

Now let us evaluate (18) under the same conditions and using the same $p_y(g)$, except for setting an arbitrary p_i and with $S(g)$ reduced by multiplying it by a small factor α . This yields a narrow conical section. The changes will scale $p_y(g)$ by $\frac{p_i}{\alpha}$. Then (18) yields

$$\begin{aligned} h_i &= \int_0^\infty \frac{p_i}{\alpha} p_y(g) \alpha S(g) \log \frac{\alpha}{p_i \cdot p_y(g)} dg \\ h_i &= p_i \int_0^\infty p_y(g) S(g) \log \frac{1}{p_y(g)} dg + p_i \log \frac{\alpha}{p_i} \int_0^\infty p_y(g) S(g) dg \\ h_i &= p_i h_n + p_i \log \frac{\alpha}{p_i} \end{aligned} \quad (20)$$

For the sake of upper-bounding $h(\mathbf{y})$, we shall replace each section i by an equivalent section having the same h_i , the same probability p_i , and the same shape $S_i(g)$ as the original section, but limited to the region $0 < g < R_f$. The volume of the equivalent section is denoted V'_i . The vector $\mathbf{y}$ occurs in the equivalent section with probability p_i with the uniform PDF $p_i(\mathbf{y}) = p_i / V'_i$. R_f is set so that h_i is preserved. The equivalence of p_i and h_i in each section renders the global new PDF legitimate. The new PDF has support of a limited volume V_W , the logarithm of which is an upper bound on entropy, and (12) follows. The contribution of the equivalent section to entropy is

$$h_i = -p_i \log(p_i / V'_i) = p_i \log(V'_i) - p_i \log(p_i) \quad (21)$$

Let V'_i fit into the conical section (16) above up to a radius R_f . Then

$$V'_i = \alpha V_B R_f^N$$

where V_B is the volume of a unity radius N -ball and (21) becomes

$$h_i = p_i \log(V_B R_f^N) + p_i \log \frac{\alpha}{p_i}$$

So, to upper bound the section entropy (20) by that of the uniformly distributed V_i' we need

$$\log(V_B R_f^N) = h_i = h_n \quad (22)$$

and this is independent of α and p_i . As presented in [9] and later termed the ‘noise ball hardening effect’, this yields, for large N , the radius of the noise sphere of

$$R_f \approx \sigma_n \sqrt{N}.$$

This is the approximate value of R_f stated in Theorem 2, and is easily verified using Stirling’s formula, as presented in [27]. So, placing a cone limited to radius R_f with the signal $\mathbf{y}$ distributed uniformly within provides at least the same entropy gain h_i as the original cone with unlimited g ; thus, the conical sections which extend to infinity in Figure 5 can be replaced by the same conical sections, but with limited radius R_f , as illustrated by the dashed line in Figure 5 and with uniform pdf of $\mathbf{y}$ in each section. The exact evaluation of R_f from (22), without the large N approximation, starts with the exact volume of an N -ball with a unity radius (e.g., [27]):

$$V_B = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2} + 1\right)}$$

from (22)

$$\begin{aligned} \log\left(\frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2} + 1\right)} R_f^N\right) &= N \log(\sigma_n \cdot \sqrt{2\pi e}) \\ R_f^N &= \frac{(\sigma_n \cdot \sqrt{2\pi e})^N}{\pi^{N/2}} \Gamma\left(\frac{N}{2} + 1\right) \\ R_f &= \sigma_n \cdot \sqrt{2e} \left[\Gamma\left(\frac{N}{2} + 1\right)\right]^{1/N} \end{aligned} \quad (23)$$

as stated in the theorem. The other extreme case is the cylindrical section, that is, constant $S(g) = S_0$; we shall have from (15) and (19)

$$p_y(g) = \frac{2p_i}{S_0} p_G(g)$$

where $p_G(g)$ is the PDF of the normal variable with zero mean and standard deviation of σ_n . Then, from (18),

$$h_i = S_0 \int_0^\infty \frac{2p_i}{S_0} p_G(g) \log \frac{1}{\frac{2p_i}{S_0} p_G(g)} dg = p_i \log(\sigma_n \cdot \sqrt{2\pi e}) + p_i \log\left(\frac{S_0}{2p_i}\right)$$

while the volume of a cylinder of length R_f is

$$V_i = S_0 R_f.$$

Then, equating the last h_i with the entropy (21) of $\mathbf{y}$ distributed uniformly in V_i ,

$$p_i \log(\sigma_n \cdot \sqrt{2\pi e}) + p_i \log\left(\frac{S_0}{2p_i}\right) = p_i \log(S_0 R_f) - p_i \log(p_i)$$

