Group codes outperform binary coset codes on non-binary symmetric memoryless channels

Giacomo Como*

October 6, 2008

Abstract

Typical minimum distances and error exponents are analyzed on the 8-PSK Gaussian channel for two capacity-achieving code ensembles with different algebraic structure. It is proved that the ensemble of group codes over $\mathbb{Z}_8$ achieves both the Gilbert-Varshamov bound and the expurgated error exponent with probability one. On the other hand, the ensemble of binary coset codes (under any labeling) is shown to be bounded away, with probability one, both from the Gilbert-Varshamov bound (at any rate) and the expurgated exponent (at low rates). The reason for this phenomenon is shown to rely on the symmetry structure of the 8-PSK constellation, which is known to match $\mathbb{Z}_8$, but not $\mathbb{Z}_2^3$.

The presented results indicate that designing group codes matching the symmetry of the channel guarantees better typical-code performance than designing codes whose algebraic structure does not match the channel. This stands in partial contrast with the well-known fact that the average binary coset code achieves both the capacity and the random-coding error exponent of any discrete memoryless channel.

Keywords: random codes, linear codes, group codes, coset codes, minimum distance, error exponent, Gilbert-Varshamov bound, expurgated exponent.

1 Introduction

As low-complexity modern coding schemes are based on random constructions of linear codes with sparse graphical representation [33], the analysis of random codes with algebraic structure has recently attracted renewed attention from the research community [2]. In fact, a precise evaluation of the performance of random linear codes (with no constraints on their density) is propaedeutic to the theory of low-density parity-check (LDPC) and turbo codes, since it allows to quantify the loss in performance due to the sparsity constraint.

On the other hand, it has long been known that random constructions of algebraically structured codes can outperform purely random code constructions. For instance, this is the case in some problems in multi-terminal information theory, where random linear codes allow to achieve larger capacity regions than purely random codes do (see [27], or the more recent work [32] and references therein). Restricting attention to point-to-point communication, which will be the framework of the present paper, random binary-linear codes are known to outperform purely random codes on binary-input symmetric-output memoryless channels in terms of typical minimum distances and error exponents (see [2]).

[*] Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. Email: giacomo@mit.edu

The present paper is concerned with the performance analysis of code ensembles with group or coset structure, when employed over non-binary discrete-input memoryless channels (DMCs). In this case, while structured code ensembles are expected to outperform purely random code constructions, it is not a priori clear which algebraic structure is the optimal one: indeed, many non-isomorphic groups typically exist of order equal to some non-prime number [26]. As it will be shown in this paper, it turns out that the choice of the algebraic structure is critical for the typical code performance of the ensemble. Rather than presenting a general theory, we shall focus on a specific case, the additive white Gaussian noise channel (AWGNC) with input restricted to the 8-Phase Shift Keying (8-PSK) signal constellation: our choice is motivated both by the applicative interest of this channel, and by the fact that it presents most of the key characteristics of the general case. While the arguments of [2] can be easily extended to show that the typical-code performance of the random coding ensemble (RCE) is suboptimal, we shall provide precise results for the ensemble of group codes over the cyclic group $\mathbb{Z}_8$ (GCE), and the ensemble of binary-coset codes (BCE), respectively (see Sect.2.3 for their formal definitions). These results will show that the typical group code has both better minimum distance and better error exponent than the typical binary-coset code.

The Gilbert-Varshamov (GV) bound [23, 37] is one of the most famous lower bounds on the achievable minimum Hamming distance of binary codes. Given a rate R in $(0,1)$, and defined $\gamma_2(R)$ as the unique solution in $(0,1/2)$ of the equation $\mathrm{H}(x) = 1 - R$ [1], it states that for every $n \geq 1$ there exist codes of block-length n, rate R, and minimum distance at least $n\gamma_2(R)$. Its asymptotic tightness is considered one of longest-standing unproved conjectures in coding theory. A closely related issue concerns the tightness of the expurgated exponent, which is conjectured by many to coincide with the reliability function of the DMC, i.e. the highest achievable error exponent (see [19, 30, 31, 5, 38]). Although both the classical GV bound and expurgated bound are mere existence results, for binary symmetric memoryless channels it is known that the binary-linear coding ensemble asymptotically achieves both the GV bound and the expurgated exponent, with probability one (see [18, 2]). It is also known that the same does not hold true [2] for the RCE, whose typical-code performance is bounded away from the GV bound, as well as (at low rates) from the expurgated error exponent.

Generalizations of the above issues to non-binary DMCs are considered in the present paper. Here, the GV distance and the expurgated bound are defined as solutions of simple finite-dimensional convex optimization problems, having the form of distortion-rate functions for the Bhattacharyya distance (see (12) and (22)). Analogously to the binary case, the RCE can be easily shown to be bounded away with probability one from both the GV distance and the expurgated error exponent of the 8-PSK AWGNC. The main results of the this paper show that, with probability one, the GCE achieves the GV bound (Theorem 1), while the BCE is bounded away from it (Theorem 2). Similarly, the GCE asymptotically achieves the expurgated exponent (Theorem 3), while the BCE does not (Theorem 4), with probability one.

As it will be clarified in the sequel, the reason for the outperformance of the GCE over the BCE resides in the symmetry structure of the 8-PSK AWGNC. Such a channel is symmetric with respect to the action of two groups of order 8, the cyclic group $\mathbb{Z}_8$ and the non-Abelian dihedral group D_4, none of which supports Galois field structure. In contrast, the additive group of the Galois field with 8 elements, which is isomorphic to $\mathbb{Z}_2^3$, does not match the 8-PSK in the sense of [29]. Thus, the results of the present paper suggest that random group codes matching the symmetry of the channel outperform random codes whose algebraic structure does not match that symmetry.

[1] Here $\mathrm{H}(x)$ denotes the binary entropy.

It is well known that, despite not matching the symmetry of the channel, the BCE achieves the capacity and the random-coding exponent of the 8-PSK AWGNC, likewise of any other DMC [19, pagg.206-209]. Recent works [25, 3, 4], analyzing the performance of binary-coset LDPC codes on non-binary input DMCs, find information-theoretical basis in the aforementioned fundamental results. In contrast, Theorem 2 and Theorem 4 imply that, when the symmetry of the channel is not matched, the BCE is suboptimal in terms the typical minimum distance and the typical error exponent.

On the other hand, group codes for symmetric channels have been widely investigated in the channel coding literature. They allow to use more spectrally efficient signal constellations, while inheriting many of the structural properties enjoyed by binary-linear codes: uniform error property, invariant distance profiles, congruent Voronoi regions, minimal encoders, syndrome formers and trellis representations. The reader is referred to [34, 16, 29, 7, 15, 17] and references therein. It is well known [14] that group codes over Abelian groups admitting Galois field structure (i.e. isomorphic to $\mathbb{Z}_p^r$ for some prime p) allow to achieve the capacity and the random coding exponent. More recently, information-theoretic limits of finite Abelian group codes were investigated in [8], where it was shown that group codes over $\mathbb{Z}_m$ allow to achieve capacity on the m-PSK AWGNC when m is the power of a prime (thus including the case $m = 8$). Theorem 1 and Theorem 3 show that, at least on the 8-PSK AWGNC, random group codes matching the symmetry of the channel are optimal in terms of typical-code performance. They provide theoretical foundation for the analysis and design of bandwidth-efficient high-performance coding schemes based on LDPC or turbo codes matched to geometrically uniform constellations [3, 35, 21, 9, 22]. It was empirically observed in [35] that LDPC codes over $\mathbb{Z}_8$ perform better than their binary counterparts on the 8-PSK AWGNC: the results of the present paper point out to an analytical explanation for this phenomenon.

Observe that, despite the cyclic group $\mathbb{Z}_8$ matches the 8-PSK constellation, in [8] the average error exponent of the GCE was shown to be strictly smaller than the random-coding error exponent at low rates (more in general this is the case for group code ensembles over finite Abelian groups not admitting Galois field structure, confirming an early conjecture of [14]). Since, as already mentioned, the average error exponent of the BCE coincides instead with the random-coding error exponent, it turns out that, at low rates, the BCE outperforms the GCE in terms of average error exponent, while the latter outperforms the former in terms of typical error exponent. While this phenomenon might appear paradoxical at a first glance, it can be easily explained by the fact that the average error exponent only provides a lower bound on the typical error exponent (by Markov's inequality). This estimation fails to be tight at rates not close to capacity, where the average error exponent is dragged down by an asymptotically vanishing fraction of codes with poor performance.

The remainder of the paper is organized as follows. In Sect.2, after introducing all the notation (Sect.2.1), analyzing the symmetry properties of the 8-PSK constellation (Sect.2.2), and formally introducing the GCE and the BCE (Sect.2.3), the main results of the paper are stated in Sect.2.4 and Sect.2.5. They are: Theorem 1, characterizing the typical asymptotic minimum distance of the GCE; Theorem 2, providing upper and lower bounds to the typical minimum distance of the BCE; Theorem 3, providing a lower bound to the typical error exponent of the GCE; Theorem 4 providing an upper bound to the typical error exponent of the BCE. In Sect.3 the most relevant part of Theorem 1, showing that the GCE achieves the GV bound, is proved by an application of the first-moment method followed by some considerations on the geometry of the 8-PSK constellation. Proving the tightness of this result requires a second-moment method and is technically more involved: for the sake of completeness, a proof is provided in Sect.B. Theorem 2 is proved in Sect.4 by applying the second-moment method (Sect.4.1) and some convex optimization techniques (Sect.4.2). The proof of Theorem 3is provided in Sect.5, while Theorem 4 is proved in Sect.6. Finally, Sect.7 presents some concluding remarks and points out to generalizations of the results to balanced DMCs. Sect.A is of a technical nature and discusses some continuity issues. Sect.C contains the proofs of some of the results of Sect.4.

Some of the material of this paper has been presented at ISIT 2007 [10].

2 Problem statement and main results

2.1 Notation

For two sets $A \subseteq B$, $\mathbb{1}_A : B \to \mathbb{R}$ will denote the indicator function of A in B, i.e. $\mathbb{1}_A(x) = 1$ for x in A and $\mathbb{1}_A(x) = 0$ for all $x \in B \setminus A$.

For a non-empty finite set A, the inner product of two functions $\boldsymbol{f}, \boldsymbol{g} : A \to \mathbb{R}$ will be denoted by $\langle \boldsymbol{f}, \boldsymbol{g} \rangle := \sum_{a \in A} \boldsymbol{f}(a)\boldsymbol{g}(a)$, while $\mathrm{supp}(\boldsymbol{f}) := \{a \in A : \boldsymbol{f}(a) \neq 0\}$ will denote the support of $\boldsymbol{f}$. We shall consider the set $\mathcal{P}(A)$ of probability measures over A, which can be identified with the simplex of functions $\boldsymbol{\theta} : A \to [0, +\infty)$ satisfying the linear constraint $\sum_{a \in A} \boldsymbol{\theta}(a) = 1$. In particular, for $a \in A$, $\delta_a \in \mathcal{P}(A)$ will denote the delta distribution concentrated in a, defined by $\delta_a(b) = 0$ for $b \neq a$, $\delta_a(a) = 1$. If $\boldsymbol{\theta}$ is in $\mathcal{P}(A)$ and $B \subseteq A$ is such that $\boldsymbol{\theta}(B) := \sum_{b \in B} \boldsymbol{\theta}(b) > 0$, the conditioned measure $\boldsymbol{\theta}|_B \in \mathcal{P}(B)$ is defined by

$$\boldsymbol{\theta}|_B(b) := \boldsymbol{\theta}(B)^{-1}\boldsymbol{\theta}(b) \,.$$

The entropy function $\mathrm{H} : \mathcal{P}(A) \to \mathbb{R}^+$ is defined as [2] [3]

$$\mathrm{H}(\boldsymbol{\theta}) := -\sum_{a \in \mathrm{supp}(\boldsymbol{\theta})} \boldsymbol{\theta}(a) \log \boldsymbol{\theta}(a) \,.$$

To any function $\pi : A \to B$ between two nonempty finite sets A and B we can associate a map $\pi_\sharp : \mathcal{P}(A) \to \mathcal{P}(B)$ sending the probability measure $\boldsymbol{\theta}$ in $\mathcal{P}(A)$ to its image measure $\pi_\sharp\boldsymbol{\theta} \in \mathcal{P}(B)$ defined by $[\pi_\sharp\boldsymbol{\theta}](b) := \boldsymbol{\theta}\left(f^{-1}(b)\right) = \sum_{a : f(a) = b} \boldsymbol{\theta}(a)$.The entropy of a measure $\boldsymbol{\theta}$ and that of its image measure $\pi_\sharp\boldsymbol{\theta}$ are related by the following equality

$$\mathrm{H}(\boldsymbol{\theta}) = \mathrm{H}(\pi_\sharp\boldsymbol{\theta}) + \sum_{b \in \mathrm{supp}(\pi_\sharp\boldsymbol{\theta})} \pi_\sharp\boldsymbol{\theta}(b)\, \mathrm{H}\left(\boldsymbol{\theta}|_{\pi^{-1}(b)}\right) \,. \tag{1}$$

A special case which will be considered in the paper is when $A = B \times B$, and $\pi^1, \pi^2 : A \to B$, are the projection operators, $\pi^j(b_1, b_2) := b_j$. In this case, the image measures $\pi^1_\sharp\boldsymbol{\theta}$ and $\pi^2_\sharp\boldsymbol{\theta}$ are simply the two marginals of the joint measure $\boldsymbol{\theta} \in \mathcal{P}(B \times B)$.

The type or empirical frequency of a string $\boldsymbol{a}$ in A^n is the probability measure $\boldsymbol{v}_A(\boldsymbol{a})$ in $\mathcal{P}(A)$ defined by $[\boldsymbol{v}_A(\boldsymbol{a})](a) := \frac{1}{n}\left|\{1 \leq i \leq n \,|\, \boldsymbol{a}(i) = a\}\right|$ for all a in A. For every positive integer n, $\mathcal{P}_n(A)$ will denote the set of the types of all length-n A-strings, i.e. $\mathcal{P}_n(A) := \boldsymbol{v}_A(A^n) \subseteq \mathcal{P}(A)$. It is immediate to check that the set of all A-types $\mathcal{P}_{\mathbb{N}}(A) := \cup_{n \geq 1}\mathcal{P}_n(A)$ is dense in $\mathcal{P}(A)$. Given a type $\boldsymbol{\theta}$ in $\mathcal{P}_n(A)$, the set of all length-n A-strings of type $\boldsymbol{\theta}$ will be denoted by $A^n_{\boldsymbol{\theta}} := \boldsymbol{v}_A^{-1}(\boldsymbol{\theta}) \cap A^n$. Its cardinality, equal to the multinomial $\binom{n}{n\boldsymbol{\theta}} := n!/\prod_{a \in A}(n\boldsymbol{\theta}(a))!$, grows exponentially fast in n with exponent given by $\mathrm{H}(\boldsymbol{\theta})$. More precisely,

[2] With an abuse of notation for any x in $[0, 1]$ we will sometimes denote by $\mathrm{H}(x)$ the entropy of the binary measure $\boldsymbol{\theta}$ in $\mathcal{P}(\{0, 1\})$ defined by $\boldsymbol{\theta}(1) = x$, $\boldsymbol{\theta}(0) = 1 - x$.

[3] Throughout the paper the base of log and exp is understood to be the same arbitrarily chosen $b > 1$.

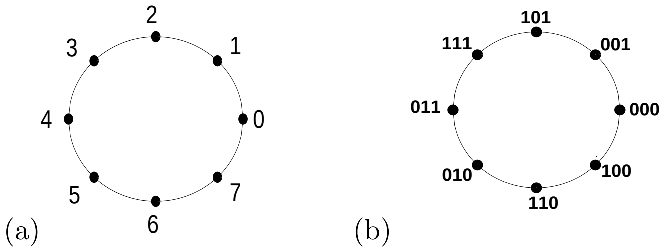

Figure 1: The 8-PSK constellation with: (a) the isometric labeling $\mu : \mathbb{Z}_8 \to \mathcal{X}$; (b) a binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$. The latter is a so-called Gray labeling: neighbor signals are assigned labels differing in one bit only.

given a sequence $(\boldsymbol{\theta}_n) \subseteq \mathcal{P}(A)$ such that $\lim_n \boldsymbol{\theta}_n = \boldsymbol{\theta}$ and $\boldsymbol{\theta}_n$ belongs to $\mathcal{P}_n(A)$ for all n, it holds

$$\left|A^n_{\boldsymbol{\theta}_n}\right| = \binom{n}{n\boldsymbol{\theta}_n} \le \exp(n\,\mathrm{H}(\boldsymbol{\theta}_n))\,, \qquad \lim_n \frac{1}{n} \log \binom{n}{n\boldsymbol{\theta}_n} = \mathrm{H}(\boldsymbol{\theta})\,. \tag{2}$$

On the other hand, the number of types $|\mathcal{P}_n(A)| = \binom{n+|A|-1}{|A|-1}$ grows polynomially fast with n. We refer to [12] for proofs of these facts.

2.2 The 8-PSK AWGNC and its symmetries

We shall consider transmission over a memoryless AWGNC with input constrained on the 8-PSK signal constellation $\mathcal{X} := \{e^{i\frac{2\pi}{8}k} | \, 0 \le k < 8\}$ and output space $\mathcal{Y} = \mathbb{R}^2$. For $x \in \mathcal{X}$, $P(\,\cdot\,|x) := \frac{1}{2\pi\sigma^2} e^{-||x-\,\cdot\,||^2/2\sigma^2}$ will denote the conditional probability density of the channel output given that the input x has been transmitted. The Bhattacharyya distance function associated to the 8-PSK AWGNC is [4]

$$\boldsymbol{D} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+\,, \qquad \boldsymbol{D}(x_1, x_2) := \frac{\log e}{8\sigma^2} ||x_1 - x_2||^2\,. \tag{3}$$

The *symmetry group* of $\mathcal{X}$, i.e. the automorphism group of its distance function $\boldsymbol{D}$ [5]

$$\mathrm{Aut}(\boldsymbol{D}) := \{\pi \in S_{\mathcal{X}} \mid \boldsymbol{D}\,(\pi(x_2), \pi(x_1)) = \boldsymbol{D}(x_1, x_2)\,,\ \forall\, x_1, x_2 \in \mathcal{X}\} \tag{4}$$

is isomorphic to the dihedral group D_8 with 16 elements [16, 29], generated by the rotation around the origin by an angle of $\frac{2\pi}{8}$ and the reflection through a straight line forming an angle of $\frac{2\pi}{16}$ with the real axis. [6] The constellation $\mathcal{X}$ is said to be geometrically uniform [16], meaning that for every $x_1, x_1 \in \mathcal{X}$ there exists $\pi \in \mathrm{Aut}(\boldsymbol{D})$ such that $\pi(x_1) = x_2$.

Moreover, the cyclic group $\mathbb{Z}_8$ is said to be a generating group of $\mathcal{X}$ [29], meaning that $\mathrm{Aut}(\boldsymbol{D})$ has a subgroup G isomorphic to $\mathbb{Z}_8$ such that for all $x_1, x_2 \in \mathcal{X}$ there exists a unique $\pi \in \mathrm{Aut}(\boldsymbol{D})$ such that $\pi(x_1) = x_2$. In particular, let

$$\mu : \mathbb{Z}_8 \to \mathcal{X}\,, \qquad \mu(z) := e^{i\frac{2\pi}{8}z}$$

be the standard isometric labeling, and consider the function $\boldsymbol{D}_\mu : \mathbb{Z}_8 \times \mathbb{Z}_8 \to \mathbb{R}$, $\boldsymbol{D}_\mu(z_1, z_2) = \boldsymbol{D}(\mu(z_1 + z_2), \mu(z_2))$. Then, all the columns $\boldsymbol{D}_\mu(\,\cdot\,, z)$ coincide with the distance profile

$$\boldsymbol{d} : \mathbb{Z}_8 \to \mathbb{R}\,, \qquad \boldsymbol{d}(z) := \boldsymbol{D}(\mu(0), \mu(z))\,. \tag{5}$$

[4] Here and throughout the paper for a vector $\boldsymbol{z}$ in $\mathbb{R}^d$, $||\boldsymbol{z}|| = \sqrt{\sum_{i=1}^d \boldsymbol{z}(i)^2}$ will denote its L_2-norm.

[5] $S_{\mathcal{X}}$ denotes the group of permutations of $\mathcal{X}$.

[6] The reader is referred to the textbook [26] for standard notions of algebra, and in particular of group actions on sets.

