Problem Set 3 Solutions

Problem 3.1 (Invariance of coding gain)

(a) Show that in the power-limited regime the nominal coding gain $\gamma_c(\mathcal{A})$ of (5.9), the UBE (5.10) of $P_b(E)$, and the effective coding gain $\gamma_{\rm eff}(\mathcal{A})$ are invariant to scaling, orthogonal transformations and Cartesian products.

In the power-limited regime, the nominal coding gain is defined as

$$\gamma_c(\mathcal{A}) = \frac{d^2_{\min}(\mathcal{A})}{4E_b(\mathcal{A})}.$$

Scaling $\mathcal{A}$ by $\alpha > 0$ multiplies both $d^2_{\min}(\mathcal{A})$ and $E_b(\mathcal{A})$ by α^2, and therefore leaves $\gamma_c(\mathcal{A})$ unchanged. Orthogonal transformations of $\mathcal{A}$ do not change either $d^2_{\min}(\mathcal{A})$ or $E_b(\mathcal{A})$. As we have seen in Problem 2.1, taking Cartesian products also does not change either $d^2_{\min}(\mathcal{A})$ or $E_b(\mathcal{A})$. Therefore $\gamma_c(\mathcal{A})$ is invariant under all these operations.

The UBE of $P_b(E)$ involves $\gamma_c(\mathcal{A})$ and $K_b(\mathcal{A}) = K_{\min}(\mathcal{A})/(|\log|\mathcal{A}|)$. $K_{\min}(\mathcal{A})$ is also obviously unchanged under scaling or orthogonal transformations. Problem 2.1 showed that $K_{\min}(\mathcal{A})$ increases by a factor of K under a K-fold Cartesian product, but so does $\log|\mathcal{A}|$, so $K_b(\mathcal{A})$ is also unchanged under Cartesian products.

The effective coding gain is a function of the UBE of $P_b(E)$, and therefore it is invariant also.

(b) Show that in the bandwidth-limited regime the nominal coding gain $\gamma_c(\mathcal{A})$ of (5.14), the UBE (5.15) of $P_s(E)$, and the effective coding gain $\gamma_{\rm eff}(\mathcal{A})$ are invariant to scaling, orthogonal transformations and Cartesian products.

In the bandwidth-limited regime, the nominal coding gain is defined as

$$\gamma_c(\mathcal{A}) = \frac{(2^{\rho(\mathcal{A})} - 1)d^2_{\min}(\mathcal{A})}{6E_s(\mathcal{A})}.$$

Scaling $\mathcal{A}$ by $\alpha > 0$ multiplies both $d^2_{\min}(\mathcal{A})$ and $E_s(\mathcal{A})$ by α^2 and does not change $\rho(\mathcal{A})$, and therefore leaves $\gamma_c(\mathcal{A})$ unchanged. Orthogonal transformations of $\mathcal{A}$ do not change $d^2_{\min}(\mathcal{A})$, $E_b(\mathcal{A})$ or $\rho(\mathcal{A})$. As we have seen in Problem 2.1, taking Cartesian products also does not change $d^2_{\min}(\mathcal{A})$, $E_b(\mathcal{A})$ or $\rho(\mathcal{A})$. Therefore $\gamma_c(\mathcal{A})$ is invariant under all these operations.

The UBE of $P_s(E)$ involves $\gamma_c(\mathcal{A})$ and $K_s(\mathcal{A}) = (2/N)K_{\min}(\mathcal{A})$. $K_{\min}(\mathcal{A})$ is also obviously unchanged under scaling or orthogonal transformations. Problem 2.1 showed that $K_{\min}(\mathcal{A})$ increases by a factor of K under a K-fold Cartesian product, but so does N, so $K_s(\mathcal{A})$ is also unchanged under Cartesian products.

The effective coding gain is a function of the UBE of $P_s(E)$, and therefore it is invariant also.