$$R_f = \sigma_n \cdot \sqrt{\frac{\pi e}{2}} \quad (24)$$

which is much smaller than for the conical section. Another trackable section type is a cone-like shape with $S(g) = C_5 \cdot g^{M-1}$, $2 \leq M < N$ and with a volume being M -th power of R_f by (17). Its radius R_f is readily derived as in (23) with N replaced by M ; such a radius is smaller than R_f in (23). Indeed, inserting $N = 1$ into (23) yields (24), as expected, considering that (16) with unity N yields the cylindrical shape. The combination of the conical and cylindrical sections above yields an equation of the volume of the part external to $\mathcal{X}$ of $\mathcal{X} + R\mathcal{B}_N$ for any positive R since, using the Steiner formula (e.g., [28]), such volumes are polynomials of degree N in R . It is possible that additional shape types are needed for the partition of the whole volume into the sections; if so, arbitrary additional types conforming to the sections description above can be formed with no increase in R_f , as presented in Appendix A. End of proof of Theorem 2. $\square$

Theorem 2 has some similarity to section V in [26]; the main difference is that Theorem 2 does not involve the limit of an infinite number of channel uses. In principle, the radius R_f can be reduced relative to (23) according to the local curvature of the surface of $\mathcal{X}$ (see Figure 5). However, to control the complexity and reliability of the numerical procedures, we shall use the maximal radius (23) all over, this will render the bound less tight. Equivalently, using the ball of the radius given by (23) is a tight bound around the vertexes and it exaggerates the entropy in the vicinity of flat faces, and the Minkowski addition of a ball in Theorem 2 overestimates the capacity.

Note that, at an infinite SNR, the set $\mathcal{W}$ is identical to $\mathcal{X}$, as evident from the vanishing R_f in (23), thus, the upper bound becomes equal to the lower bound (5), regardless of the shape of $\mathcal{X}$, as predicted also by [26].

3.3.2. Numerical Analysis of the Upper Bound

We need to upper bound the volume V_w of $\mathcal{W}$ to upper-bound the entropy (12) and then the capacity (11). Since we do not have a simple expression of the shape of $\mathcal{X}$ (see Figure 2), we need a numerical evaluation, as presented next.

The numerical evaluation of V_w is based on the method used to evaluate V_x presented in [18] with an extension. As in [18], we partition the space of the signals $\mathbf{y}$ and $\mathbf{x}$ into narrow conical sections (N -cones) with apexes at the origin, as illustrated by the red lines in Figures 2 and 6. The conical sections are visited at random by generating random $x(t)$, as in [18]; each $x(t)$ is scaled so that $\max_t |x(t)| = \sqrt{P}$ so the vector $\mathbf{x}$ of samples is on the surface of $\mathcal{X}$. Then, we evaluate $r = |\mathbf{x}|$, the local radius of $\mathcal{X}$, that is, the distance from the origin to the point $\mathbf{x}$ on the surface of $\mathcal{X}$. The volume V_x in [18] is evaluated from many values of r . Here, instead of using r directly to evaluate V_x , as in [18], we use r_w , which is the distance from the origin to the surface of $\mathcal{W}$, to evaluate V_w . The radius r_w is upper bounded as

$$r_w = r \frac{\sqrt{P} + R_f}{\sqrt{P}} \quad (25)$$

justified as follows: At some $t = t_m$, we have $|x(t_m)| = \sqrt{P}$. We represent $x(t)$ in its vector form by the sequence of the Nyquist-rate samples at times $kT + \tau$ shifted by $0 < \tau < T$, such that t_m is one of the sampling times. Then, as explained in Section 3.1 above, $\mathbf{x}$ is a point on a supporting hyperplane of $\mathcal{X}$. Next, we upscale $\mathbf{x}$ to produce the point $\mathbf{y} = \mathbf{x} \frac{\sqrt{P} + R_f}{\sqrt{P}}$. Clearly $|y(t_m)| = \sqrt{P} + R_f$; thus, $\mathbf{y}$ is at a distance of at least R_f from $\mathbf{x}$. And this also holds for all other signals $x(t)$ due to the peak power constraint at time t_m . Then, by the Minkowski sum (13), $\mathbf{y}$ is on the surface of $\mathcal{W}$ or outside $\mathcal{W}$. Since r scales identically to $x(t)$, r_w in (25) is guaranteed to be a local radius reaching the surface of $\mathcal{W}$ or beyond.