On the other hand, observe that D_8 has no subgroup isomorphic to $\mathbb{Z}_2^3$. [7] This implies that, for any binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$, not all the columns of the induced distance function

$$\boldsymbol{D}_\eta : \mathbb{Z}_2^3 \times \mathbb{Z}_2^3 \to \mathbb{R}^+ , \qquad \boldsymbol{D}_\eta(z_1, z_2) := \boldsymbol{D}(\eta(z_2), \eta(z_2 + z_1)) , \tag{6}$$

coincide, i.e.

$$\exists z_1, z_2 \in \mathbb{Z}_2^3 : \qquad \boldsymbol{D}_\eta(\cdot, z_1) \neq \boldsymbol{D}_\eta(\cdot, z_2) . \tag{7}$$

2.3 Two capacity-achieving code ensembles with different algebraic structure

We shall consider block-codes $\mathcal{C} \subseteq \mathcal{X}^n$ of rate $R(\mathcal{C}) := \frac{1}{n} \log |\mathcal{C}|$, whose *minimum distance* is defined by

$$d_{\min}(\mathcal{C}) = \min \left\{ \boldsymbol{D}_n(\boldsymbol{x}, \boldsymbol{z}) \,\middle|\, \boldsymbol{x} \neq \boldsymbol{z} \in \mathcal{C} \right\} ,$$

where, for $\boldsymbol{x}, \boldsymbol{z} \in \mathcal{X}^n$, $\boldsymbol{D}_n(\boldsymbol{x}, \boldsymbol{z}) := \sum_{i=1}^n \boldsymbol{D}(x_i, z_i)$. The *error probability* of $\mathcal{C}$ is given by

$$p_e(\mathcal{C}) := \frac{1}{|\mathcal{C}|} \sum_{\boldsymbol{x} \in \mathcal{C}} \int_{\Lambda_{\boldsymbol{x}}} \frac{1}{(2\pi\sigma^2)^n} e^{-\frac{||\boldsymbol{y}-\boldsymbol{x}||^2}{2\sigma^2}} \mathrm{d}\boldsymbol{y} ,$$

where $\Lambda_{\boldsymbol{x}} := \bigcup_{\boldsymbol{w} \neq \boldsymbol{x} \in \mathcal{C}} \{\boldsymbol{y} \in \mathcal{Y}^n \,:\, ||\boldsymbol{y} - \boldsymbol{x}|| \geq ||\boldsymbol{y} - \boldsymbol{w}||\}$ is the error event conditioned on the transmission of $\boldsymbol{x} \in \mathcal{C}$.

The focus of this paper will be on block-codes with algebraic structure compatible with $\mathbb{Z}_8$ or $\mathbb{Z}_2^3$, respectively. Specifically, a *group code* (over $\mathbb{Z}_8$) is the image of a subgroup K of the direct group product $\mathbb{Z}_8^n$ through the componentwise extension $\mu_n : \mathbb{Z}_8^n \to \mathcal{X}^n$ of the isometric labeling μ. As a consequence of the symmetry properties discussed in Sect.2.2, it is easy to check that the minimum distance of a group code $\mathcal{G} := \mu_n(K)$ coincides with its minimum weight, i.e.

$$\mathrm{d}_{\min}(\mathcal{G}) = \min\{ \sum_{1 \leq j \leq n} \boldsymbol{d}(x_j) | \boldsymbol{x} \neq \boldsymbol{0} \in K \} .$$

Similarly, group codes are known to enjoy the uniform error property, i.e.

$$p_e(\mathcal{G}|\boldsymbol{x}_1) = p_e(\mathcal{G}|\boldsymbol{x}_2) , \qquad \forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{G} .$$

A *binary coset code* is the image $\mathcal{B}$ of a coset J of the direct group product $\mathbb{Z}_2^{3n}$ through the componentwise extension $\eta_n : \mathbb{Z}_2^{3n} \to \mathcal{X}^n$ of an arbitrary binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$. As opposed to group codes, in general, neither binary coset codes enjoy the uniform error property, nor their minimum distance coincide with their minimum weight. In the sequel, we shall see as this reflects on the performance of random group and coset codes respectively.

For every design rate R in $[0, \log 8]$, set

$$\overline{R} := \log 8 - R , \qquad l := \left\lfloor \frac{\overline{R}}{\log 8} n \right\rfloor , \qquad n \in \mathbb{N} .$$

and define the two following code ensembles:

Group coding ensemble (GCE) For $n \geq 1$, let Φ_n^R be a random variable (r.v.) uniformly distributed over $\mathrm{Hom}\left(\mathbb{Z}_8^n, \mathbb{Z}_8^l\right)$, the set of all group homomorphisms from $\mathbb{Z}_8^n$ to $\mathbb{Z}_8^l$. The GCE of rate R is the sequence of random codes $\left(\mathcal{G}_n^R\right)$, where $\mathcal{G}_n^R$ is defined as the image through μ_n of the kernel of Φ_n^R, i.e.

$$\mathcal{G}_n^R := \mu_n \left(\ker \Phi_n^R\right) ;$$

[7] In fact, the only other generating group of $\mathcal{X}$ is the non-Abelian dihedral group D_4. Notice that group codes over non-Abelian groups are known to have poor minimum distance properties [28].

Binary coset ensemble (BCE) Let $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$ be an arbitrary labeling, and, for $n \geq 1$, consider a r.v. Ψ_n^R uniformly distributed over $\mathrm{Hom}\left(\mathbb{Z}_2^{3n}, \mathbb{Z}_2^{3l}\right)$, the set of $\mathbb{Z}_2$-linear maps from $\mathbb{Z}_2^{3n}$ to $\mathbb{Z}_2^{3l}$, and let $\boldsymbol{Z}_n$ be a r.v. uniformly distributed over $\mathbb{Z}_2^{3l}$, independent from Ψ_n^R. The BCE of rate R is the sequence of random codes $\left(\mathcal{B}_n^R\right)$, where $\mathcal{B}_n^R$ is defined as the image through η_n of the preimage of $\boldsymbol{Z}_n$ through Ψ_n^R, i.e.

$$\mathcal{B}_n^R := \eta_n \left(\left(\Psi_n^R\right)^{-1} \boldsymbol{Z}_n \right) . \tag{8}$$

Notice that both, for a given n, the probability that Ψ_n^R fails to be surjective is nonzero. In particular, it is possible that $\mathcal{B}_n^R$ is empty. [8] However, the following standard result shows that this event almost surely occurs only for finitely many values of n, and, therefore, does not affect the typical asymptotic performance of these ensembles.

Lemma 1 *For all* $0 < R < \log 8$*, with probability one there exists some* $n_0 \geq 0$ *such that* Ψ_n^R *is surjective for all* $n \geq n_0$*.*

Proof Let A_n be the event 'Ψ_n^R is surjective'. We can identify Ψ_n^R in the standard way with a random binary matrix H uniformly distributed over $\mathbb{Z}_2^{3n \times 3l}$, by defining $H_{ij} := (\Psi_n^R e_i)_j$, where $\{e_i\}_{1 \leq i \leq 3n}$ is the canonical bases of $\mathbb{Z}_2^{3n}$. Then Ψ_n^R is surjective iff H is full-rank. Since the rows of H are i.i.d. uniformly on $\mathbb{Z}_2^{3n}$, we have that, for $1 \leq j \leq 3l$, the probability that the j-th row of H is linear dependent on the other $(3l-1)$ rows is bounded from above by $2^{-3n} 2^{3l-1} \leq \exp(-nR)$. Then a standard union-bounding technique gives us that $\mathbb{P}(A_n) \leq n \exp(-nR)$, so that the series $\sum_{n \geq 1} \mathbb{P}(A_n)$ is convergent and the Borel-Cantelli lemma [6, pag.12] implies that, with probability one, A_n fails to occur at most for finitely many $n \in \mathbb{N}$. ■

As immediate consequence of the symmetry properties discussed in Sect.2.2 is that the optimal input distribution is the uniform one on $\mathcal{X}$, both for the 8-PSK AWGNC Shannon capacity C_8 and for its random coding error exponent [19] $E_8^r(R)$.

It is not hard to show BCE achieves capacity and

$$\mathbb{E}\left[p_e(\mathcal{B}_n^R)\right] \leq \exp(-nE_8^r(R)) , \qquad \lim_n -\frac{1}{n} \log \mathbb{E}\left[p_e(\mathcal{B}_n^R)\right] = E_8^r(R) . \tag{9}$$

In fact, the standard random coding averaging arguments of [19, pagg.206-207] as well as the tightness considerations of [20] apply, upon observing that the events $A_{\boldsymbol{z}} := \left\{\Psi_n^R \boldsymbol{z} = \boldsymbol{Z}_n\right\}$, for $\boldsymbol{z} \in \mathbb{Z}_2^{3n}$, have probability 8^{-l} each and are such that, $A_{\boldsymbol{z}_1}$, $A_{\boldsymbol{z}_2}$ and $A_{\boldsymbol{z}_3}$ are mutually independent for all three distinct n-tuples $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3 \in \mathbb{Z}_2^{3n}$.

For the GCE $\left(\mathcal{G}_n^R\right)$ the situation is different due to the presence of zero-divisors in $\mathbb{Z}_8$. In fact, the event $B_{\boldsymbol{x}} := \left\{\Phi_n^R \boldsymbol{x} = \boldsymbol{0}\right\}$, for $\boldsymbol{x} \in \mathbb{Z}_8^n$, does not have probability 8^{-l} whenever $\boldsymbol{x}$ lies in a proper subgroup of $\mathbb{Z}_8^n$. Nevertheless, it has been shown in [8] that the GCE achieves capacity, and its average error probability can be upper-bounded by a term exponentially decreasing in the block-length n

$$\mathbb{E}\left[p_e(\mathcal{G}_n^R)\right] \leq \exp(-nE_{\mathbb{Z}_8}^r(R)) . \tag{10}$$

The exponent appearing in the righthand side of the above inequality is given by

$$E_{\mathbb{Z}_8}^r(R) := \min\left\{E_8^r(R), E_4^r(\tfrac{2}{3}R), E_2^r(\tfrac{1}{3}R)\right\} \,,$$

[8] In fact, the latter problem could be avoided by considering a r.v. $\boldsymbol{V}_n$ uniformly distributed over $\mathbb{Z}_2^{3n}$ and independent of Ψ_n^R. Then, $\ker \Psi_n^R + \boldsymbol{V}_n$ and $\left(\Psi_n^R\right)^{-1} \boldsymbol{Z}_n$ can be shown to be identically distributed, conditioned on Ψ_n^R being surjective.

with $E_4^r(\frac{2}{3}R)$ and $E_2^r(\frac{1}{3}R)$ respectively denoting the random coding error exponents of the AWGNCs with input restricted over the 4-PSK and the 2-PSK constellation. As shown in [8], the bound (10) is necessarily tight for the average error probability both at rates close to C, where $E_{\mathbb{Z}_8}^r(R) = E_8^r(R)$, and at low rates, where instead $E_{\mathbb{Z}_8}^r(R) := E_2^r(\frac{1}{3}R) < E_8^r(R)$.

Thus, the average error exponent of the BCE coincides with the random coding exponent $E_8^r(R)$, while the average error exponent of the GCE is strictly smaller than it for low rates. In other words, even if algebraic constraints do not affect the capacity achievable by group codes over the 8-PSK AWGNC, they do lower the average error exponent achievable by the GCE. In fact, we argue that this claim can be somehow misleading. Indeed, it refers to the performance of the average code rather than to the performance of the typical code sampled from the two ensembles. The results stated in the two following subsections show that, instead, the typical code sampled from the GCE outperforms the typical code sampled from the BCE, thus reversing the hierarchies outlined by the average-code analysis.

2.4 Gilbert-Varshamov bound and typical minimum distances

Let $\Omega := \mathcal{P}(\mathbb{Z}_8)$ be the space of $\mathbb{Z}_8$-types, and, for $0 \leq R \leq \log 8$, define

$$\Omega_{(R)} := \left\{ \boldsymbol{\omega} \in \Omega : \mathrm{H}(\boldsymbol{\omega}) \geq \overline{R} \right\} \tag{11}$$

$$\gamma_8(R) := \min \left\{ \langle \boldsymbol{\omega}, \boldsymbol{d} \rangle \middle| \, \boldsymbol{\omega} \in \Omega_{(R)} \right\} , \tag{12}$$

where $\boldsymbol{d}$ is the squared Euclidean weight function defined in (5). In Sect.A $\gamma_8(R)$ is proved to be continuous and non-increasing as a function of the rate R. The GV bound for the 8-PSK AWGNC states that for every design rate $0 \leq R \leq \log 8$, and any $n \geq 1$,

$$\exists \mathcal{C}_n \subseteq \mathcal{X}^n : \qquad R(\mathcal{C}_n) \geq R \, , \qquad \mathrm{d}_{\min}(\mathcal{C}_n) \geq n\gamma_8(R) \, . \tag{13}$$

While (13) above is an existence result, the question we want to address is whether $\gamma_8(R)$ is achieved by random codes. In fact, using arguments analogous to those in [2], it is not difficult to see that the RCE does not achieve the GV bound with probability one: its typical asymptotic normalized minimum distance can be shown to coincide with $\gamma_8(2R)$. We shall therefore concentrate on the performance of the GCE and BCE introduced in Sect.2.3. The following result states that a typical code sequence sampled from the GCE asymptotically meets the GV-bound.

Theorem 1 (Typical minimum distance of the GCE) *For all* $0 < R < \log 8$, *with probability one*

$$\lim_n \frac{1}{n} \mathrm{d}_{\min}(\mathcal{G}_n^R) = \gamma_8(R) \, . \tag{14}$$

Proof See Sect.3 and Sect.B. ■

For the BCE instead, we will prove that a typical code sequence almost surely does not meet the GV-bound. More precisely, let $\Theta := \mathcal{P}(\mathbb{Z}_2^3 \times \mathbb{Z}_2^3)$ be the set of joint $\mathbb{Z}_2^3$-types. For $0 \leq R \leq \log 8$, define the sets

$$\underline{\Theta}_{(R)} := \left\{ \boldsymbol{\theta} \in \Theta \, : \, \mathrm{H}(\boldsymbol{\theta}) \geq 2\overline{R} \, , \, \mathrm{H}\left(\pi_\sharp^1 \boldsymbol{\theta}\right) \geq \overline{R} \right\} , \tag{15}$$

$$\overline{\Theta}_{(R)} := \left\{ \boldsymbol{\theta} \in \Theta \, : \, \mathrm{H}(\boldsymbol{\theta}) - \mathrm{H}\left(\pi_\sharp^1 \boldsymbol{\theta}\right) \geq \overline{R} \, , \, \mathrm{H}\left(\pi_\sharp^1 \boldsymbol{\theta}\right) \geq \overline{R} \right\} , \tag{16}$$

where we recall that $[\pi_\sharp^1 \boldsymbol{\theta}](\,\cdot\,) = \sum_z \boldsymbol{\theta}(\,\cdot\,, z)$ is the marginal of $\boldsymbol{\theta}$. Define the functions

$$\underline{\gamma}_\eta(R) = \min \left\{ \langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle | \, \boldsymbol{\theta} \in \underline{\Theta}_{(R)} \right\} , \tag{17}$$

$$\overline{\gamma}_\eta(R) := \min \left\{ \langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle | \, \boldsymbol{\theta} \in \overline{\Theta}_{(R)} \right\} , \tag{18}$$

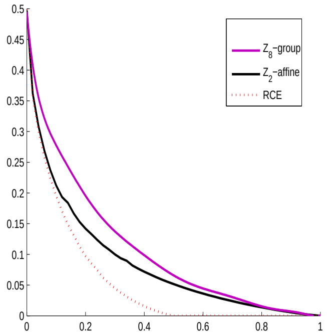

Figure 2: A comparison of $\gamma_8(R)$ (purple line) and $\overline{\gamma}_\eta(R)$ where $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$ is the Gray labeling described in Fig.1. As a reference $\gamma_8(2R)$ (which is the typical normalized minimum distance of the RCE) is plotted in dotted red.

Theorem 2 (Typical minimum distance of the BCE) *For every design rate* $0 < R < \log 8$*, with probability one*

$$\liminf_n \frac{1}{n} d_{\min}\left(\mathcal{B}_n^R\right) \geq \underline{\gamma}_\eta(R)\,. \tag{19}$$

$$\limsup_n \frac{1}{n} d_{\min}\left(\mathcal{B}_n^R\right) \leq \overline{\gamma}_\eta(R)\,. \tag{20}$$

Moreover, for all $0 < R < \log 8$

$$\underline{\gamma}_\eta(R) \leq \overline{\gamma}_\eta(R) < \gamma_8(R)\,. \tag{21}$$

Proof See Sect 4. ■

As an immediate consequence of (20) and (21) we have that, with probability one, the BCE is asymptotically bounded away from the GV-bound.

In Fig.2 the typical minimum distance of the GCE and of the BCE are plotted as a function of the rate, together with that of the RCE. For the specific choice of the binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$, and the chosen resolution, it seems that $\overline{\gamma}_\eta(R) = \underline{\gamma}_\eta(R)$.

2.5 Expurgated bound and typical error exponents

For every rate $0 \leq R \leq \log 8$ the *expurgated exponent* of the 8-PSK AWGNC is

$$E_8^x(R) := \min\left\{\langle \boldsymbol{\omega}, \boldsymbol{d}\rangle + \overline{R} - \mathrm{H}(\boldsymbol{\omega}) | \boldsymbol{\omega} \in \Omega_{(R)}\right\}\,. \tag{22}$$

The expurgated exponent $E_8^x(R)$ and the GV distance $\gamma_8(R)$ coincide at small rates. Indeed, let $\boldsymbol{\omega}^x := e^{-\boldsymbol{d}} / \sum_z e^{-\boldsymbol{d}(z)}$ be the minimizer of $\langle \boldsymbol{\omega}, \boldsymbol{d}\rangle + \overline{R} - \mathrm{H}(\boldsymbol{\omega})$ over the whole type space Ω, $R_8^x := \mathrm{H}(\boldsymbol{\omega}^x) > 0$ be the minimum rate R for which $\boldsymbol{\omega}^x \in \Omega_{(R)}$, and $R_8^0 := \log \sum_z \frac{1}{8} e^{-\boldsymbol{d}(z)}$ denote the so-called cut-off rate. We have that:

- for rates $R_8^x \leq R \leq R_8^0$, the minimum in (22) is achieved by $\boldsymbol{\omega}^x$, and $E_8^x(R) = R_8^0 - R$;
- for rates $0 \leq R \leq R_8^x$, Lemma 11 implies that the minimum in (22) is achieved by some type $\boldsymbol{\omega}$ such that $\mathrm{H}(\boldsymbol{\omega}) = \overline{R}$, so that

$$E_8^x(R) = \gamma_8(R)\,, \qquad \forall 0 \leq R \leq R_8^x\,. \tag{23}$$

The expurgated bound (see [19, pagg.153-157], and [11, pagg.185-186,192-195]) guarantees, for all rates $0 < R < \log 8$, the existence of code sequences $(\mathcal{C}_n \subseteq \mathcal{X}^n)$ with

$$R(\mathcal{C}_n) \geq R\,, \qquad \liminf_n -\frac{1}{n}\log p_e(\mathcal{C}_n) \geq E_8^x(R)\,.$$

Similarly to the GV bound, the expurgated bound is a mere existence result, while we are interested in whether the expurgated exponent $E_8^x(R)$ is achieved by random codes. In fact, arguments as in the binary case [2] show that the expurgated exponent is not achieved, at low rates, by the RCE. Therefore, we shall be concerned with the typical error exponents of the GCE and the BCE, respectively. The following result states that with probability one the GCE asymptotically achieves the expurgated exponent.

Theorem 3 (Typical error exponent of the GCE) *For every rate* $0 < R < R_8^x$, *with probability one*

$$\liminf_n -\frac{1}{n}\log p_e(\mathcal{C}_n) \geq E_8^x(R)\,. \tag{24}$$

Proof See Sect.5. ■

In contrast, the following result shows that, at sufficiently low rates, the BCE does not achieve the expurgated exponent.

Theorem 4 (Typical error exponent of the BCE) *There exists some* $R_\eta^x > 0$ *such that, for every rate* $0 < R < R_\eta^x$, *with probability one*

$$\limsup_n -\frac{1}{n}\log p_e(\mathcal{B}_n^R) < E_8^x(R)\,.$$

Proof See Sect.6. ■

3 The typical group code achieves the Gilbert-Varshamov bound

In this section we shall show that the sequence $\left(\frac{1}{n}\,\mathrm{d}_{\min}\left(\mathcal{G}_n^R\right)\right)$ is asymptotically bounded from below by $\gamma_8(R)$ with probability one. The tightness of this result will instead be proven in Sect.B, completing the proof of Theorem 1. Throughout, the notation $\Omega = \mathcal{P}(\mathbb{Z}_8)$, and $\Omega_n := \mathcal{P}_n(\mathbb{Z}_8)$ will be adopted for the sets of $\mathbb{Z}_8$-types.

We shall apply the first-moment method [1] to the type-enumerator function

$$G_n^R : \Omega \to \mathbb{Z}_+\,, \qquad G_n^R(\boldsymbol{\omega}) := \left|(\mathbb{Z}_8)^n_{\boldsymbol{\omega}} \cap \ker \Phi_n^R\right|\,.$$

It is not difficult to express the minimum distance of the GCE in terms of its type-enumerating function. Indeed, we have

$$\begin{aligned}\mathrm{d}_{\min}\left(\mathcal{G}_n^R\right) &= \min\left\{\textstyle\sum_{1\leq j\leq n} \boldsymbol{d}(x_j) \mid \boldsymbol{x} \in \ker \Phi_n^R \setminus \{\boldsymbol{0}\}\right\}\\ &= n \min\left\{\langle \boldsymbol{\omega}, \boldsymbol{d}\rangle \;\middle|\; \boldsymbol{\omega} \in \Omega \setminus \{\delta_0\} \;:\; G_n^R(\boldsymbol{\omega}) \geq 1\right\}\,.\end{aligned}$$

As a first step in our analysis, we evaluate the expected type-enumarating function $\mathbb{E}[G_n^R(\boldsymbol{\omega})]$. For $\boldsymbol{\omega} \in \Omega$, it is convenient to denote by [9]

$$\zeta(\boldsymbol{\omega}) := \frac{8}{\gcd(\mathrm{supp}(\boldsymbol{\omega}))}\,, \tag{25}$$

[9] Here gcd denotes the great common divisor of a set of positive integers.

the order of the smallest subgroup of $\mathbb{Z}_8$ supporting $\boldsymbol{\omega}$. Observe that the map $\zeta : \Omega \to \mathbb{N}$ takes values only on the set of divisors of 8. Moreover it is lower semicontinuous as it jumps to lower values when approaching $\mathbb{Z}_8$-types supported on smaller subgroups of $\mathbb{Z}_8$. The following result motivates definition (25).