Problem 3.2 (Orthogonal signal sets)

An orthogonal signal set *is a set* $\mathcal{A} = \{\mathbf{a}_j, 1 \le j \le M\}$ *of* M *orthogonal vectors in* $\mathbb{R}^M$ *with equal energy* $E(\mathcal{A})$*; i.e.,* $\langle \mathbf{a}_j, \mathbf{a}_{j'} \rangle = E(\mathcal{A})\delta_{jj'}$ *(Kronecker delta).*

(a) Compute the nominal spectral efficiency ρ *of* $\mathcal{A}$ *in bits per two dimensions. Compute the average energy* E_b *per information bit.*

The rate of $\mathcal{A}$ is $\log_2 M$ bits per M dimensions, so the nominal spectral efficiency is

$$\rho = (2/M)\log_2 M \text{ bits per two dimensions.}$$

The average energy per symbol is $E(\mathcal{A})$, so the average energy per bit is

$$E_b = \frac{E(\mathcal{A})}{\log_2 M}.$$

(b) Compute the minimum squared distance $d^2_{\min}(\mathcal{A})$*. Show that every signal has* $K_{\min}(\mathcal{A}) = M - 1$ *nearest neighbors.*

The squared distance between any two distinct vectors is

$$||\mathbf{a}_j - \mathbf{a}_{j'}||^2 = ||\mathbf{a}_j||^2 - 2\langle \mathbf{a}_j, \mathbf{a}_{j'} \rangle + ||\mathbf{a}_{j'}||^2 = E(\mathcal{A}) - 0 + E(\mathcal{A}) = 2E(\mathcal{A}),$$

so $d^2_{\min}(\mathcal{A}) = 2E(\mathcal{A})$, and every vector has all other vectors as nearest neighbors, so $K_{\min}(\mathcal{A}) = M - 1$.

(c) Let the noise variance be $\sigma^2 = N_0/2$ *per dimension. Show that the probability of error of an optimum detector is bounded by the UBE*

$$\Pr(E) \le (M-1)Q^{\surd}(E(\mathcal{A})/N_0).$$

The pairwise error probability between any two distinct vectors is

$$\Pr\{\mathbf{a}_j \to \mathbf{a}_{j'}\} = Q^{\surd}(||\mathbf{a}_j - \mathbf{a}_{j'}||^2/4\sigma^2) = Q^{\surd}(2E(\mathcal{A})/2N_0) = Q^{\surd}(E(\mathcal{A})/N_0).$$

By the union bound, for any $\mathbf{a}_j \in \mathcal{A}$,

$$\Pr(E \mid \mathbf{a}_j) \le \sum_{j' \ne j} \Pr\{\mathbf{a}_j \to \mathbf{a}_{j'}\} = (M-1)Q^{\surd}(E(\mathcal{A})/N_0),$$

so the average $\Pr(E)$ also satisfies this upper bound.

(d) Let $M \to \infty$ *with* E_b *held constant. Using an asymptotically accurate upper bound for the* $Q^{\surd}(\cdot)$ *function (see Appendix), show that* $\Pr(E) \to 0$ *provided that* $E_b/N_0 > 2\ln 2$ *(1.42 dB). How close is this to the ultimate Shannon limit on* E_b/N_0*? What is the nominal spectral efficiency* ρ *in the limit?*

By the Chernoff bound of the Appendix, $Q^{\surd}(x^2) \le e^{-x^2/2}$. Therefore

$$\Pr(E) \le (M-1)e^{-E(\mathcal{A})/2N_0} < e^{(\ln M)}e^{-(E_b \log_2 M)/2N_0}.$$

Since $\ln M = (\log_2 M)(\ln 2)$, as $M \to \infty$ this bound goes to zero provided that

$$E_b/2N_0 > \ln 2,$$

or equivalently $E_b/N_0 > 2\ln 2$ (1.42 dB).

The ultimate Shannon limit on E_b/N_0 is $E_b/N_0 > \ln 2$ (-1.59 dB), so this shows that we can get to within 3 dB of the ultimate Shannon limit with orthogonal signalling. (It was shown in 6.450 that orthogonal signalling can actually achieve $\Pr(E) \to 0$ for any $E_b/N_0 > \ln 2$, the ultimate Shannon limit.)

Unfortunately, the nominal spectral efficiency $\rho = (2\log_2 M)/M$ goes to 0 as $M \to \infty$.

Problem 3.3 (Simplex signal sets)

Let $\mathcal{A}$ be an orthogonal signal set as above.

(a) Denote the mean of $\mathcal{A}$ by $\mathbf{m}(\mathcal{A})$. Show that $\mathbf{m}(\mathcal{A}) \neq \mathbf{0}$, and compute $||\mathbf{m}(\mathcal{A})||^2$.