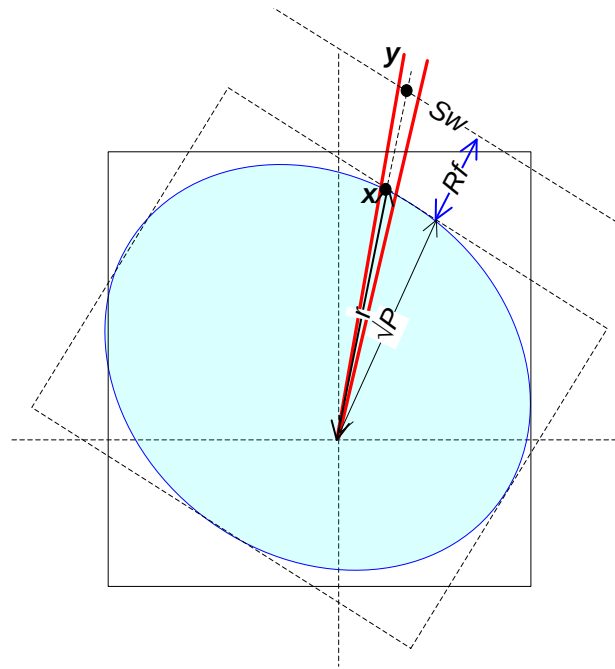


Figure 6. Elements of the upper-bound numerical analysis. The light blue ellipsoid is an illustration of $\mathcal{X}$. The red lines illustrate the N -cone. The full-line square represents the hypercube confining $\mathcal{X}$. The dashed square illustrates the hypercube formed by time-shifted Nyquist-rate samples. S_w is the surface of $\mathcal{W}$.

Additionally to (25), r_w is limited at each trial not to exceed an N -ball with a radius of $\sqrt{PN} + R_f$. The evaluation of V_W is performed as the evaluation of V_x in [18] including the importance sampling while r in [18] is replaced by $r_w(25)$.

The results are presented in Figure 7 compared to the EPI-based lower bound in (5) and (9), to the average power-limited capacity (2), and to the discrete-time upper bound using Equation (1) in Kellips [16]. We also compared to the unit process upper bound [23], which is equivalent to a power loss by a ratio of 0.9259 relative to (2), which might be improved, but not below a ratio of 0.5 due to lemma 3 in [22]; thus, the CGC-based approach cannot provide an upper bound on a capacity that is tight at a high SNR.

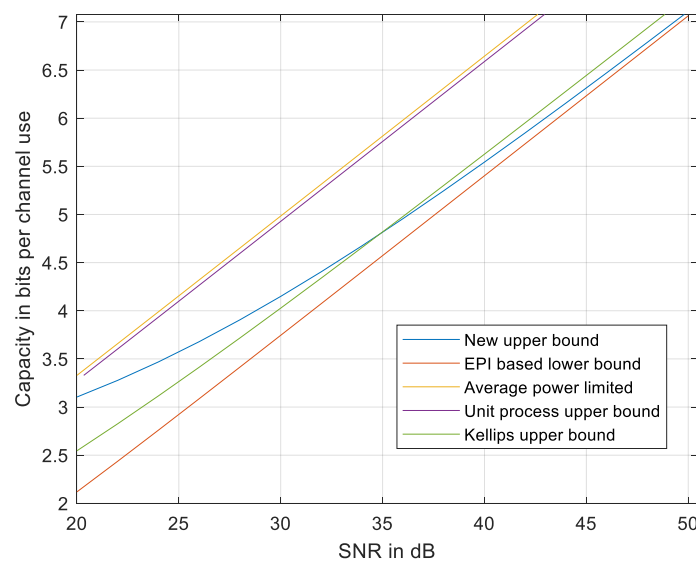


Figure 7. The new upper bound compared to the EPI-based lower bound [18], to the average power-limited capacity, to the unit process upper bound [23], and to the discrete-time upper bound using Kellips [16]. $N = 101$.

The gap between the upper and the lower bounds at an intermediate SNR is wider at high N and decreases at lower N as seen in Figure 8 with $N = 31$. Thus, the upper bound may be less tight at high N . Further improvement of the upper bound might be addressed in a future work that would refine the inequalities used in developing the upper bound, e.g., adapting R_f to the local curvature of the signal set $\mathcal{X}$ and evaluating h_i in each section, depicted in Figure 5, by a new method yet to be developed.