Lemma 2 *For every design rate* $0 < R < \log 8$ *and* $\mathbb{Z}_8$*-type* $\boldsymbol{\omega} \neq \delta_0$ *in* $\mathcal{P}_n(\mathbb{Z}_8)$, *we have*

$$\mathbb{E}\left[G_n^R(\boldsymbol{\omega})\right] = \binom{n}{n\boldsymbol{\omega}} \frac{1}{\zeta(\boldsymbol{\omega})^l}\,.$$

Proof Let $\{e_i\}_{1\leq i\leq n}$ be the canonical basis of $\mathbb{Z}_8^n$, and define $h := \frac{8}{\zeta(\boldsymbol{\omega})}$. Writing $\boldsymbol{x} = \sum_{1\leq i\leq n} x_i e_i$, we have that x_i belongs to $h\mathbb{Z}_8$ for every $1 \leq i \leq n$, and there exists some $1 \leq i^* \leq n$ such that $\gcd(x_{i^*}, 8) = h$. Since $\left\{W_i := \Phi_N^R e_i\right\}_{1\leq i\leq n}$ is a collection of i.i.d. r.v.s uniformly distributed over $\mathbb{Z}_8^l$, it follows that $K_{i^*} := x_{i^*} W_{i^*}$ is uniformly distributed over $h\mathbb{Z}_8^l$, and is independent from the r.v. $K_{-i^*} := \sum_{i\neq i^*} x_i W_i$, which in turn takes values in $h\mathbb{Z}_8^l$. Then, for every $\boldsymbol{z}$ in $h\mathbb{Z}_8^l$ we have

$$\begin{aligned}
\mathbb{P}(\Phi_n^R \boldsymbol{x} = \boldsymbol{z}) &= \sum_{\boldsymbol{w}\in h\mathbb{Z}_8^l} \mathbb{P}\left(K_{i^*} = \boldsymbol{z} - \boldsymbol{w}\,, K_{-i^*} = \boldsymbol{w}\right)\\
&= \sum_{\boldsymbol{w}\in h\mathbb{Z}_8^l} \tfrac{1}{\zeta(\boldsymbol{\omega})^l} \mathbb{P}\left(K_{-i^*} = \boldsymbol{w}\right)\\
&= \zeta(\boldsymbol{\omega})^{-l}\,,
\end{aligned}$$

which shows that $\Phi_n^R \boldsymbol{x}$ is uniformly distributed over $\frac{8}{\zeta(\boldsymbol{\omega})}\mathbb{Z}_8^l$. Then, from the linearity of the expectation, we have

$$\mathbb{E}\left[G_n^R(\boldsymbol{\omega})\right] = \mathbb{E}[\sum_{\boldsymbol{x}\in(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}} \mathbb{1}_{\{\Phi_n^R \boldsymbol{x} = \boldsymbol{0}\}}] = \sum_{\boldsymbol{x}\in(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}} \mathbb{P}\left(\Phi_n^R \boldsymbol{x} = \boldsymbol{0}\right)\,,$$

and the claim easily follows from (2). ■

For $0 \leq R \leq \log 8$, define the sets

$$\begin{aligned}
\Omega''_{(R)} &:= \left\{\boldsymbol{\omega} \in \Omega : \mathrm{supp}(\boldsymbol{\omega}) \subseteq 2\mathbb{Z}_8,\, \mathrm{H}(\boldsymbol{\omega}) \geq \tfrac{2}{3}\overline{R}\right\}\,,\\
\Omega'_{(R)} &:= \left\{\boldsymbol{\omega} \in \Omega : \mathrm{supp}(\boldsymbol{\omega}) \subseteq 4\mathbb{Z}_8,\, \mathrm{H}(\boldsymbol{\omega}) \geq \tfrac{1}{3}\overline{R}\right\}\,.
\end{aligned} \tag{26}$$

and let

$$\begin{aligned}
\gamma_4\left(\tfrac{2}{3}R\right) &:= \min\{\langle\boldsymbol{\omega}, \boldsymbol{d}\rangle \,|\, \boldsymbol{\omega} \in \Omega''_{(R)}\}\,,\\
\gamma_2\left(\tfrac{1}{3}R\right) &:= \min\{\langle\boldsymbol{\omega}, \boldsymbol{d}\rangle \,|\, \boldsymbol{\omega} \in \Omega'_{(R)}\}
\end{aligned}$$

be the GV-distances associated to the subconstellations 4-PSK and 2-PSK, respectively.

The following almost sure lower bound on the asymptotic normalized minimum distance of the GCE is proved by applying a first-moment method and Lemma 2.

Proposition 1 *For every design rate* $0 < R < \log 8$, *and every* $0 < \varepsilon < \overline{R}$

$$\liminf_n \frac{1}{n} d_{\min}(\mathcal{G}_n^R) \geq \min\left\{\gamma_8(R+\varepsilon), \gamma_4(\tfrac{2}{3}(R+\varepsilon)), \gamma_2(\tfrac{1}{3}(R+\varepsilon))\right\} \tag{27}$$

with probability one.

Proof As a consequence of Lemma 2 and (2) we have that, for every $\boldsymbol{\omega} \in \Omega$,

$$\mathbb{E}\left[G_n^R(\boldsymbol{\omega})\right] \leq \exp\left(n\left(\mathrm{H}(\boldsymbol{\omega}) - \tfrac{\log \zeta(\boldsymbol{\omega})}{\log 8}\overline{R}\right)\right) . \tag{28}$$

Define the set $\tilde{\Omega} := \left\{\boldsymbol{\omega} \in \Omega : \mathrm{H}(\boldsymbol{\omega}) < \frac{\log \zeta(\boldsymbol{\omega})}{\log 8}(\overline{R} - \varepsilon)\right\}$, and for $n \geq 1$, $\tilde{\Omega}_n := \Omega_n \cap \tilde{\Omega}$. Consider the events $F_{n,\varepsilon} := \cup_{\boldsymbol{\omega}\in\tilde{\Omega}_n}\left\{G_n^R(\boldsymbol{\omega}) \geq 1\right\}$, and $F_\varepsilon := \{F_{n,\varepsilon} \text{ i. o.}\}$. [10] Then, using a standard union bound, Markov's inequality and (28), we have

$$\begin{aligned} \mathbb{P}\left(F_{n,\varepsilon}\right) &\leq \textstyle\sum_{\boldsymbol{\omega}\in\tilde{\Omega}_n} \mathbb{E}\left[G_n^R(\boldsymbol{\omega})\right] \\ &\leq \textstyle\sum_{\boldsymbol{\omega}\in\tilde{\Omega}_n} \exp\left(n\left(\mathrm{H}(\boldsymbol{\omega}) - \tfrac{\log \zeta(\boldsymbol{\omega})}{\log 8}\overline{R}\right)\right) \\ &\leq |\Omega_n| \exp(-n\varepsilon) . \end{aligned}$$

Since the number of $\mathbb{Z}_8$-types $|\Omega_n| = \binom{n+7}{7}$ grows only polynomially fast with n, we have that the series $\sum_{n\geq 1}\mathbb{P}\left(F_{n,\varepsilon}\right)$ is convergent. Hence, the Borel-Cantelli lemma implies that $\mathbb{P}(F_\varepsilon) = 0$. Then, the claim follows upon observing that $\Omega \setminus \tilde{\Omega} \subseteq \Omega_{(R+\varepsilon)} \cup \Omega''_{(R+\varepsilon)} \cup \Omega'_{(R+\varepsilon)}$. ■

Observe that the proofs of Lemma 2 and Proposition 1 only depend on the algebraic structure of $\mathbb{Z}_8$. Instead, the following result relies on the geometric structure of $\mathcal{X}$. In fact, counterexamples can be constructed as in [8] showing that Lemma 3 fails to hold true for other DMCs with the same symmetry structure of the 8-PSK AWGNC.

Lemma 3 *For every design rate* $0 \leq R \leq \log 8$

$$\gamma_2(\tfrac{1}{3}R) = \gamma_4(\tfrac{2}{3}R) \geq \gamma_8(R) , \tag{29}$$

Proof For $R = \log 8$, trivially $\gamma_2(\frac{1}{3}R) = \gamma_4(\frac{2}{3}R) = \gamma_8(R) = 0$.

For $R < \log 8$, since the entropy function is concave and the unique minimum of the linear map $\boldsymbol{\omega} \mapsto \langle\boldsymbol{\omega}, \boldsymbol{d}\rangle$ on Ω is achieved in $\boldsymbol{\omega} = \delta_0$, we can apply Lemma 11 and claim that a minimizer $\boldsymbol{\omega} \in \Omega_{(R)}$ in the definition (12) of $\gamma_8(R)$ necessarily satisfies $\mathrm{H}(\boldsymbol{\omega}) = \overline{R}$. Then, using Lagrangian multipliers, we obtain

$$\gamma_8(R) = \frac{\sum_{x\in\mathbb{Z}_8} \boldsymbol{d}(x) e^{-\lambda_8 \boldsymbol{d}(x)}}{\sum_{x\in\mathbb{Z}_8} e^{-\lambda_8 \boldsymbol{d}(x)}} ,$$

where $\lambda_8 > 0$ solves the equation $\mathrm{H}\left(\frac{e^{-\lambda_8 \boldsymbol{d}}}{\sum_x e^{-\lambda_8 \boldsymbol{d}(x)}}\right) = \overline{R}$. Similarly, we have $\gamma_2\left(\frac{1}{3}R\right) = \boldsymbol{d}(4)\alpha$, where

$$\alpha := \frac{1}{Z_2(\lambda_2)} e^{-\lambda_2 \boldsymbol{d}(4)} , \qquad Z_2(\lambda_2) := 1 + e^{-\lambda_2 \boldsymbol{d}(4)} ,$$

and $\lambda_2 > 0$ solves $\mathrm{H}\left(Z_2(\lambda_2)^{-1} e^{-\lambda_2 \boldsymbol{d}(4)}\right) = \frac{1}{3}\overline{R}$, while

$$\gamma_4\left(\tfrac{2}{3}R\right) = \frac{\sum_{x\in 2\mathbb{Z}_8} e^{-\lambda_4 \boldsymbol{d}(x)} \boldsymbol{d}(x)}{Z_4(\lambda_4)} .$$

where $Z_4(\lambda_4) := \sum_{x\in 2\mathbb{Z}_8} e^{-\lambda \boldsymbol{d}(x)}$, and $\lambda_4 > 0$ is the solution of $\mathrm{H}\left(Z_4(\lambda_4)^{-1} e^{-\lambda_4 \boldsymbol{d}} \mathbb{1}_{2\mathbb{Z}_8}\right) = \frac{2}{3}\overline{R}$.

Elementary geometrical considerations based on Pythagoras' theorems allow to show that

$$\boldsymbol{d}(4) = 2\boldsymbol{d}(2) = 2\boldsymbol{d}(6) \tag{30}$$

[10] For a sequence of events (A_n), the event $\{A_n \text{ i. o.}\} := \cap_{k\geq 0} \cup_{n\geq k} A_k$ denotes the event 'A_n occurs infinitely often'.

$$\boldsymbol{d}(1)=\boldsymbol{d}(7), \quad \boldsymbol{d}(3)=\boldsymbol{d}(5), \quad \boldsymbol{d}(1)=\boldsymbol{d}(4)-\boldsymbol{d}(3)<\frac{1}{4}\boldsymbol{d}(4). \tag{31}$$

It follows from (30) that $Z_4(2s)=\left(1+e^{-s\boldsymbol{d}(4)}\right)^2=Z_2\left(s\right)^2$, for all $s\geq 0$.Then, it follows from (30) that

$$\frac{e^{-2\lambda_2\boldsymbol{d}(0)}}{Z_4(2\lambda_2)}=\frac{1}{Z_2(\lambda_2)^2}=(1-\alpha)^2\,, \qquad \frac{e^{-2\lambda_2\boldsymbol{d}(2)}}{Z_4(2\lambda_2)}=\frac{e^{-2\lambda_2\boldsymbol{d}(6)}}{Z_4(2\lambda_2)}=\alpha(1-\alpha)\,, \qquad \frac{e^{-2\lambda_2\boldsymbol{d}(4)}}{Z_4(2\lambda_2)}=\alpha^2\,.$$

Therefore, $\mathrm{H}\left(\frac{e^{-2\lambda_2\boldsymbol{d}|_{2\mathbb{Z}_8}}}{Z_4(2\lambda_2)}\right)=2\,\mathrm{H}\left(\alpha\right)=2\,\mathrm{H}\left(\frac{e^{-\lambda_2\boldsymbol{d}|_{4\mathbb{Z}_8}}}{Z_2(\lambda_2)}\right)$, so that $2\lambda_2=\lambda_4$. Hence,

$$\gamma_4\left(\tfrac{2}{3}R\right)=\frac{\left\langle e^{-\lambda_4\boldsymbol{d}}\mathbb{1}_{2\mathbb{Z}_8},\boldsymbol{d}\right\rangle}{Z_4(\lambda_4)}=\alpha^2\boldsymbol{d}(4)+2\alpha(1-\alpha)\boldsymbol{d}(2)=\alpha\boldsymbol{d}(4)=\frac{\boldsymbol{d}(4)e^{-\frac{\lambda_4}{2}\boldsymbol{d}(4)}}{Z_2(\lambda_4/2)}=\gamma_2\left(\tfrac{1}{3}R\right)\,,$$

thus showing the equality in (29). It remains to show the inequality in (29). In order to do that, we introduce the $\mathbb{Z}_8$-type $\hat{\boldsymbol{\omega}}$ defined by

$$\begin{array}{c} \hat{\boldsymbol{\omega}}(0):=(1-\alpha)^3\,, \qquad \hat{\boldsymbol{\omega}}(1):=\hat{\boldsymbol{\omega}}(2):=\hat{\boldsymbol{\omega}}(7):=\alpha(1-\alpha)^2\,, \\ \hat{\boldsymbol{\omega}}(4):=\alpha^3\,, \qquad \hat{\boldsymbol{\omega}}(6):=\hat{\boldsymbol{\omega}}(5):=\hat{\boldsymbol{\omega}}(3):=\alpha^2(1-\alpha)\,, \end{array} \tag{32}$$

It is straightforward to verify that $\mathrm{H}(\hat{\boldsymbol{\omega}})=3\,\mathrm{H}\left(\alpha\right)=\overline{R}$. Moreover, it follows from (30) and (31) that

$$\begin{array}{rcl} \langle\hat{\boldsymbol{\omega}},\boldsymbol{d}\rangle &=& \sum_{x\in\mathbb{Z}_8}\boldsymbol{d}(x)\hat{\boldsymbol{\omega}}(x) \\ &=& \alpha^3\boldsymbol{d}(4)+2\alpha^2(1-\alpha)\left(\boldsymbol{d}(4)-\boldsymbol{d}(1)\right)+\alpha(1-\alpha)\frac{1}{2}\boldsymbol{d}(4)+2\alpha(1-\alpha)^2\boldsymbol{d}(1) \\ &=& \alpha\boldsymbol{d}(4)\frac{1}{2}\left(-2\alpha^2+3\alpha+1\right)+\alpha\boldsymbol{d}(1)2\left(1+2\alpha^2-3\alpha\right) \\ &=& \alpha\boldsymbol{d}(4)+\alpha\boldsymbol{d}(4)\left(2\boldsymbol{d}(1)-\frac{1}{2}\boldsymbol{d}(4)\right)\left(2\alpha^2-3\alpha+1\right) \\ &\leq& \alpha\boldsymbol{d}(4)\,, \end{array}$$

last inequality following from (31) and the fact that $2\alpha^2-3\alpha+1>0$ for every $\alpha>0$. It follows that

$$\gamma_8(R)\leq\langle\hat{\boldsymbol{\omega}},\boldsymbol{d}\rangle\leq\alpha\boldsymbol{d}(4)=\gamma_2\left(\tfrac{1}{3}R\right)\,,$$

thus concluding the proof. ■

Remark: From the proof of Lemma 3 it is evident that the rightmost inequality in (29) is strict for all $0<R<\log 8$.

As immediate consequence of Proposition 1 and Lemma 3, we have that, for every rate $0<R<\log 8$ and every $0<\varepsilon<\overline{R}$,

$$\mathbb{P}\left(\liminf_n\frac{1}{n}\,\mathrm{d}_{\min}\left(\mathcal{G}_n^R\right)\geq\gamma_8(R+\varepsilon)\right)=1\,.$$

Then, it follows from the monotonicity and continuity properties of $\gamma_8(R)$ that

$$\begin{array}{rcl} \mathbb{P}\left(\liminf_n\frac{1}{n}\,\mathrm{d}_{\min}\left(\mathcal{G}_n^R\right)\geq\gamma_8(R)\right) &=& \mathbb{P}\left(\bigcup_{k\geq 1}\left\{\liminf_n\frac{1}{n}\,\mathrm{d}_{\min}\left(\mathcal{G}_n^R\right)\geq\gamma_8(R+\frac{1}{k})\right\}\right) \\ &=& \lim_k\mathbb{P}\left(\liminf_n\frac{1}{n}\,\mathrm{d}_{\min}\left(\mathcal{G}_n^R\right)\geq\gamma_8(R+\frac{1}{k})\right) \\ &=& 1\,, \end{array} \tag{33}$$

i.e. almost surely the asymptotic normalized minimum distance of the GCE is not smaller than the GV-distance $\gamma_8(R)$. In order to complete the proof of Theorem 1, it remains to prove that almost surely $\frac{1}{n}\,\mathrm{d}_{\min}(\mathcal{G}_n^R)$ is asymptotically upper-bounded by $\gamma_8(R)$. This will be the object of Sect.B.

4 The typical binary-coset code does not achieve the GV bound

In the present section, we shall prove that the normalized minimum distance of the BCE is asymptotically bounded away from the GV distance. We shall proceed in two steps. First, in Sect.4.1, we shall use a second-moment method [1] in order to prove the upper bound (20). Then, in Sect.4.2, we shall prove the rightmost inequality in (21): this will involve some convex optimization arguments. Throughout, we shall assume to have fixed an arbitrary labeling

$$\eta : \mathbb{Z}_2^3 \to \mathcal{X} \,,$$

and use the notation $\Theta = \mathcal{P}(\mathbb{Z}_2^3 \times \mathbb{Z}_2^3)$ and $\Theta_n := \mathcal{P}_n(\mathbb{Z}_2^3 \times \mathbb{Z}_2^3)$ for the spaces of joint $\mathbb{Z}_2^3$-types, and $\Upsilon := \mathcal{P}(\mathbb{Z}_2^3)$ and $\Upsilon_n := \mathcal{P}_n(\mathbb{Z}_2^3)$ for the spaces of $\mathbb{Z}_2^3$-types.

4.1 Bounds on the typical minimum distance of the BCE

A first observation is that, since binary-coset codes are not GU, their minimum distance does not in general coincide with their minimum weight, as it is the case for $\mathbb{Z}_8$-group codes. Rather, it is necessary to look at all pairs of codewords of a binary coset code in order to evaluate its minimum distance. It is therefore convenient to introduce the joint-type-enumerating functions $U_n^R : \Theta \to \mathbb{Z}_+$,

$$U_n^R(\boldsymbol{\theta}) := \left|\left\{(\boldsymbol{x}, \boldsymbol{y}) \in (\mathbb{Z}_2^3)^n_{\boldsymbol{\theta}} : \Psi_n^R \boldsymbol{x} = \mathbf{0}, \Psi_n^R \boldsymbol{y} = \boldsymbol{Z}_n\right\}\right| \,,$$

counting the number of pairs $(\boldsymbol{x}, \boldsymbol{y})$ of different joint types such that both $\boldsymbol{y}$ and $\boldsymbol{x}+\boldsymbol{y}$ belong to coset of $\mathbb{Z}_2^{3n}$ given by the counter-image of $\boldsymbol{Z}_n$ through Ψ_n^R. We also introduce the enumerating function

$$V_n^R : \Upsilon \to \mathbb{Z}_+ \,, \qquad V_n^R(\boldsymbol{\upsilon}) := \left|\left\{\boldsymbol{x} \in (\mathbb{Z}_2^3)^n_{\boldsymbol{\upsilon}} : \, \Psi_n^R \boldsymbol{x} = \mathbf{0}\right\}\right|$$

counting the number of n-tuples in the kernel of Ψ_n^R of different types. It is straightforward to check that the minimum normalized distance of the random code $\mathcal{B}_n^R$ can be rewritten as

$$\frac{1}{n} \mathrm{d}_{\min}\left(\mathcal{B}_n^R\right) = \inf\left\{\langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle \mid \boldsymbol{\theta} \in \Theta : \, \pi^1_\sharp \boldsymbol{\theta} \neq \delta_0, \, U_n^R(\boldsymbol{\theta}) \geq 1\right\} \,,$$

where $[\pi^1_\sharp \boldsymbol{\theta}](\,\cdot\,) = \sum_x \boldsymbol{\theta}(\,\cdot\,, x) \in \Upsilon$ is the marginal of $\boldsymbol{\theta}$ on the first component.

The average value of the enumerating functions $U_n^R(\boldsymbol{\theta})$ and $V_n^R(\boldsymbol{\omega})$ is easily evaluated as shown in the following result.