By definition,

$$\mathbf{m}(\mathcal{A}) = \frac{1}{M}\sum_j \mathbf{a}_j.$$

Therefore, using orthogonality, we have

$$||\mathbf{m}(\mathcal{A})||^2 = \frac{1}{M^2}\sum_j ||\mathbf{a}_j||^2 = \frac{E(\mathcal{A})}{M} \neq 0.$$

By the strict non-negativity of the Euclidean norm, $||\mathbf{m}(\mathcal{A})||^2 \neq 0$ implies that $\mathbf{m}(\mathcal{A}) \neq \mathbf{0}$.

The zero-mean set $\mathcal{A}' = \mathcal{A} - \mathbf{m}(\mathcal{A})$ (as in Exercise 2) is called a simplex signal set. *It is universally believed to be the optimum set of M signals in AWGN in the absence of bandwidth constraints, except at ridiculously low SNRs.*

(b) For $M = 2, 3, 4$, sketch $\mathcal{A}$ and $\mathcal{A}'$.

For $M = 2, 3, 4$, $\mathcal{A}$ consists of M orthogonal vectors in M-space (hard to sketch for $M = 4$). For $M = 2$, $\mathcal{A}'$ consists of two antipodal signals in a 1-dimensional subspace of 2-space; for $M = 3$, $\mathcal{A}'$ consists of three vertices of an equilateral triangle in a 2-dimensional subspace of 3-space; and for $M = 4$, $\mathcal{A}'$ consists of four vertices of a regular tetrahedron in a 3-dimensional subspace of 4-space.

(c) Show that all signals in $\mathcal{A}'$ have the same energy $E(\mathcal{A}')$. Compute $E(\mathcal{A}')$. Compute the inner products $\langle \mathbf{a}_j, \mathbf{a}_{j'} \rangle$ for all $\mathbf{a}_j, \mathbf{a}_{j'} \in \mathcal{A}'$.

The inner product of $\mathbf{m}(\mathcal{A})$ with any $\mathbf{a}_j$ is

$$\langle \mathbf{m}(\mathcal{A}), \mathbf{a}_j \rangle = \frac{1}{M}\sum_{j'} \langle \mathbf{a}_{j'}, \mathbf{a}_j \rangle = \frac{E_{\mathcal{A}}}{M}.$$

The energy of $\mathbf{a}'_j = \mathbf{a}_j - \mathbf{m}(\mathcal{A})$ is therefore

$$||\mathbf{a}'_j||^2 = ||\mathbf{a}_j||^2 - 2\langle \mathbf{m}(\mathcal{A}), \mathbf{a}_j \rangle + ||\mathbf{m}(\mathcal{A})||^2 = E(\mathcal{A}) - \frac{E(\mathcal{A})}{M} = \frac{M-1}{M}E(\mathcal{A}).$$

For $j \neq j'$, the inner product $\langle \mathbf{a}'_j, \mathbf{a}'_{j'} \rangle$ is

$$\langle \mathbf{a}'_j, \mathbf{a}'_{j'} \rangle = \langle \mathbf{a}_j - \mathbf{m}(\mathcal{A}), \mathbf{a}_{j'} - \mathbf{m}(\mathcal{A}) \rangle = 0 - 2\frac{E(\mathcal{A})}{M} + \frac{E(\mathcal{A})}{M} = -\frac{E(\mathcal{A})}{M}.$$

In other words, the inner product is equal to $\frac{M-1}{M}E(\mathcal{A})$ if $j' = j$ and $-\frac{1}{M}E(\mathcal{A})$ for $j' \neq j$.

(d) [Optional]. Show that for ridiculously low SNR, a signal set consisting of $M-2$ zero signals and two antipodal signals $\{\pm\mathbf{a}\}$ has a lower $\Pr(E)$ than a simplex signal set. [Hint: see M. Steiner, "The strong simplex conjecture is false," IEEE TRANSACTIONS ON INFORMATION THEORY, *pp. 721-731, May 1994.]*

See the cited article.

Problem 3.4 (Biorthogonal signal sets)

The set $\mathcal{A}'' = \pm\mathcal{A}$ of size $2M$ consisting of the M signals in an orthogonal signal set $\mathcal{A}$ with symbol energy $E(\mathcal{A})$ and their negatives is called a biorthogonal signal set.

(a) Show that the mean of $\mathcal{A}''$ is $\mathbf{m}(\mathcal{A}'') = \mathbf{0}$, and that the average energy is $E(\mathcal{A})$.

The mean is

$$\mathbf{m}(\mathcal{A}'') = \sum_j (\mathbf{a}_j - \mathbf{a}_j) = \mathbf{0},$$

and every vector has energy $E(\mathcal{A})$.