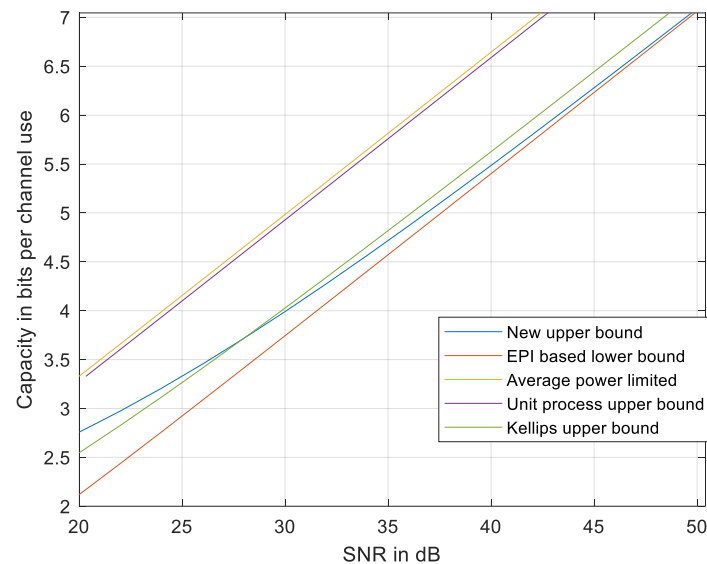


Figure 8. The new upper bound compared to the EPI-based lower bound [18], to the average power-limited capacity, to the unit process upper bound [21], and to the discrete-time upper bound using Kellips [16]. $N = 31$.

To verify the method and the software, we applied the peak limit on the Nyquist samples only. The results (see Figure 9) were expected to match the works on peak-limited discrete-time channels in [14,16,17], as presented in the Introduction. Indeed, the new upper bound and the upper bound in [16] on discrete-time channels plotted in Figure 9 do match at a high SNR while showing that, in this simple case, the tight upper bound in Kellips [16] is superior to ours at low and intermediate SNRs.

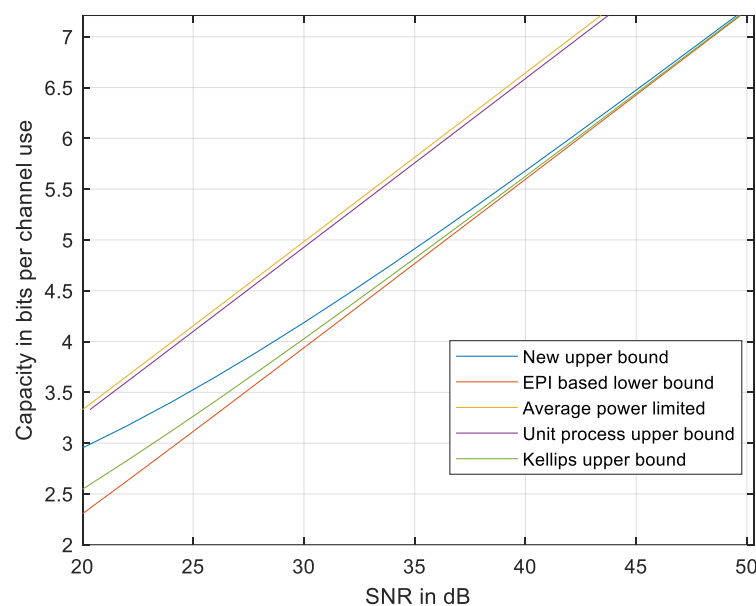


Figure 9. The bounds if the peak limit on $x(t)$ is applied at the sampling times kT only. $N = 31$.

4. Band-Pass Signals

4.1. Signal Representation

The analysis is readily extended to band-pass signals represented in the complex-valued baseband. We reuse the previous notation with modifications, as follows. The encoder produces a complex-valued low-pass signal $x(t)$ in the frequencies $|f| < 0.5B$. The noise is complex-valued with a power spectral density N_0 (two-sided). The signal is peak-limited, that is, $|x(t)| \leq \sqrt{P}$ for all t . The signal to noise ratio is defined as $\rho = \frac{P}{BN_0}$. The classical capacity per Nyquist-rate sample of the APL channel is

$$C_a = \log_2 \left(\frac{P}{N_0 B} + 1 \right) \quad (26)$$

The real and imaginary components of the noise have a power of $\sigma_n^2/2 = N_0 B/2$ each, and the per-symbol noise entropy is $h(n) = \log(\sigma_n^2 \cdot \pi e)$. The complex samples are represented by pairs of real numbers, and the volume V_x is computed in an $2N$ -dimensional real-valued space. The N -cubes containing $\mathcal{X}$ in the low-pass case, which can be considered Cartesian products of line segments each representing one peak-limited real-valued sample, are replaced by Cartesian products of N disks with a radius of $\sqrt{P}$, each representing one complex sample peak-limited to $\sqrt{P}$. Denote the set defined by this product for $P = 1$ by $\mathcal{D}$.