Lemma 4 *For every $\boldsymbol{\theta} \in \Theta_n$ and $\boldsymbol{\upsilon} \in \Upsilon_n$ such that*

$$\pi^1_\sharp \boldsymbol{\theta} \neq \delta_0 \,, \qquad \boldsymbol{\upsilon} \neq \delta_0 \,,$$

it holds

$$\mathbb{E}\left[U_n^R(\boldsymbol{\theta})\right] = \binom{n}{n\boldsymbol{\theta}} \frac{1}{8^{2l}} \,, \qquad \mathbb{E}\left[V_n^R(\boldsymbol{\upsilon})\right] = \binom{n}{n\boldsymbol{\upsilon}} \frac{1}{8^{l}} \,.$$

Proof For every $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{Z}_2^{3n}$ such that $\boldsymbol{x} \neq \mathbf{0}$ we have that $\Psi_n^R \boldsymbol{x}$ and $\Psi_n^R \boldsymbol{y} - \boldsymbol{Z}_n$ are independent and both uniformly distributed over $\mathbb{Z}_2^{3l}$. It follows that

$$\begin{array}{rcl} \mathbb{E}\left[U_n^R(\boldsymbol{\theta})\right] & = & \mathbb{E}\Big[\sum_{(\boldsymbol{x},\boldsymbol{y})} \mathbb{1}_{\{\Psi_n^R \boldsymbol{x} = \mathbf{0}\}} \mathbb{1}_{\{\Psi_n^R \boldsymbol{y} = \boldsymbol{Z}_n\}}\Big] \\ & = & \sum_{(\boldsymbol{x},\boldsymbol{y})} \mathbb{P}\left(\Psi_n^R \boldsymbol{x} = \mathbf{0} \,, \; \Psi_n^R \boldsymbol{y} - \boldsymbol{Z}_n = \mathbf{0}\right) \\ & = & \binom{n}{n\boldsymbol{\theta}} \frac{1}{8^{2l}} \,, \end{array}$$

where the summations above are extended to all pairs $(\boldsymbol{x},\boldsymbol{y})$ of joint type $\boldsymbol{\theta}$. $\mathbb{E}[V_n^R(\boldsymbol{v})]$ is computed analogously. ■

A first-moment method based on Lemma 4 allows to prove the following lower bound on the typical asymptotic minimum distance of the BCE.

Proposition 2 *For every design rate* $0<R<\log 8$ *and* $0<\varepsilon<\overline{R}$, *with probability one*

$$\liminf_n \frac{1}{n} d_{\min}\left(\mathcal{B}_n^R\right) \geq \underline{\gamma}_\eta(R+\varepsilon)\,.$$

Proof See Sect.C.1. ■

From the continuity of $\underline{\gamma}_\eta(R)$ as a function of the design rate R (see Sect.A), (19) of Theorem 2 follows by an argument analogous to the one used in (33).

We now want to obtain an upper bound on the typical asymptotic normalized minimum distance of the BCE using a second-order method [1]. Toward this end, we need to estimate the variance of the joint-type-enumerating functions $U_n^R(\boldsymbol{\theta})$.

Lemma 5 *For all* $n\geq 1$, *and every joint type* $\boldsymbol{\theta}\in\Theta_n$ *such that* $\pi^1_\sharp\boldsymbol{\theta}\neq\delta_0$,

$$\operatorname{Var}\left[U_n^R(\boldsymbol{\theta})\right] \leq \binom{n}{n\boldsymbol{\theta}}\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}}\frac{16}{8^{3l}} + \binom{n}{n\boldsymbol{\theta}}^2\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}}^{-1}\frac{1}{8^{3l}} + \binom{n}{n\boldsymbol{\theta}}\frac{8}{8^{2l}}\,, \tag{34}$$

Proof See Sect.C.2. ■

A second-order method based on Lemma 4 and Lemma 5 allows one to show that, given a sequence of joint types $(\boldsymbol{\theta}_n)$ converging to some $\boldsymbol{\theta}$ lying in the interior of the set $\overline{\Omega}_{(R)}$, with probability one $U_n^R(\boldsymbol{\theta})\geq 1$ for all but finitely many n. This idea is exploited in the proof of the following upper bound on the typical asymptotic normalized minimum distance of the BCE.

Proposition 3 *For every design rate* $0<R<\log 8$ *and every* $0<\varepsilon<R$,

$$\mathbb{P}\left(\limsup_n \frac{1}{n}\,\mathrm{d}_{\min}(\mathcal{B}_n^R)\leq\overline{\gamma}_\eta(R-\varepsilon)\right)=1\,.$$

Proof Let $\boldsymbol{\theta}_\varepsilon$ in $\overline{\Omega}_{(R-\varepsilon)}$ be such that $\overline{\gamma}_\eta(R-\varepsilon)=\langle\boldsymbol{\theta}_\varepsilon,\boldsymbol{D}_\eta\rangle$. Consider a sequence of joint types $(\boldsymbol{\theta}_n)$ converging to $\boldsymbol{\theta}_\varepsilon$, with $\boldsymbol{\theta}_n$ in Θ_n for every $n\geq 1$. Define the event $A_n:=\{U_n^R(\boldsymbol{\theta}_n)=0\}$. We can apply Chebyshev inequality and use Lemma 4 and Lemma 5 obtaining

$$\mathbb{P}\left(A_n\right)\leq\frac{\operatorname{Var}\left[U_n^R(\boldsymbol{\theta}_n)\right]}{\mathbb{E}\left[U_n^R(\boldsymbol{\theta}_n)\right]^2}\leq 16\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}_n}\binom{n}{n\boldsymbol{\theta}_n}^{-1}8^l+\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}_n}^{-1}8^l+8\binom{n}{n\boldsymbol{\theta}_n}^{-1}8^{2l}\,.$$

It follows that

$$\limsup_n\frac{1}{n}\log\mathbb{P}\left(U_n^R\left(\boldsymbol{\theta}_n\right)=0\right)\leq\max\left\{\overline{R}+\mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}_\varepsilon)-\mathrm{H}(\boldsymbol{\theta}_\varepsilon),\overline{R}-\mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}_\varepsilon),2\overline{R}-2\,\mathrm{H}(\boldsymbol{\theta}_\varepsilon)\right\}\leq-\varepsilon\,.$$

so that the series $\sum_{n\geq 1}\mathbb{P}\left(A_n\right)$ is convergent, so that $\mathbb{P}\left(A_n\,\text{i.o.}\right)=0$ by the Borel-Cantelli lemma, and the claim easily follows. ■

Observe now that Proposition 3 and the monotonicity and continuity properties of $\overline{\gamma}_\eta(R)$ (see Sect.A) imply the second claim (20) of Theorem 2.

4.2 Comparing $\overline{\gamma}_\eta(R)$, $\underline{\gamma}_\eta(R)$ and $\gamma_8(R)$

We now want to compare the distance bounds $\underline{\gamma}_\eta(R)$, $\overline{\gamma}_\eta(R)$, and $\gamma_8(R)$ defined in (17), (18) and (12) respectively. First, observe that any joint type $\boldsymbol{\theta} \in \overline{\Theta}_{(R)}$ trivially satisfies $\mathrm{H}(\boldsymbol{\theta}) \geq 2\overline{R}$, so that $\overline{\Theta}_{(R)} \subseteq \underline{\Theta}_{(R)}$. From this, it immediately follows that $\overline{\gamma}_\eta(R) \geq \underline{\gamma}_\eta(R)$. Notice also that the inequality above holds as an equality whenever $\underline{\gamma}_\eta(R) = \langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle$ for some joint type $\boldsymbol{\theta}$ belonging to $\overline{\Theta}_{(R)}$. It can be shown that this is the case for every binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$ for large enough values of R, so that often $\overline{\gamma}_\eta(R)$ and $\underline{\gamma}_\eta(R)$ do coincide. However, we will now concentrate on comparing $\overline{\gamma}_\eta(R)$ with the GV-distance $\gamma_8(R)$, in particular showing that the former is strictly below the latter.

In order to do that, we start by considering the $\mathbb{Z}_8$-type $\boldsymbol{\omega}_{(R)}$ in $\Omega_{(R)}$ giving the GV-distance, i.e. such that $\gamma_8(R) = \langle \boldsymbol{\omega}_{(R)}, \boldsymbol{d} \rangle$. Since the entropy function is concave and the map $\boldsymbol{\omega} \mapsto \langle \boldsymbol{\omega}, \boldsymbol{d} \rangle$ is linear and it achieves its global minimum in δ_0, Lemma 11 can be applied to guarantee that $\mathrm{H}(\boldsymbol{\omega}_{(R)}) = \overline{R}$. Hence, using Lagrangian multipliers we get

$$\boldsymbol{\omega}_{(R)}(x) = \frac{e^{-\lambda \boldsymbol{d}(x)}}{Z(\lambda)}, \qquad Z(\lambda) := \sum_{x \in \mathbb{Z}_8} e^{-\lambda \boldsymbol{d}(x)}, \tag{35}$$

where $\lambda \in (0, +\infty)$ is the unique solution of the equation $\mathrm{H}\left(Z(\lambda)^{-1} e^{-\lambda \boldsymbol{d}}\right) = \overline{R}$.

From $\boldsymbol{\omega}_{(R)} \in \Omega$ we define a joint type $\boldsymbol{\theta}_{(R)}$ in Θ as follows. For every z in $\mathbb{Z}_2^3$, consider the bijection

$$\sigma_z : \mathbb{Z}_2^3 \to \mathbb{Z}_8, \qquad \sigma_z(x) := \mu^{-1}\left(\eta(x+z)\right) - \mu^{-1}\left(\eta(z)\right).$$

[11] Now define

$$\boldsymbol{\theta}_{(R)}(x,z) := \frac{1}{8} \boldsymbol{\omega}_{(R)}\left(\sigma_z(x)\right), \qquad x, z \in \mathbb{Z}_2^3, \tag{36}$$

and let $\boldsymbol{\upsilon}_{(R)} := \pi^1_\sharp \boldsymbol{\theta}_{(R)}$ in Υ be its marginal measure. A few simple properties of $\boldsymbol{\theta}_{(R)}$ and $\boldsymbol{\upsilon}_{(R)}$ are gathered in the following result.

Lemma 6 *For all* $0 < R < \log 8$

$$\boldsymbol{\theta}_{(R)}(x,z) > 0, \qquad \boldsymbol{\upsilon}_{(R)}(x) > 0, \qquad \forall\, x,\, z \in \mathbb{Z}_2^3, \tag{37}$$

$$\langle \boldsymbol{\theta}_{(R)}, \boldsymbol{D}_\eta \rangle = \gamma_8(R). \tag{38}$$

$$\mathrm{H}(\boldsymbol{\theta}_{(R)}) = \log 8 + \overline{R}. \tag{39}$$

$$\mathrm{H}(\boldsymbol{\upsilon}_{(R)}) > \overline{R}. \tag{40}$$

Proof See Sect. C.3. ■

We are now ready to prove the rightmost inequality in (21).

Proposition 4 *For every labeling* $\eta : \mathbb{Z}_2^3 \to \mathbb{Z}_8$ *and every rate* $0 < R < \log 8$,

$$\overline{\gamma}_\eta(R) < \gamma_8(R).$$

[11] Observe that, in the expression above, the $+$ sign refers to addition in $\mathbb{Z}_8$, while the $-$ refers to difference in $\mathbb{Z}_2^3$.

Proof For $x \in \mathbb{Z}_2^3$, define $m_x := \min\{\boldsymbol{D}_\eta(x,z) | \, z \in \mathbb{Z}_2^3\}$, $M_x := \{z \in \mathbb{Z}_2^3 : \boldsymbol{D}_\eta(x,z) = m_x\}$. Observe that, since $\boldsymbol{D}_\eta(0,z) = 0$ for every binary labeling η and any $z \in \mathbb{Z}_2^3$, we have that $m_0 = 0$ and $|M_0| = 8$. However, since no binary labeling η is isometric, we have that

$$\exists \, x, z \in \mathbb{Z}_2^3 : \; m_x < \boldsymbol{D}_\eta(x,z) \, . \tag{41}$$

For $\boldsymbol{v} \in \Upsilon$, consider the set $\Theta_{\boldsymbol{v}} := \{\boldsymbol{\theta} \in \Theta : \pi^1_\sharp \boldsymbol{\theta} = \boldsymbol{v}\}$ of joint measures with marginal $\boldsymbol{v}$, and define $f : \Upsilon \to \mathbb{R}$, $f(\boldsymbol{v}) := \min\left\{\langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle | \, \boldsymbol{\theta} \in \Theta_{\boldsymbol{v}} : \mathrm{H}(\boldsymbol{\theta}) - \mathrm{H}(\boldsymbol{v}) \geq \overline{R}\right\}$. As an immediate consequence of (40), we have that $\overline{\gamma}_\eta(R) \leq f(\boldsymbol{v}_{(R)})$, so that, in order to prove the claim, it is sufficient to show that $f(\boldsymbol{v}_{(R)}) < \gamma_8(R)$.

First, suppose that $\sum_x \boldsymbol{v}_{(R)}(x) \log n_x \geq \overline{R}$. Then, it is easy to check that $f(\boldsymbol{v}_{(R)}) = \sum_x \boldsymbol{v}_{(R)}(x) m_x$. Hence, it follows from, (37), (41) and (38) that

$$\begin{array}{rcl} f(\boldsymbol{v}_{(R)}) & = & \sum_x m_x \boldsymbol{v}_{(R)}(x) \\ & = & \sum_x \sum_z \frac{1}{8} \boldsymbol{\theta}_{(R)}(\sigma_z(x)) m_x \\ & < & \sum_x \sum_z \frac{1}{8} \boldsymbol{\theta}_{(R)}(\sigma_z(x)) \boldsymbol{D}_\eta(x,z) \\ & = & \gamma_8(R) \, , \end{array}$$

thus proving the claim.

Now, assume that $\sum_x \boldsymbol{v}_{(R)}(x) \log n_x < \overline{R}$. For any $x \neq 0$ in $\mathbb{Z}_2^3$, we have

$$\boldsymbol{v}_{(R)}(0) = \sum_{z \in \mathbb{Z}_2^3} \boldsymbol{\theta}_{(R)}(0,z) = \frac{1}{Z(\lambda)} > \sum_{z \in \mathbb{Z}_2^3} \frac{e^{-\lambda \boldsymbol{d}(\sigma_z(x))}}{8Z(\lambda)} = \sum_{z \in \mathbb{Z}_2^3} \boldsymbol{\theta}_{(R)}(x,z) = \boldsymbol{v}_{(R)}(x) \, .$$

Hence, $\boldsymbol{v}_{(R)}$ is not the uniform measure over $\mathbb{Z}_2^3$ and, as a consequence $\mathrm{H}\left(\boldsymbol{v}_{(R)}\right) < \log 8$. Therefore, from (39) and (40), $\mathrm{H}(\boldsymbol{\theta}_{(R)}) = \log 8 + \overline{R} > \mathrm{H}(\boldsymbol{v}_{(R)}) + \overline{R}$. Then, thanks to the concavity of the entropy function, we can apply Lemma 11, obtaining that $f(\boldsymbol{v}_{(R)}) < \langle \boldsymbol{\theta}_{(R)}, \boldsymbol{D}_\eta \rangle = \gamma_8(R)$, the last equality following from (38). ■

5 The typical group code achieves the expurgated exponent

In this section, we shall show that with probability one the GCE achieves the expurgated error exponent $E_8^x(R)$, thus proving Theorem 3. We shall use the union-Bhattacharyya bound [38], in order to estimate of the error probability of the GCE in terms of its type-enumerating functions:

$$p_e(\mathcal{G}_n^R) \leq \sum_{\boldsymbol{\omega} \in \Omega_n} G_n^R(\boldsymbol{\omega}) \exp(-n\langle \boldsymbol{\omega}, \boldsymbol{d} \rangle) \, . \tag{42}$$

The reader is referred to [36, 8, 24] for tighter bounds on the error probability of group codes based on their type-enumerating functions.

We start by introducing the expurgated exponents of the 4-PSK and 2-PSK AWGNC, respectively given by

$$\begin{array}{rcl} E_4^x(\frac{2}{3}R) & := & \min\left\{\langle \boldsymbol{\omega}, \boldsymbol{d} \rangle - \mathrm{H}(\boldsymbol{\omega}) + \frac{2}{3}\overline{R} | \boldsymbol{\omega} \in \Omega''_{(R)}\right\} \, , \\ E_2^x(\frac{1}{3}R) & := & \min\left\{\langle \boldsymbol{\omega}, \boldsymbol{d} \rangle - \mathrm{H}(\boldsymbol{\omega}) + \frac{1}{3}\overline{R} | \boldsymbol{\omega} \in \Omega'_{(R)}\right\} \, , \end{array}$$

where $\Omega'_{(R)}$ and $\Omega''_{(R)}$ have been defined in (26).

The following result can be proved using the first-moment method based on Lemma 2 and (42).

Proposition 5 *For every* $0 < R < \log 8$ *and every* $0 < \varepsilon < \overline{R}$*, with probability one*

$$\liminf_{n} -\frac{1}{n} \log p_e(\mathcal{G}_n^R) \geq \min\left\{E_8^x(R+\varepsilon), E_4^x(\tfrac{2}{3}(R+\varepsilon)), E_2^x(\tfrac{1}{3}(R+\varepsilon))\right\}$$

Proof For any $\varepsilon > 0$, consider the events

$$A_{n,\varepsilon} := \left\{\cup_{\boldsymbol{\omega}\in\Omega_n} G_n^R(\boldsymbol{\omega}) \geq \exp\left(n\left[\mathrm{H}(\boldsymbol{\omega}) - (\overline{R}-\varepsilon)\tfrac{\log\zeta(\boldsymbol{\omega})}{\log 8}\right]\right)\right\}\,, \qquad n \geq 1\,,$$

and define $A_\varepsilon := \{A_{n,\varepsilon} \text{ i. o.}\}$. From Markov's inequality, Lemma 2 and (2), we have

$$\mathbb{P}\left(A_{n,\varepsilon}\right) \leq \sum_{\boldsymbol{\omega}\in\Omega_n} \mathbb{E}\left[G_n^R(\boldsymbol{\omega})\right] \exp\left(-n\left(\mathrm{H}(\boldsymbol{\omega}) - (\overline{R}-\varepsilon)\frac{\log\zeta(\boldsymbol{\omega})}{\log 8}\right)\right) \leq |\Omega_n| \exp\left(-n\frac{\varepsilon}{3}\right)\,,$$

so that the series $\sum_{n\geq 1} \mathbb{P}(A_{n,\varepsilon})$ is convergent, and the Borel-Cantelli lemma implies $\mathbb{P}(A_\varepsilon) = 0$. Hence, with probability one there exists n_0 such that, for all $n \geq n_0$, $G_n^R(\boldsymbol{\omega}) = 0$ for $\boldsymbol{\omega} \in \tilde{\Omega}$ (where $\tilde{\Omega}$ is defined as in the proof of Proposition 1), and $G_n^R(\boldsymbol{\omega}) \leq \exp(n\,\mathrm{H}(\boldsymbol{\omega}) - (\overline{R} - \varepsilon)\frac{\log\zeta(\boldsymbol{\omega})}{\log 8})$ for all $\boldsymbol{\omega} \in \Omega_n$. It follows from (42) that, for $n \geq n_0$,

$$\begin{array}{rcl}
p_e(\mathcal{G}_n^R) & \leq & \sum_{\boldsymbol{\omega}\in\Omega_n\setminus\tilde{\Omega}} G_n^R(\boldsymbol{\omega}) \exp(-n\langle\boldsymbol{\omega},\boldsymbol{d}\rangle) \\
& \leq & |\Omega_n| \exp\left(-n\min\left\{\langle\boldsymbol{\omega},\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}) + (\overline{R}-\varepsilon)\frac{\log\zeta(\boldsymbol{\omega})}{\log 8} | \boldsymbol{\omega}\in\Omega\setminus\tilde{\Omega}\right\}\right) \\
& \leq & |\Omega_n| \exp\left(-n\min\left\{E_8^x(R+\varepsilon), E_4^x(\frac{2}{3}(R+\varepsilon)), E_2^x(\frac{1}{3}(R+\varepsilon))\right\}\right)\,,
\end{array}$$

and the claim follows since $|\Omega_n|$ grows polynomially fast in n. ■

The following result constitutes an analogous of Lemma 3, showing that the proper subgroups of $\mathbb{Z}_8$ cause no algebraic obstruction to the error exponent achievable by the typical group code over $\mathbb{Z}_8$.