(b) How much greater is the nominal spectral efficiency ρ of $\mathcal{A}''$ than that of $\mathcal{A}$?

The rate of $\mathcal{A}''$ is $\log_2 2M = 1 + \log_2 M$ bits per M dimensions, so its nominal spectral efficiency is $\rho = (2/M)(1 + \log_2 M)$ b/2D, which is $2/M$ b/2D greater than for $\mathcal{A}$. This is helpful for small M, but negligible as $M \to \infty$.

(c) Show that the probability of error of $\mathcal{A}''$ is approximately the same as that of an orthogonal signal set with the same size and average energy, for M large.

Each vector in $\mathcal{A}''$ has $2M-2$ nearest neighbors at squared distance $2E(\mathcal{A})$, and one antipodal vector at squared distance $4E(\mathcal{A})$. The union bound estimate is therefore

$$\Pr(E) \approx (2M-2)Q^{\surd}(E(\mathcal{A})/N_0) \approx |\mathcal{A}''|Q^{\surd}(E(\mathcal{A})/N_0),$$

which is approximately the same as the estimate $\Pr(E) \approx (2M-1)Q^{\surd}(E(\mathcal{A})/N_0) \approx |\mathcal{A}|Q^{\surd}(E(\mathcal{A})/N_0)$ for an orthogonal signal set $\mathcal{A}$ of size $|\mathcal{A}| = 2M$.

(d) Let the number of signals be a power of 2: $2M = 2^k$. Show that the nominal spectral efficiency is $\rho(\mathcal{A}'') = 4k2^{-k}$ b/2D, and that the nominal coding gain is $\gamma_c(\mathcal{A}'') = k/2$. Show that the number of nearest neighbors is $K_{\min}(\mathcal{A}'') = 2^k - 2$.

If $M = 2^{k-1}$, then the nominal spectral efficiency is

$$\rho(\mathcal{A}'') = (2/M)(1 + \log_2 M) = 2^{2-k}k = 4k2^{-k} \text{ b/2D}.$$

We are in the power-limited regime, so the nominal coding gain is

$$\gamma_c(\mathcal{A}'') = \frac{d^2_{\min}(\mathcal{A}'')}{4E_b} = \frac{2E(\mathcal{A}'')}{4E(\mathcal{A}'')/k} = \frac{k}{2}.$$

The number of nearest neighbors is $K_{\min}(\mathcal{A}'') = 2M - 2 = 2^k - 2$.

Problem 3.5 (small nonbinary constellations)

(a) For $M = 4$, the (2×2)-QAM signal set is known to be optimal in $N = 2$ dimensions. Show however that there exists at least one other inequivalent two-dimensional signal set $\mathcal{A}'$ with the same coding gain. Which signal set has the lower "error coefficient" $K_{\min}(\mathcal{A})$?

The 4-QAM signal set $\mathcal{A}$ with points $\{(\pm\alpha, \pm\alpha)\}$ has $b = 2$, $d^2_{\min}(\mathcal{A}) = 4\alpha^2$ and $E(\mathcal{A}) = 2\alpha^2$, so $\mathcal{A}$ has $E_b = E(\mathcal{A})/2 = \alpha^2$ and $\gamma_c(\mathcal{A}) = d^2_{\min}(\mathcal{A})/4E_b = 1$.

The 4-point hexagonal signal set $\mathcal{A}'$ with points at $\{(0, 0), (\alpha, \sqrt{3}\alpha), (2\alpha, 0), (3\alpha, \sqrt{3}\alpha)\}$ has mean $\mathbf{m} = (3\alpha/2, \sqrt{3}\alpha)/2)$ and average energy $E(\mathcal{A}') = 5\alpha^2$. If we translate $\mathcal{A}'$ to $\mathcal{A}'' = \mathcal{A}' - \mathbf{m}$ to remove the mean, then $E(\mathcal{A}'') = E(\mathcal{A}') - ||\mathbf{m}||^2 = 5\alpha^2 - 3\alpha^2 = 2\alpha^2$. Thus $\mathcal{A}''$ has the same minimum squared distance, the same average energy, and thus the same coding gain as $\mathcal{A}$.