4.2. Lower Bound

The power factor and the EPI-based lower bound (see [18]) are now

$$\gamma = \frac{1}{\pi e} V_x^{\frac{1}{N}} \quad (27)$$

$$C \geq \log_2 \left(\frac{\gamma \cdot P}{N_0 B} + 1 \right) \quad (28)$$

The volume V_x is evaluated numerically as in the low-pass case while the N -cube with its volume of 2^N is replaced by the body $\mathcal{D}$ confining the signal set, the volume of which is $V_D = \pi^N$. Equation (7) is replaced by $V_x = p_c \cdot (\pi \sqrt{P})^N$ accordingly. The 4σ confidence interval is calculated by the same procedure as in the low-pass case. The results are presented in Figure 10.

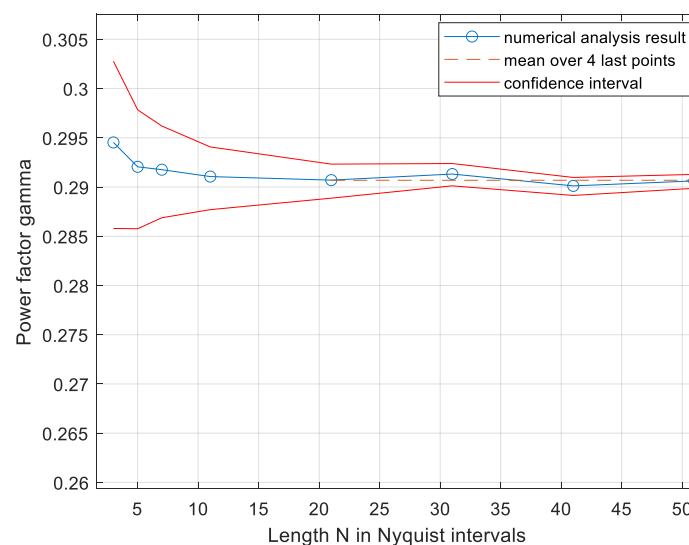


Figure 10. Lower bound on the power factor in the band-pass channel as a function of the sequence length. The red lines are the confidence interval $\pm 4\sigma_p$. The numbers of simulations N_t were increased with increasing N to yield about 1000 hits inside the signal set.

The estimated power factor is $\gamma = 0.2907$, identical to that in [18]; here, we were able to evaluate the 4σ confidence interval, which is less than 1% of the power factor.

4.3. Upper Bound

The upper bound in the band-pass case is evaluated similarly to the low-pass case. Both are evaluated in spaces of real numbers. The number of dimensions rises from N to $2N$ and the noise power in each real dimension is reduced by a factor of 2. This results in Corollary 4.

Corollary 4. *In the band-pass case,*

$$h(\mathbf{y}) \leq \log(V_w) \quad (29)$$

where V_w is volume of $\mathcal{W}$ and $\mathcal{W}$ is the Minkowski addition $\mathcal{W} = \mathcal{X} + R_f \cdot \mathcal{B}$. $\mathcal{B}$ is an $2N$ -ball of unit radius. R_f is

$$R_f = \sigma_n \cdot \sqrt{e} [\Gamma(N+1)]^{1/(2N)}, \quad (30)$$

the approximate value is $R_f \approx \sigma_n \cdot \sqrt{N}$.

Proof. Note that the proof of Theorem 2 was independent of the shape of $\mathcal{X}$. The difference in Corollary 4, relative to Theorem 2, is just replacing σ and N by $\sigma/\sqrt{2}$ and $2N$, respectively. The approximate expression of R_f follows immediately. The exact expression (30) follows by inserting $\sigma/\sqrt{2}$ and $2N$ into (23). $\square$

The numerical evaluation of the upper bound closely follows the low-pass case described above. The capacity of our continuous time channel is upper-bounded if the peak constraint is only imposed on the Nyquist-rate samples. The bounds on such a channel are presented in [17]. For simplicity, we present a simple EPI-based lower bound on the capacity of such a channel, Equation (38) in [15], which is

$$C_d \geq \log_2 \left(\frac{\text{SNR}}{e} + 1 \right) \quad (31)$$

It is close to capacity at a high SNR and is, of course, lower than any upper bound. The results are presented in Figure 11.

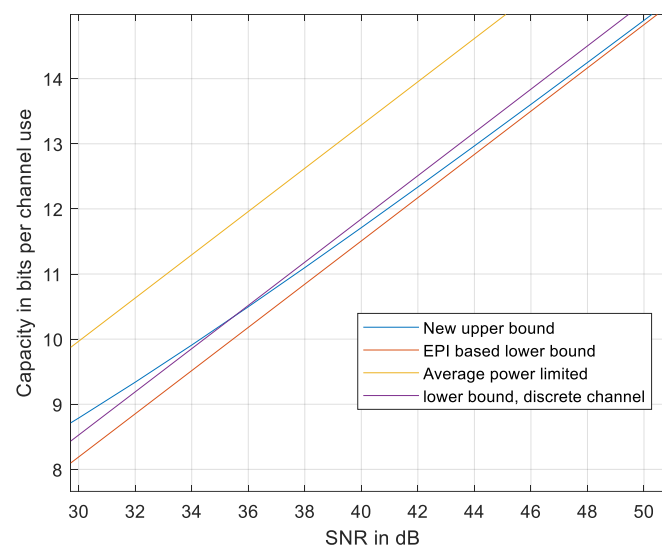


Figure 11. The new upper bound compared to the EPI-based lower bound [18], to the average power-limited capacity, and to the discrete-sampled lower bound using (31). $N = 51$.