Lemma 7 *For all* $0 \leq R \leq \log 8$,

$$E_8^x(R) \leq E_4(\tfrac{2}{3}R) \leq E_2(\tfrac{1}{3}R)\,.$$

Proof In order to prove the rightmost inequality, let $\boldsymbol{\omega}'$ be a minimizer of $\langle\boldsymbol{\omega},\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}) + \frac{1}{3}\overline{R}$ in $\Omega'_{(R)}$, and define $\boldsymbol{\omega}'' \in \Omega$ by $\boldsymbol{\omega}''(x) := 0$ for all $x \notin 2\mathbb{Z}_8$, $\boldsymbol{\omega}''(2) = \boldsymbol{\omega}''(6) := \boldsymbol{\omega}'(4)\boldsymbol{\omega}'(0)$, $\boldsymbol{\omega}''(4) := \boldsymbol{\omega}'(4)^2$, $\boldsymbol{\omega}''(0) := \boldsymbol{\omega}'(0)^2$. As in the proof of Lemma 3, it is immediate to show that $\langle\boldsymbol{\omega}',\boldsymbol{d}\rangle = \langle\boldsymbol{\omega}'',\boldsymbol{d}\rangle$, while $\mathrm{H}(\boldsymbol{\omega}'') = 2\,\mathrm{H}(\boldsymbol{\omega}')$. It follows that $\boldsymbol{\omega}'' \in \Omega''_{(R)}$, so that

$$\begin{array}{rcl}
E_4^x(\frac{2}{3}R) & = & \min\left\{\langle\boldsymbol{\omega},\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}) + \frac{2}{3}\overline{R} | \boldsymbol{\omega}\in\Omega''_{(R)}\right\} \\
& \leq & \langle\boldsymbol{\omega}'',\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}'') + \frac{2}{3}\overline{R} \\
& = & \langle\boldsymbol{\omega}',\boldsymbol{d}\rangle - 2(\mathrm{H}(\boldsymbol{\omega}') - \frac{1}{3}\overline{R}) \\
& \leq & \langle\boldsymbol{\omega}',\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}') + \frac{1}{3}\overline{R} \\
& = & E_2^x(\frac{1}{3}R)\,.
\end{array}$$

Let us now consider $E_8^x(R)$ and $E_4^x(\frac{2}{3}R)$. Define $\alpha := \sqrt{\frac{e^{-\boldsymbol{d}(4)}}{\sum_{z\in 2\mathbb{Z}_8} e^{-\boldsymbol{d}(z)}}}$, and $\hat{\boldsymbol{\omega}} \in \Omega$ as in (32). Moreover, define $\hat{\boldsymbol{\omega}}'' \in \Omega$ by $\hat{\boldsymbol{\omega}}''(x) = 0$ for $x \notin 2\mathbb{Z}_8$, $\hat{\boldsymbol{\omega}}''(4) = \alpha^2$, $\hat{\boldsymbol{\omega}}(2) = \hat{\boldsymbol{\omega}}(6) = \alpha(1-\alpha)$, while $\hat{\boldsymbol{\omega}}(0) = (1-\alpha)^2$. Then, as in the proof of Lemma 3, it is not hard to see that $\mathrm{H}(\alpha) = \frac{1}{2}\mathrm{H}(\hat{\boldsymbol{\omega}}'') = \frac{1}{3}\mathrm{H}(\hat{\boldsymbol{\omega}})$, while $\langle\hat{\boldsymbol{\omega}},\boldsymbol{d}\rangle = \langle\hat{\boldsymbol{\omega}}'',\boldsymbol{d}\rangle$. Moreover $\hat{\boldsymbol{\omega}}''$ is a global minimizer of $\langle\boldsymbol{\omega},\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}) + \frac{2}{3}\overline{R}$ in $\Omega'' := \{\boldsymbol{\omega}\in\Omega : \mathrm{supp}(\boldsymbol{\omega}) \subseteq 2\mathbb{Z}_8\}$.

Let $R_4^x = \mathrm{H}(\hat{\boldsymbol{\omega}})$ be the minimum rate R for which $\hat{\boldsymbol{\omega}} \in \Omega''_{(R)}$. By Lemma 3, for $0 < R \leq \frac{3}{2}R_4^x$, it holds

$$E_8^x(R) \leq \gamma_8(R) \leq \gamma_4(\tfrac{2}{3}R) = E_4^x(\tfrac{2}{3}R)\,.$$

Finally, for $R \geq R_4^x$, we have

$$\begin{array}{rcl} E_8^x(R) & = & \min\left\{\langle\boldsymbol{\omega},\boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}) + \overline{R} | \boldsymbol{\omega} \in \Omega_{(R)}\right\} \\ & \leq & \langle\hat{\boldsymbol{\omega}},\boldsymbol{d}\rangle - \mathrm{H}(\hat{\boldsymbol{\omega}}) + \overline{R} \\ & = & \langle\hat{\boldsymbol{\omega}}'',\boldsymbol{d}\rangle - \frac{3}{2}\left(\mathrm{H}(\hat{\boldsymbol{\omega}}'') - \frac{2}{3}\overline{R}\right) \\ & \leq & \langle\hat{\boldsymbol{\omega}}'',\boldsymbol{d}\rangle - \mathrm{H}(\hat{\boldsymbol{\omega}}'') + \frac{2}{3}\overline{R} \\ & = & E_4^x(\frac{2}{3}R)\,, \end{array}$$

which completes the proof. ■

Now, from Proposition 5, Lemma 3 and the continuity of the expurgated exponent $E_8^x(R)$ as a function of the rate R (see Sect.A), Theorem 3 follows by an argument analogous to (33).

6 The typical binary-coset code does not achieve the expurgated exponent

In this section we shall prove Theorem 4 by showing that the BCE does not achieve the expurgated bound, with probability one. First, we obtain the following upper bound on the typical error exponent which is valid at any rate R.

Proposition 6 *For all* $0 < R < \log 8$*,*

$$\limsup_n -\frac{1}{n}\log p_e(\mathcal{B}_n^R) \leq \overline{\gamma}_\eta(R) + R$$

Proof First, observe that

$$\begin{array}{rcl} p_e(\mathcal{B}_n^R) & = & \displaystyle\frac{1}{|\mathcal{B}_n^R|}\sum_{\boldsymbol{x}\in\mathcal{B}_n^R} p_e(\mathcal{B}_n^R|\boldsymbol{x}) \\ & \geq & \frac{1}{|\mathcal{B}_n^R|}\displaystyle\min_{\boldsymbol{x}\neq\boldsymbol{w}\in\mathcal{B}_n^R}\int_{\mathcal{Y}^n}\tfrac{1}{(2\pi\sigma^2)^n}e^{-\frac{||\boldsymbol{y}-\boldsymbol{x}||^2}{2\sigma^2}}\mathbb{1}_{\{||\cdot-\boldsymbol{x}||<||\cdot-\boldsymbol{w}||\}}(\boldsymbol{y})\mathrm{d}\boldsymbol{y} \\ & = & \frac{1}{|\mathcal{B}_n^R|}\varsigma\left(\sigma\sqrt{\frac{2}{\log e}}\,\mathrm{d}_{\min}(\mathcal{B}_n^R)\right)\,, \end{array}$$

where $\varsigma(t) := \int_{\sqrt{t}}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-t^2/2\pi\sigma^2}\mathrm{d}t$. It is a standard result [13, pag.2] that $\lim_n -\frac{1}{n}\log\varsigma(\alpha n) = \frac{\alpha}{\sigma}\sqrt{\frac{2}{\log e}}$. Combined with Proposition 3, this immediately implies that, with probability one,

$$\limsup_n -\frac{1}{n}\log\varsigma\left(\mathrm{d}_{\min}\left(\mathcal{B}_n^R\right)\right) \leq \limsup_n \frac{1}{n}\mathrm{d}_{\min}\left(\mathcal{B}_n^R\right) \leq \overline{\gamma}_\eta(R)\,.$$

Then, since Lemma 1 implies that $\limsup_n \frac{1}{n}\log|\mathcal{B}_N^R| \leq R$ with probability one, the claim follows. ■

Consider now the optimization problem

$$\overline{E_\eta^x}(R) := \min\left\{\langle\boldsymbol{\theta},\boldsymbol{D}_\eta\rangle - \mathrm{H}(\boldsymbol{\theta}) + \overline{R} + \log 8|\,\boldsymbol{\theta}\in\overline{\Theta}_{(R)}\right\}\,, \tag{43}$$

where $\overline{\Theta}_{(R)}$ has been defined in (16). The following result relates $\overline{E_\eta^x}(R)$ and $\overline{\gamma}_\eta(R)$ for small rates R.

Lemma 8 *There exists $\overline{R^x_8} > 0$ such that, for all $0 \leq R \leq \overline{R^x_8}$,*

$$\overline{E^x_\eta}(R) \geq \overline{\gamma}_\eta(R) + R\,,$$

and, if $\boldsymbol{\theta} \in \overline{\Theta}_{(R)}$ is a minimizer in the variational definition (43), then necessarily

$$\mathrm{H}(\boldsymbol{\theta}) - \mathrm{H}\left(\pi^1_\sharp \boldsymbol{\theta}\right) = \overline{R}\,. \tag{44}$$

Proof First, we shall show that there exists $R' > 0$ such that, for all $0 \leq R \leq R'$, if $\boldsymbol{\theta} \in \overline{\Theta}_{(R)}$ is minimizer in the definition (43) of $\overline{E^x_\eta}(R)$, then (44) holds. For a fixed $\boldsymbol{\upsilon} \in \Upsilon$, consider the set $\Theta_{\boldsymbol{\upsilon}} := \{\boldsymbol{\theta} \in \Theta : \pi^1_\sharp \boldsymbol{\theta} = \boldsymbol{\upsilon}\}$. The strictly convex function $\boldsymbol{\theta} \mapsto \langle \boldsymbol{\theta}, \boldsymbol{D}_\eta\rangle - \mathrm{H}(\boldsymbol{\theta})$ admits a unique minimizer in $\Theta_{\boldsymbol{\upsilon}}$, given by

$$\boldsymbol{\theta}^*_{\boldsymbol{\upsilon}}(x,z) := \boldsymbol{\upsilon}(x) \frac{e^{-\boldsymbol{D}_\eta(x,z)}}{\sum_z e^{-\boldsymbol{D}_\eta(x,z)}}\,.$$

Consider the uniform measure $\boldsymbol{u}$ over $\mathbb{Z}^3_2$. Then $\boldsymbol{\theta}^*_{\boldsymbol{u}}$ does not coincide with the uniform measure over $\mathbb{Z}^3_2 \times \mathbb{Z}^3_2$, so that in particular $\mathrm{H}(\boldsymbol{\theta}^*_{\boldsymbol{u}}) - \mathrm{H}(\boldsymbol{u}) < 2\log 8 - \log 8 = \log 8$. Since the map $\boldsymbol{\upsilon} \mapsto \boldsymbol{\theta}^*_{\boldsymbol{\upsilon}}$ is continuous on Υ, while the of the entropy function is convex and continuous, Lemma 10 guarantees that the map

$$R \mapsto -\overline{R} + \min\left\{\mathrm{H}(\boldsymbol{\theta}^*_{\boldsymbol{\upsilon}}) - \mathrm{H}(\boldsymbol{\upsilon}) \middle|\, \mathrm{H}(\boldsymbol{\upsilon}) \geq R\right\}$$

is continuous on the interval $[0, \log 8]$. In particular, it follows that there exists some $R' > 0$ such that, if $0 \leq R < R'$, then $\mathrm{H}(\boldsymbol{\theta}^*_{\boldsymbol{\upsilon}}) - \mathrm{H}(\boldsymbol{\upsilon}) < \overline{R}$ for all $\boldsymbol{\upsilon} \in \Omega$ with $\mathrm{H}(\boldsymbol{\upsilon}) \geq \overline{R}$. Therefore, by Lemma 11, a minimizer $\boldsymbol{\theta} \in \Theta_{\boldsymbol{\upsilon}}$ for the constrained optimization problem

$$\min\left\{\langle \boldsymbol{\theta}, \boldsymbol{D}_\eta\rangle - \mathrm{H}(\boldsymbol{\theta}) \middle|\, \boldsymbol{\theta} \in \Theta_{\boldsymbol{\upsilon}} :\, \mathrm{H}(\boldsymbol{\theta}) \geq \mathrm{H}(\boldsymbol{\upsilon}) + \overline{R}\right\}$$

necessarily satisfies $\mathrm{H}(\boldsymbol{\theta}) = \mathrm{H}(\boldsymbol{\upsilon}) + \overline{R}$ for all $\boldsymbol{\upsilon} \in \Upsilon$ such that $\mathrm{H}(\boldsymbol{\upsilon}) \geq \overline{R} > \overline{R'}$. Hence, equality (44) follows for $R < R'$, since

$$\overline{E^x_\eta}(R) = \overline{R} + \log 8 + \min\left\{\min\left\{\langle \boldsymbol{\theta}, \boldsymbol{D}_\eta\rangle - \mathrm{H}(\boldsymbol{\theta}) \middle|\, \boldsymbol{\theta} \in \Theta_{\boldsymbol{\upsilon}} :\, \mathrm{H}(\boldsymbol{\theta}) \geq \mathrm{H}(\boldsymbol{\upsilon}) + \overline{R}\right\} \middle|\, \boldsymbol{\upsilon} : \mathrm{H}(\boldsymbol{\upsilon}) \geq \overline{R}\right\}\,.$$

For any family $\overline{\boldsymbol{\vartheta}} = \{\boldsymbol{\vartheta}_x \in \Upsilon\}_{x \in \mathbb{Z}^3_2}$ of probability distributions over $\mathbb{Z}^3_2$, define $\boldsymbol{a}, \boldsymbol{h} \in \Upsilon$ by $\boldsymbol{a}_{\overline{\boldsymbol{\vartheta}}}(x) := \sum_z \boldsymbol{\vartheta}_x(z) \boldsymbol{D}_\eta(x,z)$ and $\boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}(x) := \mathrm{H}(\boldsymbol{\vartheta}_x)$, respectively. The minimum of strictly convex function $g_{\overline{\boldsymbol{\vartheta}}}(\boldsymbol{\upsilon}) := \langle \boldsymbol{\upsilon}, \boldsymbol{a}\rangle - \mathrm{H}(\boldsymbol{\upsilon}) + \overline{R}$, over the set $\Upsilon_{(\overline{\boldsymbol{\vartheta}},R)} := \left\{\boldsymbol{\upsilon} \in \Upsilon :\, \langle \boldsymbol{\upsilon}, \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}\rangle = \overline{R}\right\}$ (which is an affine subspace of Υ) is achieved by

$$\boldsymbol{\upsilon}^*_{(\overline{\boldsymbol{\vartheta}},R)} \in \Upsilon\,, \qquad \boldsymbol{\upsilon}^*_{(\overline{\boldsymbol{\vartheta}},R)}(z) := \frac{e^{-\boldsymbol{a}_{\overline{\boldsymbol{\vartheta}}}(z) - \lambda \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}(z)}}{\sum_w e^{-\boldsymbol{a}_{\overline{\boldsymbol{\vartheta}}}(w) - \lambda \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}(w)}}\,, \quad z \in \mathbb{Z}^3_2\,,$$

where $\lambda > 0$ solves the equation $\sum_x e^{-\boldsymbol{a}_{\overline{\boldsymbol{\vartheta}}}(x) - \lambda \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}(x)} \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}(x) = \overline{R} \sum_x e^{-\boldsymbol{a}_{\overline{\boldsymbol{\vartheta}}}(x) - \lambda \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}(x)}$.

Observe that (by the implicit function theorem) the map $(\overline{\boldsymbol{\vartheta}}, R) \mapsto \boldsymbol{\upsilon}^*_{(\overline{\boldsymbol{\vartheta}},R)}$ is continuous on $\Upsilon^{\mathbb{Z}^3_2} \times [0, \log 8]$. Moreover, the map

$$\Gamma : \Theta \to \Upsilon^{\mathbb{Z}^3_2}\,, \qquad \Gamma(\boldsymbol{\theta}) := \left\{\boldsymbol{\theta}|_{\{x\} \times \mathbb{Z}^3_2}\right\}_{x \in \mathbb{Z}^3_2}$$

(associating to the joint type $\boldsymbol{\theta}$ the vector of its conditioned probability distributions) is well defined and continuous on a neighborhood $\mathcal{N}$ of the uniform measure $\boldsymbol{u} \otimes \boldsymbol{u}$ on $\mathbb{Z}^3_2 \times \mathbb{Z}^3_2$. Then, the composition map $(\boldsymbol{\theta}, R) \mapsto \mathrm{H}(\boldsymbol{\upsilon}^*_{\Gamma(\boldsymbol{\theta}),R}) - \overline{R}$ is continuous on $\mathcal{N} \times [0, \log 8]$. It follows that, since H is concave on Θ and $\mathrm{H}(\boldsymbol{\theta}) = 0$ iff $\boldsymbol{\theta} = \boldsymbol{u} \otimes \boldsymbol{u}$, we can apply Lemma 10 in order to guarantee continuity of the map

$$f(R) := -\overline{R} + \min\left\{\mathrm{H}(\boldsymbol{\upsilon}^*_{(\Gamma(\boldsymbol{\theta}),R)}) \middle|\ \mathrm{H}(\boldsymbol{\theta}) \geq 2\overline{R}\right\}$$

in a nonempty interval $[0, \tilde{R})$.

Notice that, for $R = 0$ and $\overline{\boldsymbol{\vartheta}} = \overline{\boldsymbol{u}} := \{\boldsymbol{\vartheta}_x = \boldsymbol{u}\} \in \Upsilon^{\mathbb{Z}_2^3}$ constantly equal to the uniform distribution $\boldsymbol{u}$ on $\mathbb{Z}_2^3$, $\boldsymbol{h}_{\overline{\boldsymbol{u}}}(z) = \log 8$ for all z, while $0 = \boldsymbol{a}_{\overline{\boldsymbol{u}}}(0) < \boldsymbol{a}_{\overline{\boldsymbol{u}}}(x)$ for all $x \neq 0$. It follows that $\boldsymbol{\upsilon}^*_{(\overline{\boldsymbol{u}},0)} \neq \boldsymbol{u}$, so that in particular $\mathrm{H}\left(\boldsymbol{\upsilon}^*_{(\overline{\boldsymbol{u}},0)}\right) < \log 8$. Therefore, by the continuity of $f(R)$, there exists $R'' \in (0, \tilde{R})$ such that, $\mathrm{H}(\boldsymbol{\upsilon}^*_{(\Gamma(\boldsymbol{\theta}),R)}) < \overline{R}$ for all $\boldsymbol{\theta} \in \Theta$ such that $\mathrm{H}(\boldsymbol{\theta}) \geq 2\overline{R} > 2\overline{R''}$.

Hence, for all $\boldsymbol{\theta}$ and R such that $\mathrm{H}(\boldsymbol{\theta}) \geq 2\overline{R} > 2\overline{R''}$, Lemma 11 implies that a minimizer $\boldsymbol{\upsilon} \in \Upsilon_{(\Gamma(\boldsymbol{\theta}),R)}$ for the constrained optimization problem

$$\min\left\{g_{\Gamma(\boldsymbol{\theta})}(\boldsymbol{\upsilon}) - \overline{R} \middle| \boldsymbol{\upsilon} \in \Upsilon_{(\Gamma(\boldsymbol{\theta}),R)} :\ \mathrm{H}(\boldsymbol{\upsilon}) \geq R\right\} \tag{45}$$

necessarily satisfies $\mathrm{H}(\boldsymbol{\upsilon}) = \overline{R}$. Then, for all $0 < R < R_8^x := \min\{R', R''\}$, we have

$$\begin{aligned}
\overline{E^x_\eta}(R) &\overset{(a)}{=} \min\left\{\langle\boldsymbol{\theta}, \boldsymbol{D}_\eta\rangle - \mathrm{H}(\boldsymbol{\theta}) + \overline{R} + \log 8 \middle| \boldsymbol{\theta} \in \Theta :\ \mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) \geq \overline{R}, \mathrm{H}(\boldsymbol{\theta}) - \mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) = \overline{R}\right\} \\
&\overset{(b)}{=} R + \min\left\{g_{\Gamma(\boldsymbol{\theta})}(\pi^1_\sharp\boldsymbol{\theta}) - \overline{R} \middle| \boldsymbol{\theta} \in \Theta :\ \mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) \geq \overline{R}, \langle\pi^1_\sharp\boldsymbol{\theta}, \boldsymbol{h}_{\Gamma(\boldsymbol{\theta})}\rangle = \overline{R}\right\} \\
&\overset{(c)}{\geq} R + \min\left\{\min\left\{g_{\Gamma(\boldsymbol{\theta})}(\boldsymbol{\upsilon}) - \overline{R} \middle| \boldsymbol{\upsilon} \in \Upsilon_{(\Gamma(\boldsymbol{\theta}),R)} :\ \mathrm{H}(\boldsymbol{\upsilon}) \geq \overline{R}\right\} \middle| \boldsymbol{\theta} \in \overline{\Theta}_{(R)}\right\} \\
&\overset{(d)}{=} R + \min\left\{\min\left\{g_{\Gamma(\boldsymbol{\theta})}(\boldsymbol{\upsilon}) - \overline{R} \middle| \boldsymbol{\upsilon} \in \Upsilon_{(\Gamma(\boldsymbol{\theta}),R)} :\ \mathrm{H}(\boldsymbol{\upsilon}) = \overline{R}\right\} \middle| \boldsymbol{\theta} \in \overline{\Theta}_{(R)}\right\} \\
&\overset{(e)}{\geq} R + \min\left\{g_{\overline{\boldsymbol{\vartheta}}}(\boldsymbol{\upsilon}) - \overline{R} \middle| \overline{\boldsymbol{\vartheta}} \in \Upsilon^{\mathbb{Z}_2^3}, \boldsymbol{\upsilon} \in \Upsilon :\ \mathrm{H}(\boldsymbol{\upsilon}) = \overline{R}, \langle\boldsymbol{\upsilon}, \boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}\rangle = \overline{R}\right\} \\
&\overset{(f)}{=} R + \min\left\{\langle\boldsymbol{\theta}, \boldsymbol{D}_\eta\rangle \middle| \boldsymbol{\theta} \in \Theta\ :\mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) = \overline{R},\ \mathrm{H}(\boldsymbol{\theta}) - \mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) = \overline{R}\right\} \\
&\overset{(g)}{\geq} R + \overline{\gamma}_\eta(R)\,,
\end{aligned}$$

where: (a) follows from the definition (43) of $\overline{E^x_\eta}(R)$ end (44); (b) from the definitions of $g_{\overline{\boldsymbol{\vartheta}}}$, $\boldsymbol{h}_{\overline{\boldsymbol{\vartheta}}}$ and $\Gamma(\boldsymbol{\theta})$, respectively; (c) since $\pi^1_\sharp\boldsymbol{\theta} \in \Upsilon_{(\Gamma(\boldsymbol{\theta}),R)}$ for all $\boldsymbol{\theta} \in \overline{\Theta}_{(R)}$; (d) from the previous considerations on the properties of the maximizers of the problem (45); (e) from the fact that $\Gamma(\overline{\Theta}_{(R)}) \subseteq \Upsilon^{\mathbb{Z}_2^3}$; (f) by considering the surjective map

$$\varphi : \Upsilon \times \Upsilon^{\mathbb{Z}_2^3}\,, \qquad [\varphi(\boldsymbol{\upsilon}, \overline{\boldsymbol{\vartheta}})](x, z) := \boldsymbol{\upsilon}(x)\boldsymbol{\vartheta}_x(z)\,, \quad x, z \in \mathbb{Z}_2^3\,;$$

(g) from the definition (18) of $\overline{\gamma}_\eta(R)$. ■

Now, fix some rate $0 < R < \overline{R^x_8}$, and consider the joint type $\boldsymbol{\theta}_{(R)} \in \Theta$ defined by (36). It follows from Lemma 6 that $\mathrm{H}(\boldsymbol{\theta}_{(R)}) = \log 8 + \overline{R}$, $\mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) = \overline{R}$, and $\langle\boldsymbol{\theta}_{(R)}, \boldsymbol{D}_\eta\rangle = \gamma_8(R)$. In particular $\boldsymbol{\theta}_{(R)} \in \overline{\Theta}_{(R)}$ and $\mathrm{H}(\boldsymbol{\theta}) - \mathrm{H}(\pi^1_\sharp\boldsymbol{\theta}) > \overline{R}$. Then, by Lemma 8,

$$\overline{E^x_\eta}(R) < \langle\boldsymbol{\theta}_{(R)}, \boldsymbol{D}_\eta\rangle - \mathrm{H}(\boldsymbol{\theta}_{(R)}) + \overline{R} + \log 8 = \gamma_8(R)\,.$$

Therefore, again from Lemma 8, and using the fact that at low rates the expurgated exponent $E^x_8(R)$ coincides with the GV-distance $\gamma_8(R)$ (see (23)), we have

$$\overline{\gamma}_\eta(R) + R \leq \overline{E^x_\eta}(R) < \gamma_8(R) = E^x_8(R)\,, \tag{46}$$

for every $0 < R < R^x_\eta := \min\{R^x_8, \overline{R^x_\eta}\}$. Hence, by combining (46) with Proposition 6, we have finally proved Theorem 4.