In $\mathcal{A}$, each point has two nearest neighbors, so $K_{\min}(\mathcal{A}) = 2$. In $\mathcal{A}'$, two points have two nearest neighbors and two points have three nearest neighbors, so $K_{\min}(\mathcal{A}') = 2.5$. (This factor of 1.25 difference in error coefficient will cost about $(1/4) \cdot (0.2) = 0.05$ dB in effective coding gain, by our rule of thumb.)

[Actually, all parallelogram signal sets with sides of length 2α and angles between 60° and 90° have minimum squared distance $4\alpha^2$ and average energy $2\alpha^2$, if the mean is removed.]

(b) Show that the coding gain of (a) can be improved in $N = 3$ dimensions. [Hint: consider the signal set $\mathcal{A}'' = \{(1, 1, 1), (1, -1, -1), (-1, 1, -1), (-1, -1, 1)\}$.] Sketch $\mathcal{A}''$. What is the geometric name of the polytope whose vertex set is $\mathcal{A}''$?

The four signal points in $\mathcal{A}''$ are the vertices of a tetrahedron (see Chapter 6, Figure 1). The minimum squared distance between points in $\mathcal{A}''$ is $2 \cdot 4 = 8$, and the average energy is $E(\mathcal{A}'') = 3$, so $E_b = 3/2$. Thus the coding gain of $\mathcal{A}''$ is $\gamma_c(\mathcal{A}'') = d^2_{\min}(\mathcal{A}'')/4E_b = 4/3$, a factor of 4/3 (1.25 dB) better than that of $\mathcal{A}$.

However, the nominal spectral efficiency ρ of $\mathcal{A}''$ is only 4/3 b/2D, compared to $\rho = 2$ b/2D for $\mathcal{A}$; *i.e.,* $\mathcal{A}''$ is less bandwidth-efficient. Also, each point in $\mathcal{A}''$ has $K_{\min}(\mathcal{A}'') = 3$ nearest neighbors, which costs about 0.1 dB in effective coding gain.

(c) Give an accurate plot of the UBE of the $\Pr(E)$ for the signal set $\mathcal{A}''$ of (b). How much is the effective coding gain, by our rule of thumb and by this plot?

The UBE for $\Pr(E)$ is

$$\Pr(E) \approx K_{\min}(\mathcal{A}'')Q^{\surd}(2\gamma_c(\mathcal{A}'')E_b/N_0) = 3Q^{\surd}(2\frac{4}{3}E_b/N_0).$$

Since each signal sends 2 bits, the UBE for $P_b(E)$ is $\frac{1}{2}\Pr(E)$: $P_b(E) \approx 1.5Q^{\surd}(2\frac{4}{3}E_b/N_0)$. An accurate plot of the UBE may be obtained by moving the baseline curve $P_b(E) \approx Q^{\surd}(2E_b/N_0)$ to the left by 1.25 dB and up by a factor of 1.5, as shown in Figure 1. This shows that the effective coding gain is about $\gamma_{\text{eff}}(\mathcal{A}'') \approx 1.15$ dB at $P_b(E) \approx 10^{-5}$. Our rule of thumb gives approximately the same result, since 1.5 is equal to about $\sqrt{2}$.