As evident from Figure 11, the new upper bound is tighter at high enough SNR, relative to any upper bound based on the discrete-time channel. At even higher SNRs, the new upper bound meets the lower EPI-based bound (28). The bound is slightly nearer to the EPI-based lower bound at reduced N (see Figure 12).

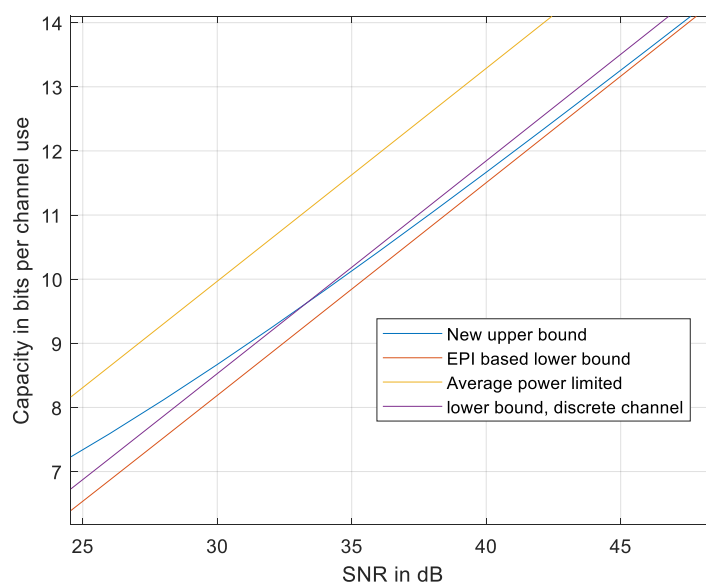


Figure 12. The new upper bound compared to the EPI-based lower bound [18], to the average power-limited capacity, and to the discrete-sampled lower bound using (31). $N = 31$.

The method was verified by applying the peak limit on the Nyquist-rate samples only with results qualitatively similar to those in the low-pass case in Figure 9.

5. Conclusions

The important problem of the capacity of Peak-Power Limited and Band Limited (PPBL) signals over the AWGN channel was investigated. We present Theorem 2 upper-bounding the entropy of the received signal leading to an upper bound on capacity; the upper bound is geometry-based and dependent on the shape of the signal set, not only on its volume. We computed an upper bound on capacity applicable to Cyclic Prefix-assisted Frequency Domain Equalization (CP-FDE) and OFDM-type signals. At a high SNR, the new bound is tighter than the previously known upper bounds and approaches the lower bound derived in previous work. We found that, at a high SNR, the change from average to peak power constraints with the same power parameters incurs a cost of about 7.5 dB in the low-pass case and about 5.4 dB in the band-pass case, as reflected by the power factor γ . The previous upper bounds on γ were $2/\pi e$ in [9] in the low-pass case and $1/e$ in [15] and in (31) in the band-pass case, corresponding to a smaller power cost of 6.3 dB and 4.34 dB, respectively. We verified the previous lower bound using a simpler numerical method and established the accuracy of the result. Previous bounds based on the volume of the transmitted signal set yielded EPI-based lower bounds, e.g., [18]. Here, we utilize the effective volume of the received signal set, which is influenced by the volume and the shape of the transmit signal set to derive an upper bound.

Future directions: The upper-bounding technique might be improved and the bound may be tightened, e.g., by more detailed exploitation of the geometry of the problem. The same geometry-based bounding technique should be applicable over the AWGN channel to other signal sets which are convex, such as some cases of MIMO, e.g., [26], the constrained Gaussian channel [20] and PPBL frequency-selective channels. The effort to find effective and practical modulation and coding techniques, e.g., [11,12], should be continued.

Author Contributions: Conceptualization, M.P. and S.S.; Methodology, M.P. and S.S.; Software, M.P.; Formal analysis, M.P. and S.S.; Investigation, M.P. and S.S.; Writing—original draft, M.P. and S.S.; Supervision, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the German Research Foundation (DFG) via the German-Israeli Project Cooperation (DIP) under Project SH 1937/1-1. 532.