7 Conclusion

In this paper we have analyzed the typical minimum distances and error exponents of two code-ensembles for the 8-PSK AWGNC with different algebraic structure. We have shown that the ensemble of group codes over $\mathbb{Z}_8$ achieves the GV bound as well as the expurgated exponent with probability one, whereas the ensemble of binary coset codes, under any possible labeling, is bounded away from the GV bound and, at low rates, from the expurgated exponent. While the paper has been focused on the specific case of the 8-PSK AWGNC, a closer look at the derivations shows that generalizations are possible to much larger classes of DMCs.

On the one hand, it is possible to consider DMCs which are symmetric with respect to the action of an arbitrary finite Abelian group G, and to characterize the typical asymptotic minimum distance achievable by the ensemble of group codes over G. This idea has been pursued in [10], where it was shown that on every $\mathbb{Z}_m$-symmetric channel, the normalized minimum distance (respectively the error exponent) of the typical group code over $\mathbb{Z}_m$ asymptotically achieves the minimum of the GV distances (the expurgated exponents) associated to all the nontrivial subgroups of $\mathbb{Z}_m$. Then, one is left to verify whether results analogous to Lemma 1 and Lemma 7 hold true, showing that proper subgroups cause no loss in the performance of the typical group code.

On the other hand, it is interesting to see how the impossibility results of Sect.4 can be generalized. Consider a DMC with input $\mathcal{X}$ of cardinality $|\mathcal{X}| = p^r$ (where p is a prime number and r a positive integer), output $\mathcal{Y}$, and transition probabilities $P(y|x)$. Define the Battacharyya distance function

$$\boldsymbol{D}(x_1, x_2) := -\log \int_{\mathcal{Y}} \sqrt{P(y|x_1)P(y|x_2)}\mathrm{d}y\,.$$

Assume that the DMC has has zero-error capacity equal to zero, so that $\boldsymbol{D}(x_1, x_2)$ is finite for every $x_1, x_2 \in \mathcal{X}$, and further that it is balanced (see [31]), i.e. that, for all $x, z \in \mathcal{X}$,

$$\{\boldsymbol{D}(x,z)|\, z \in \mathcal{X}\} = \{\boldsymbol{D}(x,z)|\, x \in \mathcal{X}\} = \{\boldsymbol{d}(x)|\, x \in \mathcal{X}\}\,,$$

for some $\boldsymbol{d} : \mathcal{X} \to \mathbb{R}$. Then, the GV distance and the expurgated exponent are respectively given by (see [11, pag.185])

$$\gamma(R) := \min_{\boldsymbol{\omega} \in \Omega_{(R)}} \{\langle \boldsymbol{\omega}, \boldsymbol{d}\rangle\}\,, \qquad E^x(R) := \min_{\boldsymbol{\omega} \in \Omega_{(R)}} \{\langle \boldsymbol{\omega}, \boldsymbol{d}\rangle - \mathrm{H}(\boldsymbol{\omega}) + R\}\,,$$

where, for $0 \leq R \leq \log|\mathcal{X}|$,

$$\overline{R} := \log|\mathcal{X}| - R\,, \qquad \Omega_{(R)} := \{\boldsymbol{\omega} \in \mathcal{P}(\mathcal{X}) : \mathrm{H}(\boldsymbol{\omega}) \geq \overline{R}\}\,.$$

Now consider the automorphism group $\mathrm{Aut}(\boldsymbol{D})$ defined as in (4). Assume that $\mathrm{Aut}(\boldsymbol{D})$ does not have any subgroup isomorphic to $\mathbb{Z}_p^r$, so that, for any binary labeling $\eta : \mathbb{Z}_p^r \to \mathcal{X}$, (7) holds, with $\boldsymbol{D}_\eta$ defined as in (6). Then, it follows that both Theorem 2 and Theorem 4 continue to hold for the ensemble of coset codes over $\mathbb{Z}_p$, which turns out to be bounded away from the GV distance at any rate, and from the expurgated exponent at low rates. Observe that, if instead $\mathrm{Aut}(\boldsymbol{D})$ does contain a subgroup isomorphic to $\mathbb{Z}_p^r$, then the arguments of [2] can be used to show that the ensemble of coset codes over $\mathbb{Z}_p$ (and in fact the ensemble of linear codes over $\mathbb{Z}_p$), achieve the GV-bound and the expurgated exponent with probability one. In other words, we have that, for balanced DMCs, having a Bhattacharyya distance function symmetric with respect to the action of the group $\mathbb{Z}_p^r$ is a necessary and sufficient condition for the typical coset codes over $\mathbb{Z}_p$ to achieve the GV-bound and the expurgated exponent.

Aknowledgements

The author is indebted to Prof. Fabio Fagnani of Politecnico di Torino for motivation, encouragement and suggestions which have led to this manuscript.

A Some lemmas on continuity

This section is devoted to the proof of the continuity of the some functions which have been defined in the paper as solutions of finite-dimensional convex optimization problems, such as the GV-distance $\gamma_8(R)$ and the expurgated error exponent $E_8^x(R)$, as well as the bounds $\overline{\gamma}_\eta(R)$ and $\underline{\gamma}_\eta(R)$. We shall obtain these results as a consequence of the general lemmas presented below.

For some fixed $d \in \mathbb{N}$, let $\Xi \subseteq \mathbb{R}^d$ be a compact and convex set. It is a standard fact that any lower semicontinuous (l.s.c.) function achieves its minimum on every closed nonempty subset $C \subseteq \Xi$. Consider two functions $g : \Xi \to \overline{\mathbb{R}}$ and $h : \Xi \to \overline{\mathbb{R}}$, and define

$$f : \mathbb{R} \to \overline{\mathbb{R}}\,, \qquad f(y) := \inf\left\{ g(\xi) \big|\, \xi \in \Xi : h(\xi) \le y \right\}\,. \tag{47}$$

It is immediate to verify that $f(y)$ is nonincreasing in y. The following simple result was proved in [9, Lemma 8.1].

Lemma 9 *If g and h are both l.s.c., then f defined in (47) is l.s.c.*

Notice that, even if g and h are both continuous, h fails in general to be continuous; in fact it is simple to provide counterexamples in this sense, when h has local minima which are not global minima. By ensuring that this cannot happen (for instance requiring that h is convex), it is possible to strengthen the previous result and prove continuity of h.

Lemma 10 *If $g : \Xi \to \mathbb{R}$ is continuous and $h : \Xi \to \mathbb{R}$ is l.s.c. and such that every local minimum is necessarily a global minimum, then f defined in (47) is continuous on $[h^*, +\infty)$ where $h^* := \min\{h(\xi)\,|\,\xi \in \Xi\}$.*

Proof Since f is nonincreasing and l.s.c. by Lemma 9, it remains to show that

$$\lim_n f(y_n) \le f(y) \tag{48}$$

for every increasing sequence $(y_n) \subset [h^*, +\infty)$ converging to some $y > h^*$. Notice that the existence of the limit in the lefthand side of (48) is guaranteed by the monotonicity of f. From the semicontinuity of g and h, there exists some ξ in Ξ such that $f(y) = g(\xi)$ and $h(\xi) \le y$. If $h(\xi) < y$, then $h(\xi) \le y_n$ for sufficiently large n, so that $f(y_n) \le g(\xi) = f(y)$ definitively in n and (48) follows. Thus we can assume that $h(\xi) = y$. Since $y > h^*$ the point ξ is not a global minimum for h. Hence, it is not even a local minimum for h, by assumption. It follows that every neighborhood of ξ in Ξ contains some $\overline{\xi}$ such that $h(\overline{\xi}) < h(x)$. It is then possible to construct a sequence (ξ_n) in Ξ converging to ξ and such that $h(\xi_n) < y$ for every n. From (ξ_n) we can extract a subsequence (ξ_{n_k}) such that $h(\xi_{n_k}) \le y_k$ for every k. Therefore we have $f(y_k) \le g(\xi_{n_k})$ and so

$$\lim_n f(y_n) \le \limsup_k g(\xi_{n_k}) \le g(\xi) = f(y)\,,$$

thus concluding the proof. ■

By considering $\Xi = \Omega$, $h(\boldsymbol{\omega}) = -\operatorname{H}(\boldsymbol{\omega})$ and $g(\boldsymbol{\omega}) = \langle \boldsymbol{\omega}, \boldsymbol{d} \rangle$ (respectively $g(\boldsymbol{\omega}) = \langle \boldsymbol{\omega}, \boldsymbol{d} \rangle + \overline{R} - \operatorname{H}(\boldsymbol{\omega})$), Lemma 10 implies the continuity of $\gamma_8(R)$ ($E_8^x(R)$). Indeed, observe that $-\operatorname{H}(\,\cdot\,)$ is convex and therefore does not admit local minima which are not global minima. Similarly, the continuity of $\underline{\gamma_\eta}(R)$ follows by taking $\Xi = \Theta$, $g(\boldsymbol{\theta}) = \langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle$, and $h(\boldsymbol{\theta}) = \max\{-\frac{1}{2}\operatorname{H}(\boldsymbol{\theta}), -\operatorname{H}(\pi^1_{\sharp\boldsymbol{\theta}})\}$, which is convex, as it is the maximum of two convex functions.

Finally, the continuity of $\overline{\gamma}_\eta(R)$ follows from Lemma 10 again with $\Xi = \Theta$, $g(\boldsymbol{\theta}) = \langle \boldsymbol{\theta}, \boldsymbol{D}_\eta \rangle$, and $h(\boldsymbol{\theta}) = \max\{-\operatorname{H}(\boldsymbol{\theta}) + \operatorname{H}(\pi^1_\sharp \boldsymbol{\theta}), -\operatorname{H}(\pi^1_{\sharp\boldsymbol{\theta}})\}$. In this last case, the absence of strictly local minima of h can be verified directly as follows. If $\boldsymbol{\theta} \in \Theta$ is a local minimum for $h(\boldsymbol{\theta}) = -\operatorname{H}(\boldsymbol{\theta}) + \operatorname{H}(\pi^1_{\sharp\boldsymbol{\theta}}) = -\sum_x [\pi^1_\sharp \boldsymbol{\theta}](x) \operatorname{H}(\boldsymbol{\theta}|_{\{x\}\times\mathbb{Z}_2^3})$, then, for every $x \in \operatorname{supp}(\pi^1_\sharp \boldsymbol{\theta})$, necessarily $\boldsymbol{\theta}|_{\{x\}\times\mathbb{Z}_2^3}$ is the uniform distribution over $\mathbb{Z}_2^3$; it follows that $h(\boldsymbol{\theta}) = -\log 8$, and therefore $\boldsymbol{\theta}$ is a global minimum.

We end this section with the following result, giving sufficient conditions for the minimizer of a convex optimization problem to satisfy the constraint with equality.

Lemma 11 *Let $g, h : \Xi \to \mathbb{R}$ be convex functions. Let $g^* := \min_{\xi\in\Xi} g(\xi)$ be the global minimum of g, and consider the set $\Xi^* := \{\xi : g(\xi) = g^*\}$ where such minimum is achieved. Then, for all $y < h^* := \min\limits_{\xi\in\Xi^*} h(\xi)$, any minimizer ξ_y for the convex optimization problem*

$$f(y) := \min_{h(\xi)\le y} g(\xi)$$

necessarily satisfies $g(\xi_y) = y$.

Proof Let $\xi \in \Xi$ be such that $h(\xi) \le y$, and $g(\xi) = f(y)$. Since $h(\xi) \le y < h^*$, necessarily $g(\xi) > g^*$. Consider some $\xi^* \in \Xi^*$ such that $h^* = h(\xi^*)$, and, for $0 \le \lambda \le 1$, define $\xi_\lambda := \lambda\xi + (1-\lambda)\xi^*$. Then, by the convexity of h, we have

$$h(\xi_\lambda) \le \lambda h(\xi) + (1-\lambda)h(\xi^*) \le \lambda y + (1-\lambda)h^*\,,$$

so that, since $y < h^*$, there exists $0 < \lambda^* < 1$ such that $h(\xi_{\lambda^*}) \le h(\xi) \le y$. From the convexity of g, it follows that

$$g(\xi_{\lambda^*}) \le \lambda^* g(\xi) + (1-\lambda^*)g(\xi^*) = \lambda^* g(\xi) + (1-\lambda^*)g^* < g(\xi)\,.$$

Then, $f(y) = \min\limits_{h(\xi)\le y} g(\xi) \le g(\xi_{\lambda^*}) < g(\xi)$, so that ξ cannot be a minimizer. ■

B An upper bound on the typical asymptotic minimum distance of the GCE

In this section we shall show that the bound (33) is tight, i.e. that the asymptotic normalized minimum distance of the GCE does not exceed the GV distance $\gamma_8(R)$, thus completing the proof of Theorem 13. Our arguments are will involve an application of the second moment method [1, pagg.43-63], and the key point consists in estimating the variance of the type-enumerating function $G_n^R(\boldsymbol{\omega})$.

We start with some preliminary considerations about the structure of the product set $(\mathbb{Z}_8)^n_{\boldsymbol{\omega}} \times (\mathbb{Z}_8)^n_{\boldsymbol{\omega}}$, $\boldsymbol{\omega} \in \Omega_n$ being some $\mathbb{Z}_8$-type. Let $m = \zeta(\boldsymbol{\omega})$ be the order of the smallest subgroup of $\mathbb{Z}_8$ supporting $\boldsymbol{\omega}$, and consider two non necessarily distinct n-tuples $\boldsymbol{x}$ and $\boldsymbol{z}$ of type $\boldsymbol{\omega}$. Let $< \boldsymbol{x} >$, $< \boldsymbol{z} >$ and $< \boldsymbol{x}, \boldsymbol{z} >$ be the subgroups of $\mathbb{Z}_8^n$ respectively generated by

$\boldsymbol{x}$, by $\boldsymbol{z}$, and by $\boldsymbol{x}$ and $\boldsymbol{z}$. It is easy to realize that both $< \boldsymbol{x} >$ and $< \boldsymbol{z} >$ are isomorphic to $\frac{8}{m}\mathbb{Z}_8$. Moreover, define: the isomorphism

$$i :< \boldsymbol{x} > \to \tfrac{8}{m}\mathbb{Z}_8 \,, \qquad i(\alpha\boldsymbol{x}) := \alpha;$$

the standard injections

$$j_1 :< \boldsymbol{x} > \to < \boldsymbol{x}, \boldsymbol{z} >, \qquad j_1(\alpha\boldsymbol{x}) = \alpha\boldsymbol{x}\,,$$

$$j_2 : \tfrac{8}{m}\mathbb{Z}_8 \to \tfrac{8}{m}\mathbb{Z}_8 \oplus \tfrac{8}{m}\mathbb{Z}_8, \qquad j_2(k) = (k, 0)\,;$$

the surjective homomorphism

$$f : \frac{8}{\zeta(\boldsymbol{\omega})}\mathbb{Z}_8 \oplus \frac{8}{\zeta(\boldsymbol{\omega})}\mathbb{Z}_8 \to < \boldsymbol{x}, \boldsymbol{z} >, \qquad f(a, b) = a\boldsymbol{x} + b\boldsymbol{z}\,.$$

Then, we have that

$$j_1 = f \circ j_2 \circ i\,.$$

In other words, the following diagram commutes

$$\begin{tikzcd}
< \boldsymbol{x} > \arrow[r, hook, "j_1"] \arrow[d, "i"] & < \boldsymbol{x}, \boldsymbol{z} > \\
\tfrac{8}{m}\mathbb{Z}_8 \arrow[r, hook, "j_2"] & \tfrac{8}{m}\mathbb{Z}_8 \oplus \tfrac{8}{m}\mathbb{Z}_8 \arrow[u, "f"']
\end{tikzcd}$$

It follows that $< \boldsymbol{x}, \boldsymbol{z} >$ contains a subgroup isomorphic to $\frac{8}{m}\mathbb{Z}_8$ and is itself isomorphic to a subgroup of $\frac{8}{m}\mathbb{Z}_8 \oplus \frac{8}{m}\mathbb{Z}_8$. An immediate consequence is that $< \boldsymbol{x}, \boldsymbol{z} >$ is isomorphic to a group of type $\frac{8}{m}\mathbb{Z}_8 \oplus \frac{8}{h}\mathbb{Z}_8$ for some h dividing m (possibly $h = 1$ when $\boldsymbol{x} = \boldsymbol{w}$). It is then possible to partition the set of ordered pairs of n-tuples of type $\boldsymbol{\omega}$ as follows:[12]

$$(\mathbb{Z}_8)^n_{\boldsymbol{\omega}} \times (\mathbb{Z}_8)^n_{\boldsymbol{\omega}} = \bigcup_{h|\zeta(\boldsymbol{\omega})} A_{n,\boldsymbol{\omega},h}\,, \tag{49}$$

with $A_{n,\boldsymbol{\omega},h}$ denoting the set of all pairs $(\boldsymbol{x}, \boldsymbol{z})$ in $(\mathbb{Z}_8)^n_{\boldsymbol{\omega}} \times (\mathbb{Z}_8)^n_{\boldsymbol{\omega}}$ such that the subgroup $< \boldsymbol{x}, \boldsymbol{z} >$ generated by $\boldsymbol{x}$ and $\boldsymbol{z}$ is isomorphic to $\frac{8}{m}\mathbb{Z}_8 \oplus \frac{8}{h}\mathbb{Z}_8$. The following lemma provides an estimation of the cardinality of $A_{n,\boldsymbol{\omega},h}$, with h ranging over the set of divisors of $\zeta(\boldsymbol{\omega})$.

Lemma 12 *For every n, $\boldsymbol{\omega}$ in $\mathcal{P}_n(\mathbb{Z}_8)$, and h dividing $\zeta(\boldsymbol{\omega})$, we have*

$$|A_{n,\boldsymbol{\omega},h}| \leq 4\binom{n}{n\boldsymbol{\omega}} \prod_{\substack{1\leq i\leq 8/h:\\ \boldsymbol{\omega}\left(i+\frac{8}{h}\mathbb{Z}_8\right)>0}} \binom{n_i}{n_i\boldsymbol{\omega}|_{i+\frac{8}{h}\mathbb{Z}_8}}\,, \tag{50}$$

where $n_i := n\boldsymbol{\omega}\left(i + \frac{8}{h}\mathbb{Z}_8\right)$ is the number of entries from the coset $i + \frac{8}{h}\mathbb{Z}_8$ in any n-tuple of type $\boldsymbol{\omega}$.

Proof Let $\boldsymbol{x}$ and $\boldsymbol{z}$ be in $(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}$. A necessary condition for the subgroup $< \boldsymbol{x}, \boldsymbol{z} >$ to be isomorphic to $\frac{8}{\zeta(\boldsymbol{\omega})}\mathbb{Z}_8 \oplus \frac{8}{h}\mathbb{Z}_8$ is the existence of some α in the set $\mathbb{Z}_8^*$ of invertible elements of Z_8, such that

$$-h\alpha\boldsymbol{x} + h\boldsymbol{z} = \boldsymbol{0}\,. \tag{51}$$

[12] For two integers a and b $a \mid b$ means that a divides b.