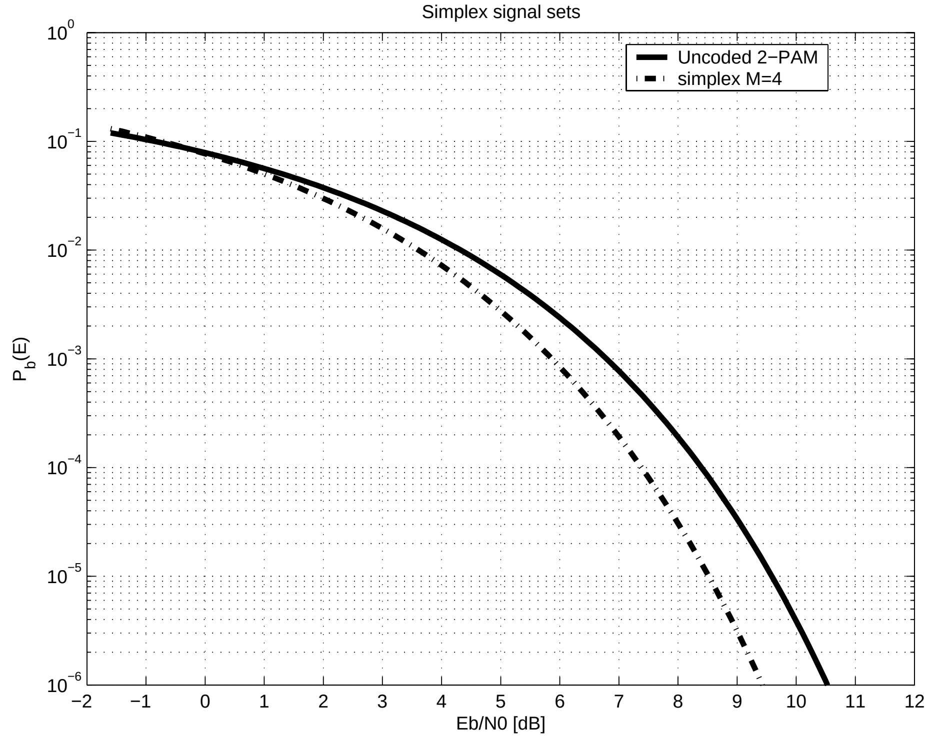

Figure 1. $P_b(E)$ *vs.* E_b/N_0 for tetrahedron (4-simplex) signal set.

(d) For $M = 8$ and $N = 2$, propose at least two good signal sets, and determine which one is better. [Open research problem: Find the optimal such signal set, and prove that it is optimal.]

Possible 8-point 2-dimensional signal sets include:

(i) 8-PSK. If the radius of each signal point is r, then the minimum distance is $d_{\min} = 2r \sin 22.5°$, so to achieve $d_{\min} = 2$ requires $r = 1/(\sin 22.5°) = 2.613$, or an energy of 6.828 (8.34 dB).

(ii) An 8-point version of the V.29 signal set, with four points of type $(1, 1)$ and four points of type $(3, 0)$. The average energy is then 5.5 (7.40 dB), about 1 dB better than 8-PSK. Even better, the minimum distance can be maintained at $d_{\min} = 2$ if the outer points are moved in to $(1 + \sqrt{3}, 0)$, which reduces the average energy to 4.732 (6.75 dB).

(iii) Hexagonal signal sets. One hexagonal 8-point set with $d_{\min} = 2$ has 1 point at the origin, 6 at squared radius 4, and 1 at squared radius 12, for an average energy of $36/8 = 4.5$ (6.53 dB). The mean $\mathbf{m}$ has length $\sqrt{12}/8$, so removing the mean reduces the energy further by $3/16 = 0.1875$ to 4.3125 (6.35 dB).

Another more symmetrical hexagonal signal set (the "double diamond") has points at $(\pm 1, 0), (0, \pm\sqrt{3})$ and $(\pm 2, \pm\sqrt{3})$. This signal set also has average energy $36/8 = 4.5$ (6.53 dB), and zero mean.

Problem 3.6 (Even-weight codes have better coding gain)

Let $\mathcal{C}$ be an (n, k, d) binary linear code with d odd. Show that if we append an overall parity check $p = \sum_i x_i$ to each codeword $\mathbf{x}$, then we obtain an $(n+1, k, d+1)$ binary linear code $\mathcal{C}'$ with d even. Show that the nominal coding gain $\gamma_c(\mathcal{C}')$ is always greater than $\gamma_c(\mathcal{C})$ if $k > 1$. Conclude that we can focus primarily on linear codes with d even.

The new code $\mathcal{C}'$ has the group property, because the mod-2 sum of two codewords $(x_1, \ldots, x_n, p = \sum_i x_i)$ and $(x'_1, \ldots, x'_n, p' = \sum_i x'_i)$ is

$$(x_1 + x'_1, \ldots, x_n + x'_n, p + p' = \sum_i x_i + x'_i),$$

another codeword in $\mathcal{C}'$. Its length is $n' = n+1$, and it has the same number of codewords (dimension). Since the parity bit p is equal to 1 for all odd-weight codewords in $\mathcal{C}$, the weight of all odd-weight codewords is increased by 1, so the minimum nonzero weight becomes $d' = d + 1$. We conclude that $\mathcal{C}'$ is a binary linear $(n+1, k, d+1)$ block code.

The nominal coding gain thus goes from $\frac{dk}{n}$ to $\frac{(d+1)k}{n+1}$. Since

$$\frac{d}{d+1} < \frac{n}{n+1}$$

if $d < n$, the nominal coding gain strictly increases unless $d = n$— *i.e.,* unless $\mathcal{C}$ is a repetition code with $k = 1$— in which case it stays the same (namely 1 (0 dB)).