Institutional Review Board Statement: Not applicable, since this work did not involve patients, humans, plants, or biological samples.

Data Availability Statement: Data is contained within the article and its figures.

Acknowledgments: We thank Luca Barletta for an insightful correspondence at the beginning of this work. We thank Neri Merhav for helpful and continual discussions.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A Reshaping the Sections

We presented above a partitioning of the region of the part of the received signal set $\mathcal{Y}$, exterior to the transmit signal set $\mathcal{X}$, into sections. The shape of section i is defined by its $S_i(g)$, its contribution to $h(\mathbf{y})$ is denoted h_i , and p_i is the probability that $\mathbf{y}$ is in the section. Each section is associated with its equivalent section having the same h_i and p_i , having the same $S_i(g)$ up to $g = R_f$, and a volume V'_i over which $\mathbf{y}$ is distributed uniformly. We presented conical-shaped and cylindrical-shaped sections. To ensure that the partition of the whole $\mathcal{Y}$ exterior to $\mathcal{X}$ into the sections is possible, we show below that the section types can be reshaped to any desired form and produce all of the shapes needed for the partition, with no need to increase R_f . As explained above, $S_i(g)$ is non-decreasing with increasing g before and after the reshaping.

The first type of reshaping involves the region $g > R_f$. We shall reshape a section characterized by $S_1(g)$ into a section characterized by $S_2(g)$ by moving some mass of $S_1(g)$ from the higher to the lower region of g . We derive $S_2(g)$ from $S_1(g)$ by increasing $S_1(g)$ by S_d in a narrow region of g , with a width dg around g_1 , and decreasing it by $S_d \frac{p_y(g_1)}{p_y(g_2)}$ around a narrow region of g around g_2 , such that $g_1 < g_2$; this keeps p_i in (19) constant. There are two cases: Case 1 is $R_f < g_1 < g_2$. In this case, h_i in (18) decreases by

$$p_y(g_1) S_d \log \frac{p_y(g_1)}{p_y(g_2)} dg \quad (32)$$

and so the R_f of the equivalent section can be reduced, as evidenced in (22). Case 2 is $g_1 < R_f < g_2$. As before, h_i decreases and, additionally, the increase in $S(g_1)$ increases the volume of the section. Then R_f decreases by both of these two factors.

The second type of reshaping is adding a volume element in the region $g < R_f$. This is performed in the following steps: First, $S_i(g)$ is increased anywhere in the region $g < R_f$ of the original set, adding volume to it and increasing its h_i and its p_i as, governed by (17), (18), and (19). Then $S_i(g)$ is modified identically in the equivalent section, but using the same $p_y(g)$ in the added region as in the original section. At this stage, the changes in the equivalent section of the volume, the shape $S_i(g)$, entropy h_i , and probability p_i , are identical to those in the original section; thus, R_f does not change. Finally, the PDF of $\mathbf{y}$ in the equivalent section is uniformized, keeping p_i constant; this final step does not decrease h_i in the equivalent section, enabling us to keep the original R_f or reduce it.