For (51) to hold, necessarily $\boldsymbol{z}$ has to belong to the coset $\alpha\boldsymbol{x}+\frac{8}{h}\mathbb{Z}_8^n$. Thus, whenever (51) holds, the set of positions of the entries of $\boldsymbol{x}$ belonging to any coset $i+\frac{8}{h}\mathbb{Z}_8$ and the set of positions of the entries of $\boldsymbol{z}$ belonging to the coset $\alpha i+\frac{8}{h}\mathbb{Z}_8$ need to coincide, i.e.

$$\boldsymbol{x}^{-1}\left(i+\tfrac{8}{h}\mathbb{Z}_8\right)=\boldsymbol{z}^{-1}\left(\alpha i+\tfrac{8}{h}\mathbb{Z}_8\right)\,,\qquad \forall\, i\in\mathbb{Z}_8\,. \tag{52}$$

Notice that since both $\boldsymbol{x}$ and $\boldsymbol{z}$ are assumed to be of type $\boldsymbol{\omega}$, (52) in particular implies

$$\boldsymbol{\omega}\left(i+\tfrac{8}{h}\mathbb{Z}_8\right)=\boldsymbol{\omega}\left(\alpha i+\tfrac{8}{h}\mathbb{Z}_8\right)\,,\qquad \forall\, i\in\mathbb{Z}_8\,. \tag{53}$$

For those α for which (53) is not satisfied there exists no pair $(\boldsymbol{x},\boldsymbol{z})$ satisfying (51). Thus, with no loss of generality we can restrict ourselves to considering values of α such that (53) is satisfied (as it is the case always for $\alpha=1$).

Notice that a necessary and sufficient condition for $\boldsymbol{x}$ and $\boldsymbol{z}$ both to belong to $(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}$ is the existence of an index permutation $\sigma\in S_n$ such that $\sigma\boldsymbol{x}:=\boldsymbol{x}\circ\sigma^{-1}=\boldsymbol{z}$. Equation (52) can be read as a constraint on the structure of σ, which has necessarily to be of the form

$$\sigma=\sigma^1\circ\sigma^2\circ\ldots\circ\sigma^{8/h}\circ\tilde{\sigma}_{\alpha,\boldsymbol{x}}\,, \tag{54}$$

where

- $\tilde{\sigma}_{\alpha,\boldsymbol{x}}$ is the index permutation mapping, for every coset $i+\frac{8}{h}\mathbb{Z}_8$, the smallest element of $\boldsymbol{x}^{-1}\left(\alpha^{-1}i+\frac{8}{h}\mathbb{Z}_8\right)$ in the smallest element of $\boldsymbol{x}^{-1}\left(i+\frac{8}{h}\mathbb{Z}_8\right)$, the second smallest element $\boldsymbol{x}^{-1}\left(\alpha^{-1}i+\frac{8}{h}\mathbb{Z}_8\right)$ in the second smallest element of $\boldsymbol{z}^{-1}\left(i+\frac{8}{h}\mathbb{Z}_8\right)$, and so on;
- for every coset $i+\frac{8}{h}\mathbb{Z}_8$ instead, $\sigma^i\in S_n$ is any permutation such that

$$\sigma^i(j)=j\,,\qquad \forall\, j\in\{1,\ldots,n\}\setminus\boldsymbol{x}^{-1}\left(i+\frac{8}{h}\mathbb{Z}_8\right)\,. \tag{55}$$

Thus, for a given $\boldsymbol{x}$ in $(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}$ and α in $\mathbb{Z}_8^*$ such that (53) is satisfied, we have that the number of $\boldsymbol{z}$ in $(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}$ satisfying (52) equals the cardinality of the orbit of $\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x}$ under the action of the group of index permutations

$$G^{(\boldsymbol{x})}:=\{\sigma=\sigma^1\circ\sigma^2\circ\ldots\circ\sigma^{8/h}\;:\;(55)\;\forall\,1\le i\le\tfrac{8}{h}\}\,.$$

Clearly the order of this group is $\left|G^{(\boldsymbol{x})}\right|=\prod_{i=1}^{8/h}n_i!$, while the cardinality of the stabilizer of $\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x}$ in $G^{(\boldsymbol{x})}$ is $\left|\mathrm{Stab}\left(\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x},G^{(\boldsymbol{x})}\right)\right|=\prod_{i=1}^{8}(n\boldsymbol{\omega}(i))!$, so that the orbit of $\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x}$ in $G^{(\boldsymbol{x})}$ has cardinality

$$\left|O\left(G^{(\boldsymbol{x})},\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x}\right)\right|=\prod_{i=1}^{8/h}n_i!/\prod_{i=1}^{8}(n\boldsymbol{\omega}(i))!=\prod_{i=1}^{8/h}\binom{n_i}{n_i\boldsymbol{\omega}|_{i+\frac{8}{h}\mathbb{Z}_8}},.$$

This allows us to estimate the cardinality of $A_{n,\boldsymbol{\omega},h}$ as follows

$$|A_{n,\boldsymbol{\omega},h}|\le|\mathbb{Z}_8^*|\sum_{\boldsymbol{x}\in(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}}\left|O\left(G^{(\boldsymbol{x})},\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x}\right)\right|=4\binom{n}{n\boldsymbol{\omega}}\prod_{i=1}^{8/h}\binom{n_i}{n_i\boldsymbol{\omega}|_{i+\frac{8}{h}\mathbb{Z}_8}}\,,$$

hence proving the claim. ■

Lemma 13 *For every $n \in \mathbb{N}$ and $\boldsymbol{\omega}$ in $\mathcal{P}_n(\mathbb{Z}_8)$*

$$\operatorname{Var}\left[G_n^R(\boldsymbol{\omega})\right] \leq 4\binom{n}{n\boldsymbol{\omega}}\left(\frac{1}{\zeta(\boldsymbol{\omega})}\right)^l \sum_{\substack{h|\zeta(\boldsymbol{\omega})\\ h<\zeta(\boldsymbol{\omega})}} \left(\frac{1}{h}\right)^l \prod_{i=1}^{8/h} \binom{n_i}{n_i\boldsymbol{\omega}|_{i+\frac{8}{h}\mathbb{Z}_8}}, \qquad (56)$$

for n_i defined as in Lemma 12.

Proof Assume that $\boldsymbol{x}, \boldsymbol{z} \in A_{n,\boldsymbol{\omega},h}$ for some $h \mid \zeta(\boldsymbol{\omega})$. Notice that, for every $1 \leq j \leq l$, the image of the evaluation homomorphism

$$\Lambda_j : \hom(\mathbb{Z}_8^n, \mathbb{Z}_8) \to \mathbb{Z}_8^2, \qquad \Lambda_j(\Phi) = ((\Phi\boldsymbol{x})_j, (\Phi\boldsymbol{z})_j)$$

coincides with the subgroup $< (x_i, z_i) >_{1\leq i\leq n}$ of $\mathbb{Z}_8^2$ generated by $\{(x_i, z_i)\}_{1\leq i\leq n}$. Hence, (see [8, Lemma 9] for instance) each pair $\left((\Phi_n^R\boldsymbol{x})_j, (\Phi_n^R\boldsymbol{z})_j\right)$ is uniformly distributed over $< (x_i, z_i) >_{1\leq i\leq n} \leq \mathbb{Z}_8 \oplus \mathbb{Z}_8$, which is isomorphic to $< \boldsymbol{x}, \boldsymbol{z} >$. As $< \boldsymbol{x}, \boldsymbol{z} >$ is in turn is isomorphic to a group of type $\frac{8}{\zeta(\boldsymbol{\omega})}\mathbb{Z}_8 \oplus \frac{8}{h}\mathbb{Z}_8$, it follows that

$$\mathbb{P}\left((\Phi_n^R\boldsymbol{x})_j, (\Phi_n^R\boldsymbol{z})_j = 0\right) = (h\zeta(\boldsymbol{\theta}))^{-1}, \qquad \forall 1 \leq j \leq l,$$

and, since the r.v.s $\left((\Phi_n^R\boldsymbol{x})_j, (\Phi_n^R\boldsymbol{z})_j\right)_{1\leq j\leq l}$ are mutually independent,

$$\mathbb{P}\left(\Phi_n^R\boldsymbol{x} = \boldsymbol{0}, \Phi_n^R\boldsymbol{z} = \boldsymbol{0}\right) = (h\zeta(\boldsymbol{\omega}))^{-l}. \qquad (57)$$

It follows from (49), (50) and (57) that

$$\begin{aligned}
\operatorname{Var}\left[G_n^R(\boldsymbol{\omega})\right] &= \sum_{\boldsymbol{x},\boldsymbol{w}\in(\mathbb{Z}_8)^n_{\boldsymbol{\omega}}} \operatorname{Cov}\left[\mathbb{1}_{\{\Phi_n^R\boldsymbol{x}=\boldsymbol{0}\}}\mathbb{1}_{\{\Phi_n^R\boldsymbol{z}=\boldsymbol{0}\}}\right] \\
&= \sum_{h|\zeta(\boldsymbol{\omega})} \sum_{(\boldsymbol{x},\boldsymbol{z})\in A_{n,\boldsymbol{\omega},h}} \mathbb{P}\left(\Phi_n^R\boldsymbol{x} = \boldsymbol{0}, \Phi_n^R\boldsymbol{z} = \boldsymbol{0}\right) - \mathbb{P}\left(\Phi_n^R\boldsymbol{x} = \boldsymbol{0}\right)\mathbb{P}\left(\Phi_n^R\boldsymbol{z} = \boldsymbol{0}\right) \\
&= \sum_{h|\zeta(\boldsymbol{\omega})} |A_{n,\boldsymbol{\omega},h}| \left(\tfrac{1}{h^l\zeta(\boldsymbol{\omega})^l} - \tfrac{1}{\zeta(\boldsymbol{\omega})^{2l}}\right),
\end{aligned}$$

and the claim follows immediately from Lemma 12. ■

For every h dividing 8, consider the projection τ^h of $\mathbb{Z}_8$ onto the quotient group $\mathbb{Z}_8\big/\frac{8}{h}\mathbb{Z}_8$, defined by

$$\tau^h(x) = y + \tfrac{8}{h}\mathbb{Z}_8 \qquad \Leftrightarrow \qquad x \in y + \tfrac{8}{h}\mathbb{Z}_8,$$

and, for every $\boldsymbol{\omega}$ in Ω denote by $\tau_\sharp^h\boldsymbol{\omega}$ in $\mathcal{P}\left(\mathbb{Z}_8\big/\frac{8}{h}\mathbb{Z}_8\right)$ the image measure under τ^h. As an immediate consequence of Lemma 2 and Lemma 13 we have

$$\begin{aligned}
\limsup_n \tfrac{1}{n}\log\left(\tfrac{\operatorname{Var}\left[G_n^R(\boldsymbol{\omega})\right]}{\mathbb{E}[G_n^R(\boldsymbol{\omega})]^2}\right) &\leq \limsup_n \tfrac{1}{n}\log\left[\tbinom{n}{n\boldsymbol{\omega}}^{-1} \sum_{\substack{h|\zeta(\boldsymbol{\omega})\\ h<\zeta(\boldsymbol{\omega})}} \left(\tfrac{\zeta(\boldsymbol{\omega})}{h}\right)^l \textstyle\prod_{i=1}^{8/h} \tbinom{n_i}{n_i\boldsymbol{\omega}|_{i+\frac{8}{h}\mathbb{Z}_8}}\right] \\
&= \max_{\substack{h|\zeta(\boldsymbol{\omega})\\ h<\zeta(\boldsymbol{\omega})}} \left\{\tfrac{\log\zeta(\boldsymbol{\omega})/h}{\log 8}\overline{R} - \mathrm{H}(\boldsymbol{\omega}) + \sum_{i=1}^{8/h} \boldsymbol{\omega}(i + \tfrac{8}{h}\mathbb{Z}_8)\,\mathrm{H}\left(\boldsymbol{\omega}|_{i+\frac{8}{h}\mathbb{Z}_8}\right)\right\} \\
&= \max_{\substack{h|\zeta(\boldsymbol{\omega})\\ h<\zeta(\boldsymbol{\omega})}} \left\{\tfrac{\log\zeta(\boldsymbol{\omega})/h}{\log 8}\overline{R} - \mathrm{H}\left(\tau_\sharp^h(\boldsymbol{\omega})\right)\right\}.
\end{aligned} \qquad (58)$$

We are now ready to state the following result, whose proof relies on relies on an application of the second moment method and the key point consists in showing that the $\mathbb{Z}_8$-type $\boldsymbol{\omega}$ minimizing the righthand side of (12) can be approximated by $\mathbb{Z}_8$-types $\boldsymbol{\omega}_\varepsilon$ such that $\mathrm{H}\left(\tau_\sharp^h\boldsymbol{\omega}_\varepsilon\right) > \frac{\log l(\boldsymbol{\omega}_\varepsilon)/h}{\log 8}\overline{R}$ for all $h = 1, 2, 4$.

Proposition 7 *For every* $0 < R < \log 8$*, and* $0 < \varepsilon < \overline{R}$*,*

$$\mathbb{P}\left(\limsup_n \frac{1}{n} d_{\min}\left(\mathcal{G}_n^R\right) \geq \gamma_8(R-\varepsilon)\right) = 1\,.$$

Proof Let $\boldsymbol{\omega} = \boldsymbol{\omega}_{(R-\varepsilon)} \in \Omega$ be the $\mathbb{Z}_8$-type achieving the GV distance, defined as in (35). Observe that (30) and (31) imply that $\boldsymbol{\omega}$ has the following ordering

$$\boldsymbol{\omega}(0) > \boldsymbol{\omega}(1) = \boldsymbol{\omega}(7) > \boldsymbol{\omega}(2) = \boldsymbol{\omega}(6) > \boldsymbol{\omega}(3) = \boldsymbol{\omega}(5) > \boldsymbol{\omega}(4)\,. \tag{59}$$

Define $A_0 := \{0,1,7,2\}\,, B_0 := \{0,1,6,3\}\,, C_0 := \{0,5,6,7\}\,,$ and let A_1, B_1 and C_1 be the complements in $\mathbb{Z}_8$ respectively of A_0, B_0 and C_0. It follows from (59) that

$$\begin{array}{ll} \boldsymbol{\omega}(A_0) \geq \boldsymbol{\omega}(2\mathbb{Z}_8)\,, & \boldsymbol{\omega}(A_0) \geq \boldsymbol{\omega}(2\mathbb{Z}_8+1)\,, \\ \boldsymbol{\omega}(B_0) \geq \boldsymbol{\omega}(2\mathbb{Z}_8)\,, & \boldsymbol{\omega}(B_0) \geq \boldsymbol{\omega}(2\mathbb{Z}_8+1)\,, \\ \boldsymbol{\omega}(C_0) \geq \boldsymbol{\omega}(2\mathbb{Z}_8)\,, & \boldsymbol{\omega}(C_0) \geq \boldsymbol{\omega}(2\mathbb{Z}_8+1)\,. \end{array} \tag{60}$$

Moreover, it is easy to check that $|A_a \cap B_b \cap C_c| = 1$, for every choice of (a,b,c) in $\{0,1\}^3$. Thus, $f : \mathbb{Z}_8 \to \{0,1\}^3$, where $f(x) = (a,b,c)$ if and only if x is in $A_a \cap B_b \cap C_c$, is a bijection. Then, from (1) and (60), it thus follows that

$$\mathrm{H}(\boldsymbol{\omega}) = \mathrm{H}\left(f_\sharp \boldsymbol{\omega}\right) \geq \mathrm{H}\left(\boldsymbol{\omega}(A_0)\right) + \mathrm{H}\left(\boldsymbol{\omega}(B_0)\right) + \mathrm{H}\left(\boldsymbol{\omega}(C_0)\right) \geq 3\,\mathrm{H}\left(\boldsymbol{\omega}(2\mathbb{Z}_8)\right) = 3\,\mathrm{H}\left(\tau_\sharp^4 \boldsymbol{\omega}\right)\,. \tag{61}$$

Let us now introduce the sets $D := \{0,2\}$ and $E := \{1,7\}$. We have from (59) that

$$\begin{array}{ll} \boldsymbol{\omega}(D) \geq \boldsymbol{\omega}(4\mathbb{Z}_8)\,, & \boldsymbol{\omega}(D) \geq \boldsymbol{\omega}(4\mathbb{Z}_8+2)\,, \\ \boldsymbol{\omega}(E) \geq \boldsymbol{\omega}(4\mathbb{Z}_8+1)\,, & \boldsymbol{\omega}(E) \geq \boldsymbol{\omega}(4\mathbb{Z}_8+3)\,. \end{array}$$

It thus follows that

$$\begin{array}{rcl} \mathrm{H}\left(\tau_\sharp^2 \boldsymbol{\omega}\right) & = & \mathrm{H}\left(\tau_\sharp^4 \boldsymbol{\omega}\right) + \boldsymbol{\omega}(2\mathbb{Z}_8)\,\mathrm{H}\left(\tau_\sharp^4 \boldsymbol{\omega}|_{2\mathbb{Z}_8}\right) + \boldsymbol{\omega}(2\mathbb{Z}_8+1)\,\mathrm{H}\left(\tau_\sharp^4 \boldsymbol{\omega}|_{2\mathbb{Z}_8+1}\right) \\ & \geq & \mathrm{H}(\boldsymbol{\omega}(2\mathbb{Z}_8)) + \boldsymbol{\omega}(2\mathbb{Z}_8)\,\mathrm{H}\left(\boldsymbol{\omega}|_{2\mathbb{Z}_8}(D)\right) + \boldsymbol{\omega}(2\mathbb{Z}_8+1)\,\mathrm{H}\left(\boldsymbol{\omega}|_{2\mathbb{Z}_8+1}(E)\right)\,. \end{array} \tag{62}$$

Observe that

$$\boldsymbol{\omega}|_{2\mathbb{Z}_8}(D) = \frac{\boldsymbol{\omega}(0)}{\boldsymbol{\omega}(2\mathbb{Z}_8)} + \frac{\boldsymbol{\omega}(2)}{\boldsymbol{\omega}(2\mathbb{Z}_8)} = \boldsymbol{\omega}|_{2\mathbb{Z}_8}(4\mathbb{Z}_8)\boldsymbol{\omega}|_{4\mathbb{Z}_8}(0) + \boldsymbol{\omega}|_{2\mathbb{Z}_8}(4\mathbb{Z}_8+2)\boldsymbol{\omega}|_{4\mathbb{Z}_8+2}(2)$$

so that, by the concavity of the entropy function, we get

$$\mathrm{H}\left(\boldsymbol{\omega}|_{2\mathbb{Z}_8}(D)\right) \geq \boldsymbol{\omega}|_{2\mathbb{Z}_8}(4\mathbb{Z}_8)\,\mathrm{H}\left(\boldsymbol{\omega}|_{4\mathbb{Z}_8}(0)\right) + \boldsymbol{\omega}|_{2\mathbb{Z}_8}(4\mathbb{Z}_8+2)\,\mathrm{H}\left(\boldsymbol{\omega}|_{4\mathbb{Z}_8+2}(2)\right)\,.$$

An analogous reasoning leads to

$$\mathrm{H}\left(\boldsymbol{\omega}|_{2\mathbb{Z}_8+1}(E)\right) \geq \boldsymbol{\omega}|_{2\mathbb{Z}_8+1}(4\mathbb{Z}_8+1)\,\mathrm{H}\left(\boldsymbol{\omega}|_{4\mathbb{Z}_8+1}(1)\right) + \boldsymbol{\omega}|_{2\mathbb{Z}_8+1}(4\mathbb{Z}_8+3)\,\mathrm{H}\left(\boldsymbol{\omega}|_{4\mathbb{Z}_8+3}(3)\right)\,.$$

Upon substituting the two inequalities above in (62), we get

$$\begin{array}{rcl} \mathrm{H}(\tau_\sharp^2 \boldsymbol{\omega}) & \geq & \mathrm{H}(\tau_\sharp^4 \boldsymbol{\omega}) + \sum_{i=0}^3 \boldsymbol{\omega}(4\mathbb{Z}_8+i)\,\mathrm{H}(\boldsymbol{\omega}|_{4\mathbb{Z}_8+i}(i)) \\ & = & \mathrm{H}(\tau_\sharp^4 \boldsymbol{\omega}) + \mathrm{H}(\boldsymbol{\omega}) - \mathrm{H}(\tau_\sharp^2 \boldsymbol{\omega}) \\ & \geq & \frac{4}{3}\,\mathrm{H}(\boldsymbol{\omega}) - \mathrm{H}(\tau_\sharp^2 \boldsymbol{\omega})\,, \end{array}$$

last inequality following from (61). Then

$$\mathrm{H}\left(\tau_\sharp^2\boldsymbol{\omega}\right) \geq \frac{2}{3}\,\mathrm{H}(\boldsymbol{\omega})\,. \tag{63}$$

Now let $(\boldsymbol{\omega}_n)$ ba a sequence of $\mathbb{Z}_8$-types converging to $\boldsymbol{\omega}$, with $\boldsymbol{\omega}_n \in \Omega_n$ for every n. By successively applying Chabyshev inequality, (58), (61) and (63), we get

$$\begin{array}{rcl}\limsup\limits_n \frac{1}{n}\log\mathbb{P}\left(G_n^R(\boldsymbol{\omega}_n)=0\right) & \leq & \limsup\limits_n \frac{1}{n}\log\frac{\mathrm{Var}\left[G_n^R(\boldsymbol{\omega}_n)\right]}{\mathbb{E}[G_n^R(\boldsymbol{\omega}_n)]^2} \\ & \leq & \max\left\{\frac{1}{3}\overline{R}-\mathrm{H}(\boldsymbol{\tau}_\sharp^4\boldsymbol{\omega}),\,\frac{2}{3}\overline{R}-\mathrm{H}(\boldsymbol{\tau}_\sharp^2\boldsymbol{\omega}),\,\overline{R}-\mathrm{H}(\boldsymbol{\omega})\right\} \\ & \leq & \max\left\{\frac{1}{3}\left(\overline{R}-\mathrm{H}\left(\boldsymbol{\omega}\right)\right),\,\frac{2}{3}\left(\overline{R}-\mathrm{H}\left(\boldsymbol{\omega}\right)\right),\overline{R}-\mathrm{H}\left(\boldsymbol{\omega}\right)\right\} \\ & \leq & -\frac{1}{3}\varepsilon\,,\end{array}$$

Thus the series $\sum_n \mathbb{P}(S_n(\boldsymbol{\omega}_n)=0)$ is convergent, and the Borel-Cantelli lemma implies the claim. ■

Finally, observe that, from Proposition 7 and the continuity of $\gamma_8(R)$, it follows that $\limsup \frac{1}{n}\,\mathrm{d}_{\min}(\mathcal{G}_n^R) \leq \gamma_8(R)$ with probability one, thus completing the proof of Theorem 1.

C Proofs for Section 4

C.1 Proof of Proposition 2

Proposition 2 *For every design rate $0 < R < \log 8$ and $0 < \varepsilon < \overline{R}$, with probability one*

$$\liminf_n \frac{1}{n} d_{\min}\left(\mathcal{B}_n^R\right) \geq \underline{\gamma}_\eta(R+\varepsilon)\,.$$

Proof For $\boldsymbol{\theta} \in \Theta$, consider the events $K_{\boldsymbol{\theta},n} := \left\{U_n^R(\boldsymbol{\theta}) \geq 1\right\}$ and $H_{\boldsymbol{\theta},n} := \{V_n^R(\pi_\sharp^1\boldsymbol{\theta}) \geq 1\}$. Observe that $K_{\boldsymbol{\theta},n}$ implies the existence of a pair $(\boldsymbol{x},\boldsymbol{y})$ such that $\Psi_n^R\boldsymbol{x} = \boldsymbol{0}$, so that

$$K_{\boldsymbol{\theta},n} \subseteq H_{\boldsymbol{\theta},n}\,. \tag{64}$$

From Lemma 4, using (2), we have that, for every joint type $\boldsymbol{\theta}$ in $\tilde{\Theta}_n := \Theta_n \setminus \Theta_{(\overline{R}-\varepsilon)}$, at least one of the two following inequalities holds true:

$$\mathbb{E}\left[U_n^R(\boldsymbol{\theta})\right] = \binom{n}{n\boldsymbol{\theta}}\frac{1}{8^{2l}} \leq \exp\left(n\left(\mathrm{H}(\boldsymbol{\theta})-2\overline{R}\right)\right) \leq \exp(-2n\varepsilon)\,, \tag{65}$$

$$\mathbb{E}\left[V_n^R\left(\pi_\sharp^1\boldsymbol{\theta}\right)\right] = \binom{n}{n\pi_\sharp^1\boldsymbol{\theta}}\frac{1}{8^l} \leq \exp\left(n\left(\mathrm{H}\left(\pi_\sharp^1\boldsymbol{\theta}\right)-\overline{R}\right)\right) \leq \exp(-n\varepsilon)\,. \tag{66}$$

Using a union bound estimation for the event $K_n := \bigcup_{\boldsymbol{\theta}\in\tilde{\Theta}_n} K_{\boldsymbol{\theta},n}$ and applying (64) and Markov's inequality, we have

$$\begin{array}{rcl}\mathbb{P}(K_n) & \leq & \mathbb{P}\left(\bigcup_{\boldsymbol{\theta}\in\tilde{\Theta}_n} K_{\boldsymbol{\theta},n}\cap B_{\boldsymbol{\theta},n}\right) \\ & \leq & \sum_{\boldsymbol{\theta}\in\tilde{\Theta}_n}\mathbb{P}\left(K_{\boldsymbol{\theta},n}\cap B_{\boldsymbol{\theta},n}\right) \\ & \leq & \sum_{\boldsymbol{\theta}\in\tilde{\Theta}_n}\min\left\{\mathbb{P}\left(K_{\boldsymbol{\theta},n}\right)\,,\,\mathbb{P}\left(B_{\boldsymbol{\theta},n}\right)\right\} \\ & \leq & \sum_{\boldsymbol{\theta}\in\tilde{\Theta}_n}\min\left\{\mathbb{E}\left[U_n(\boldsymbol{\theta})\right]\,,\,\mathbb{E}\left[V_n^R(\boldsymbol{\theta})\right]\right\} \\ & \leq & |\Omega_n|\exp(-n\varepsilon)\,.\end{array}$$

Thus the series $\sum_{n\geq 1}\mathbb{P}(K_n) \leq \sum_{n\geq 1}|\Omega_n|\exp(-n\varepsilon)$ is convergent, and, by Borel-Cantelli lemma, $\mathbb{P}\left(K_n\,\mathrm{i.\,o.}\right) = 0$, which, in turn, implies the claim. ■

C.2 Proof of Lemma 5

Lemma 5 *For all $n \geq 1$, and every joint type $\boldsymbol{\theta} \in \Theta_n$ such that $\pi^1_\sharp \boldsymbol{\theta} \neq \delta_0$,*

$$\mathrm{Var}\left[U^R_n(\boldsymbol{\theta})\right] \leq \binom{n}{n\boldsymbol{\theta}}\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}}\frac{16}{8^{3l}} + \binom{n}{n\boldsymbol{\theta}}^2\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}}^{-1}\frac{1}{8^{3l}} + \binom{n}{n\boldsymbol{\theta}}\frac{8}{8^{2l}}\,.$$

Proof We have

$$\begin{aligned}\mathrm{Var}\left[U^R_n(\boldsymbol{\theta})\right] &= \mathrm{Var}\left[\textstyle\sum_{(\boldsymbol{x},\boldsymbol{y})} \mathbb{1}_{\{\Psi^R_n \boldsymbol{x}=\boldsymbol{0}\}}\mathbb{1}_{\{\Psi^R_n \boldsymbol{y}=\boldsymbol{Z}_n\}}\right] \\ &= \sum_{(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)} c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big)\,,\end{aligned}$$

where the summations are extended to all pairs $(\boldsymbol{x},\boldsymbol{y})$, $(\boldsymbol{x}_1,\boldsymbol{y}_1)$ and $(\boldsymbol{x}_2,\boldsymbol{y}_2)$ in $(\mathbb{Z}_2^3\times\mathbb{Z}_2^3)^n_{\boldsymbol{\theta}}$, and

$$c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big) := \mathrm{Cov}\left[\mathbb{1}_{\{\Psi^R_n \boldsymbol{x}_1=\boldsymbol{0}\}}\mathbb{1}_{\{\Psi^R_n \boldsymbol{y}_1=\boldsymbol{Z}_n\}}\,,\ \mathbb{1}_{\{\Psi^R_n \boldsymbol{x}_2=\boldsymbol{0}\}}\mathbb{1}_{\{\Psi^R_n \boldsymbol{y}_2=\boldsymbol{Z}_n\}}\right].$$

We are now going to estimate the covariance terms $c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big)$, separately considering four possible different linear dependency structures among $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{y}_1$, and $\boldsymbol{y}_2$. Observe that, since $\pi^1_\sharp\boldsymbol{\theta}\neq\delta_0$, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ need to be nonzero in order for the pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1)$ and $(\boldsymbol{x}_2,\boldsymbol{y}_2)$ to have type $\boldsymbol{\theta}$.

Suppose first that $(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)$ in $\big(\mathbb{Z}_2^3\times\mathbb{Z}_2^3\big)^n_{\boldsymbol{\theta}}$ are such that $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are linear independent. Then, the r.v.s $\Psi^R_n\boldsymbol{x}_1$, $\Psi^R_n\boldsymbol{x}_2$, $\Psi^R_n\boldsymbol{y}_1$ and $\Psi^R_n\boldsymbol{y}_2$ are independent, so that

$$c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big) = 0\,.$$

Second, consider the case when $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are linear independent but $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are not linear independent. In this case we have that the random variables $\Psi^R_n\boldsymbol{x}_1$, $\Psi^R_n\boldsymbol{x}_2$ and $\Psi^R_n\boldsymbol{y}_1-\boldsymbol{Z}_n$ are independent, so that

$$c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big) \leq \mathbb{P}\left(\Psi^R_n\boldsymbol{x}_1=\boldsymbol{0},\Psi^R_n\boldsymbol{x}_2=\boldsymbol{0},\Psi^R_n\boldsymbol{y}_2=\boldsymbol{Z}_n\right) = \frac{1}{8^{3l}}\,.$$

Since there are at most $16\binom{n}{n\boldsymbol{\theta}}\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}}$ possible choices of such pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)$ in $\big(\mathbb{Z}_2^3\times\mathbb{Z}_2^3\big)^n_{\boldsymbol{\theta}}$, they contribute to the first addend in the righthand side of (34).

As a third case, consider pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)$, such that $\boldsymbol{x}_1=\boldsymbol{x}_2$, and $\boldsymbol{x}_1$, $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are linear independent. In this situation the random variables $\Psi^R_n\boldsymbol{x}_1$, $\Psi^R_n\boldsymbol{y}_1$ and $\Psi^R_n\boldsymbol{y}_2$ are independent so that

$$c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big) \leq \mathbb{P}\left(\Psi^R_n\boldsymbol{x}_1=\boldsymbol{0},\Psi^R_n\boldsymbol{y}_1=\boldsymbol{Z}_n,\Psi^R_n\boldsymbol{y}_2=\boldsymbol{Z}_n\right) = \frac{1}{8^{3l}}\,.$$

Since there are at most $\binom{n}{n\boldsymbol{\theta}}^2\binom{n}{n\pi^1_\sharp\boldsymbol{\theta}}^{-1}$ possible choices of such pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)$ in $(\mathbb{Z}_2^3\times\mathbb{Z}_2^3)^n_{\boldsymbol{\theta}}$, they contribute to the second addend in the righthand side of (34).

Finally, it remains to be considered the case when $\boldsymbol{x}_1=\boldsymbol{x}_2$, and $\boldsymbol{x}_1$, $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are linear dependent. There are at most $\binom{n}{n\boldsymbol{\theta}}8$ possible choices of pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1)$ and $(\boldsymbol{x}_2,\boldsymbol{y}_2)$ in $\big(\mathbb{Z}_2^3\times\mathbb{Z}_2^3\big)^n_{\boldsymbol{\theta}}$ satisfying these requirements and for each of them

$$c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big) \leq \mathbb{P}\left(\Psi^R_n\boldsymbol{x}_1=\boldsymbol{0},\Psi^R_n\boldsymbol{y}_1=\boldsymbol{Z}_n\right) = \frac{1}{8^{2l}}\,.$$

Therefore, they contribute to the third addend in the righthand side of (34). ■

C.3 Proof of Lemma 6

Lemma 6 *For all* $0 < R < \log 8$

$$\boldsymbol{\theta}_{(R)}(x,z) > 0\,, \qquad \boldsymbol{\upsilon}_{(R)}(x) > 0\,, \qquad \forall\, x,\, z \in \mathbb{Z}_2^3\,, \tag{67}$$

$$\langle \boldsymbol{\theta}_{(R)}, \boldsymbol{D}_\eta \rangle = \gamma_8(R)\,. \tag{68}$$

$$\mathrm{H}(\boldsymbol{\theta}_{(R)}) = \log 8 + \overline{R}\,. \tag{69}$$

$$\mathrm{H}(\boldsymbol{\upsilon}_{(R)}) > \overline{R}\,. \tag{70}$$

Proof (67) follows immediately from (35).

It is easy to verify that

$$\begin{aligned} \boldsymbol{D}_\eta(x,z) &= \boldsymbol{D}(\eta(z), \eta(x+z)) \\ &= \boldsymbol{d}\left(\mu^{-1}(\eta(x+z)) - \mu^{-1}(\eta(z))\right) \\ &= \boldsymbol{d}(\sigma_z(x))\,. \end{aligned}$$

Then (68) follows, since

$$\begin{aligned} \langle \boldsymbol{\theta}_{(R)}, \boldsymbol{D}_\eta \rangle &= \sum_{x,z\in\mathbb{Z}_2^3} \boldsymbol{\theta}_{(R)}(x,z)\boldsymbol{D}_\eta(x,z) \\ &= \sum_{z\in\mathbb{Z}_2^3} \tfrac{1}{8}\boldsymbol{\omega}_{(R)}\left(\sigma_z(x)\right)\boldsymbol{d}\left(\sigma_z(x)\right) \\ &= \langle \boldsymbol{\omega}_{(R)}, \boldsymbol{d} \rangle \\ &= \gamma_8(R)\,. \end{aligned}$$

From (36) we have $\sum_x \boldsymbol{\theta}_{(R)}(x,z) = \frac{1}{8}\sum_x \boldsymbol{\omega}_{(R)}(\sigma_z(x)) = \frac{1}{8}$, so that the marginal $\pi_\sharp^2\boldsymbol{\theta}_{(R)}$ is the uniform measure over $\mathbb{Z}_2^3$. Again from (36) we have that the conditioned measures satisfy $\boldsymbol{\theta}_{(R)}|_{\mathbb{Z}_2^3\times\{z\}} = \boldsymbol{\omega}_{(R)} \circ \sigma_z$ for every z in $\mathbb{Z}_2^3$. Then, by applying (1) we have

$$\begin{aligned} \mathrm{H}(\boldsymbol{\theta}_{(R)}) &= \mathrm{H}\left(\pi_\sharp^2\boldsymbol{\theta}_{(R)}\right) + \sum_{x\in\mathbb{Z}_2^3} \boldsymbol{\theta}_{(R)}\left(\mathbb{Z}_2^3 \times \{x\}\right)\mathrm{H}\left(\boldsymbol{\theta}_{(R)}|_{\mathbb{Z}_2^3\times\{x\}}\right) \\ &= \log 8 + \mathrm{H}(\boldsymbol{\omega}_{(R)}) \\ &= \log 8 + \overline{R}\,, \end{aligned}$$

showing (69).

Finally, observe that $\boldsymbol{\upsilon}_{(R)} = \pi_\sharp^1\boldsymbol{\theta}_{(R)} = \frac{1}{8}\sum_x \boldsymbol{\omega}_{(R)} \circ \sigma_x$ is a convex combination of permutations of the vector $\boldsymbol{\omega}_{(R)}$. As argued in Sect.2.2, for every labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$ there exists at least a pair of nonequal columns of the matrix $\boldsymbol{D}_\eta$, say $\boldsymbol{D}_\eta(\cdot, z_1) \neq \boldsymbol{D}_\eta(\cdot, z_2)$. As a consequence, we have $\boldsymbol{d} \circ \sigma_{z_1} \neq \boldsymbol{d} \circ \sigma_{z_2}$ which, together with (35), implies $\boldsymbol{\omega}_{(R)} \circ \sigma_{z_1} \neq \boldsymbol{\omega}_{(R)} \circ \sigma_{z_2}$. Hence, from the strict concavity and the permutation invariance of the entropy function H it follows that

$$\begin{aligned} \mathrm{H}(\boldsymbol{\upsilon}_{(R)}) &= \mathrm{H}\left(\tfrac{1}{8}\textstyle\sum_{x\in\mathbb{Z}_2^3} \boldsymbol{\omega}_{(R)} \circ \sigma_x\right) \\ &> \tfrac{1}{8}\textstyle\sum_{x\in\mathbb{Z}_2^3} \mathrm{H}\left(\boldsymbol{\omega}_{(R)} \circ \sigma_x\right) \\ &= \mathrm{H}(\boldsymbol{\omega}_{(R)}) \\ &= \overline{R}\,, \end{aligned}$$

showing (70). ■

References

[1] N. Alon and J.H. Spencer *The probabilistic method*, 3rd edition, Wiley, Hoboken, NJ, 2008.

[2] A. Barg and G.D. Forney, Jr., "Random codes: minimum distances and error exponents", *IEEE Trans. Inf. Theory*, vol. 48, pp. 2568-2573, 2001.

[3] A. Bennatan and D. Burshtein, "On the application of LDPC codes to arbitrary discrete memoryless channels", *IEEE Trans. Inf. Theory*, vol. 50, pp. 417-438, 2004.

[4] A. Bennatan and D. Burshtein, "Design and analysis of nonbinary LDPC codes for arbitrary discrete memoryless channels", *IEEE Trans. Inf. Theory*, vol. 52, pp. 549-583, 2006.

[5] R. Blahut, "Composition bounds for channel block codes", *IEEE Trans. Inf. Theory*, vol. 23, 656-674, 1977.

[6] V.S. Borkar, *Probability theory: an advanced course*, Springer, New York, 1995.

[7] G. Caire and E. Biglieri, "Linear block codes over cyclic groups", *IEEE Trans. Inf. Theory*, vol. 41, pp. 1246-1256, 1995.

[8] G. Como and F. Fagnani, "The capacity of finite Abelian group codes over memoryless symmetric channels", *IEEE Trans. Inf. Theory*, submitted, 2005, [online] av. at http://calvino.polito.it/∼fagnani/groupcodes/capacitygroupcodes.pdf..

[9] G. Como and F. Fagnani, "Average spectra and minim distances of low-density parity-check codes over Abelian groups", *SIAM J. Discr. Math.*, accepted, 2008.

[10] G. Como and F. Fagnani, "On the Gilbert-Varshamov distance of Abelian group codes", in Proc. of IEEE ISIT 2007, Nice 26-30 June 2007, pp. 2651-2655.

[11] I. Csiszàr and J. Korner, *Information theory: coding theorems for memoryless systems*, Academic press, New York, 1981.

[12] I. Csiszàr, "The method of types", *IEEE Trans. Inf. Theory*, vol. 44, pp. 2505-2523, 1998.

[13] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Jones and Bartlett, Boston, 1993.

[14] R. L. Dobrushin, "Asymptotic optimality of group and systematic codes for some channels", *Theor. Probab. Appl.*, vol. 8, pp. 47-59, 1963.

[15] F. Fagnani and S. Zampieri, "Minimal syndrome formers for group codes", *IEEE Trans. Inf. Theory*, vol. 45, pp. 1-31, 1998.

[16] G.D. Forney, Jr., "Geometrically Uniform Codes", *IEEE Trans. Inf. Theory*, vol. 37, pp. 1241-1260, 1991.

[17] G.D. Forney, Jr. and M. D. Trott, "The dynamics of group codes: dual Abelian group codes and systems", *IEEE Trans. Inf. Theory*, vol. 50, pp. 2935-2965, 2004.

[18] R.G. Gallager, *Low density parity check codes*, MIT Press, Cambridge MA, 1963.

[19] R.G. Gallager, *Information theory and reliable communication*, Wiley, New York, 1968.

[20] R.G. Gallager, “The random coding bound is tight for the average code”, *IEEE Trans. Inf. Theory*, vol. 19, pp. 244-246, 1973.

[21] R. Garello, G. Montorsi, S. Benedetto, D. Divsalar and F. Pollara, “Labelings and encoders with the uniform bit error property with applications to serially concatenated trellis codes”, *IEEE Trans. Inf. Theory*, vol. 48, pp. 123-136, 2002.

[22] F. Garin and F. Fagnani, “Analysis of serial turbo codes over Abelian groups for symmetric channels”, *SIAM J. Discr. Math.*, 2008.

[23] E.N. Gilbert, “A comparison of signalling alphabets”, *Bell Syst. Tech. J.*, vol. 31, pp. 504-522, 1952.

[24] E. Hof, I. Sason and S. Shamai, “Gallager-type bounds for non-binary linear block codes over memoryless symmetric channels”, *IEEE Trans. Inf. Theory*, submitted, 2008.

[25] J. Hou, P.H. Siegel, L.B. Milstein and H.D. Pfister, “Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes”, *IEEE Trans. Inf. Theory*, vol. 49, pp. 2141-2155, 2003.

[26] T.W. Hungerford, *Algebra*, Springer Verlag, New York, 1974.

[27] J. Korner and K. Marton, “How to encode the modulo-two sum of binary sources”, IEEE Trans. Inf. Theory, vol. 25, pp. 219221, March 1979.

[28] J. C. Interlando, R. Palazzo and M. Elia, “Group bloack codes over non-Abelian groups are asymptotically bad”, *IEEE Trans. Inf. Theory*, vol. 42, pp. 1277-1280, 1996.

[29] H.-A. Loeliger, “Signal sets matched to groups”, *IEEE Trans. Inf. Theory*, vol. 37, pp. 1675-1679, 1991.

[30] J. K. Omura, “On general Gilbert bounds”, *IEEE Trans. Inf. Theory*, vol. 19, pp. 661-666, 1973.

[31] J. K. Omura, “Expurgated bounds, Bhattacharyya distance and rate distortion functions”, *Information and Control*, vol. 24, pp. 358-383, 1974.

[32] B. Nazer and M. Gatspar, “The case for structured random codes in network capacity theorems”, *European Trans. on Telecommunication*, vol. 19, pp. 455-474, 2008.

[33] T. Richardson, R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[34] D. Slepian, “Group codes for the Gaussian channel”, *Bell System Tech. J.*, vol. 47, pp. 575-602, 1968.

[35] D. Sridhara and T.E. Fuja, LDPC codes over rings for PSK modulation, *IEEE Trans. Inf. Theory*, vol. 51, pp. 3209-3220, 2005.

[36] N. Shulman and M. Feder, “Random coding techniques for nonrandom codes”, *IEEE Trans. Inf. Theory*, vol. 45, pp. 2001-2004, 1999.

[37] R.R. Varshamov, “Estimate of the number of signals in error correcting codes”, *Dokl. Acad. Nauk.*, vol. 117, pp. 739-741, 1957.

[38] A. J. Viterbi and J. Omura, *Principles of digital communication and coding*, McGraw-Hill, New York, 1979.