References

- Haider, M.F.; You, F.; He, S.; Rahkonen, T.; Aikio, J.P. Predistortion-Based Linearization for 5G and Beyond Millimeter-Wave Transceiver Systems: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 2029–2072. [\[CrossRef\]](#)
- Ugolini, A. Physical and MAC Layer Techniques for Next Generation Satellite Communications. Ph.D. Thesis, Università Degli Studi Di Parma, Parma, Italy, 2016.
- Lei, L. Advances in New Physical Layer Technologies for Next-Generation Wireless Communications. *Entropy* **2025**, *27*, 616. [\[CrossRef\]](#)
- Biglieri, E.; Proakis, J.; Shamai, S. Fading channels: Information-theoretic and communications aspects. *IEEE Trans. Inf. Theory* **1998**, *44*, 2619–2692. [\[CrossRef\]](#)
- Ochiai, H.; Imai, H. On the distribution of the peak-to-average power ratio in OFDM signals. *IEEE Trans. Commun.* **2001**, *49*, 282–289. [\[CrossRef\]](#)
- Savaux, V.; Louet, Y. PAPR Analysis as a ratio of two random variables: Application to multicarrier systems with low subcarriers number. *IEEE Trans. Commun.* **2018**, *66*, 5732–5739. [\[CrossRef\]](#)
- Ochiai, H. Peak Power Properties of Band-Limited Signals: With Pulse Shaping or Windowing. *IEEE Trans. Veh. Technol.* **2024**, *73*, 15200–15213. [\[CrossRef\]](#)
- Sayyari, R.; Pourrostam, J.; Ahmadi, H. Efficient PAPR reduction scheme for OFDM-NOMA systems based on DSI & precoding methods. *Phys. Commun.* **2021**, *47*, 101372. [\[CrossRef\]](#)
- Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.
- Modonesi, R.; Dalai, M.; Migliorati, P.; Leonardi, R. A Note on Probabilistic and Geometric Shaping for the AWGN Channel. *IEEE Commun. Lett.* **2020**, *24*, 2119–2122. [\[CrossRef\]](#)
- Kamal, S.; Khan, I.A.; Ali, A.; Farooqi, A.H.; Ahmad, A.; Estevez, C. Novel Nyquist-I filter to reduce PAPR for SC-FDMA scheme. *Telecommun. Syst.* **2025**, *88*, 34. [\[CrossRef\]](#)
- Levi, O.; Raphaeli, D.; Tate, Y. A Novel Shaping Scheme for PAPR Reduction in Single-Carrier Modulation. *IEEE Trans. Commun.* **2018**, *66*, 4222–4233. [\[CrossRef\]](#)
- Shamai, S. On the capacity of a Gaussian channel with peak power and bandlimited input signals. *AEU—Int. J. Electron. Commun.* **1998**, *42*, 340–346.
- Smith, J.G. The information capacity of amplitude- and variance-constrained scalar gaussian channels. *Inf. Control* **1971**, *18*, 203–219. [\[CrossRef\]](#)
- Shamai, S.; Bar-David, I. The Capacity of Average and Peak-Power-Limited Quadrature Gaussian Channels. *IEEE Trans. Inf. Theory* **1995**, *41*, 1060–1071. [\[CrossRef\]](#)
- McKellips, A.L. Simple tight bounds on capacity for the peak-limited discrete-time channel. In Proceedings of the International Symposium on Information Theory. ISIT 2004, Chicago, IL, USA, 27 June–2 July 2004; p. 348. [\[CrossRef\]](#)
- Thangaraj, A.; Kramer, G.; Böcherer, G. Capacity Bounds for Discrete-Time, Amplitude-Constrained, Additive White Gaussian Noise Channels. *IEEE Trans. Inf. Theory* **2017**, *63*, 4172–4182. [\[CrossRef\]](#)
- Peleg, M.; Shamai, S. On the Capacity of the Peak-Limited and Band-Limited Channel. *Entropy* **2024**, *26*, 1049. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, J.; Yang, L.-L.; Hanzo, L.; Gharavi, H. Advances in Cooperative Single-Carrier FDMA Communications: Beyond LTE-Advanced. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 730–756. [\[CrossRef\]](#)
- Shamai, S.; Bar-David, I. Upper bounds on capacity for a constrained Gaussian channel. *IEEE Trans. Inf. Theory* **1989**, *35*, 1079–1084. [\[CrossRef\]](#)
- Shepp, L.A. Covariance of unit processes. In Proceedings of the Working Conference on Stochastic Processes, Santa Barbara, CA, USA, 18–20 April 1967; pp. 205–218.
- Shamai, S.; Chayat, N. The autocorrelation function of a peak-power-limited process. *Signal Process.* **1991**, *24*, 127–136. [\[CrossRef\]](#)
- Tzachy, I.; Dabora, R.; Shamai, S. An Upper Bound on the Capacity of Bandlimited LTI AWGN Channels subject to Peak-Amplitude Constraint. In Proceedings of the 2025 IEEE International Symposium on Information Theory (ISIT 2025), Ann Arbor, MI, USA, 22–27 June 2025.
- Bar-David, I. On the Capacity of Peak Power Constrained Gaussian Channels. In *Performance Limits in Communication Theory and Practice*; Springer: Dordrecht, The Netherlands, 1988. [\[CrossRef\]](#)
- Dooley, S.R.; Nandi, A.K. Notes on the interpolation of discrete periodic signals using sinc function related approaches. *IEEE Trans. Signal Process.* **2000**, *48*, 1201–1203. [\[CrossRef\]](#)
- Favano, A.; Ferrari, M.; Magarini, M.; Barletta, L. A Sphere Packing Bound for Vector Gaussian Fading Channels Under Peak Amplitude Constraints. *IEEE Trans. Inf. Theory* **2023**, *69*, 238–250. [\[CrossRef\]](#)

27. Ball, K.M. An Elementary Introduction to Modern Convex Geometry. In *Flavors of Geometry*; MSRI Publications: Berkeley, CA, USA, 1997; Volume 31, pp. 1–58. ISBN 10-0-521-62048-1.
28. Hug, D.; Weil, W. *Lectures on Convex Geometry*; Springer: Cham, Switzerland, 2021; ISBN 978-3-030-50182-2.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.