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

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

Giacomo Como

***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 typical group code over the the cyclic group of order eight achieves both the Gilbert-Varshamov bound and the expurgated error exponent. On the other hand, the typical binary-coset codes (under any labeling) is shown to be bounded away 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 the cyclic group of order eight, but not the direct product of three copies of the binary group.**

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 contrasts 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.

I. Introduction

As low-complexity modern coding has emerged, based on random constructions of linear codes with sparse graphical representation [34], the analysis of random codes with algebraic structure has recently attracted renewed attention from the research community [2], [30]. 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 one 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 [26]. Confining 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 [2].

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 [25]. 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 (GCE) over the cyclic group of order eight, $\mathbb{Z}_8$, and the ensemble of binary-coset codes (BCE), respectively (see Sect. II-A 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 [22], [38] is one of the most well known and fundamental results of coding theory. Given a rate R in $(0,1)$, and defined $\delta_2(R)$ as the unique solution in $(0,1/2)$ of the equation $\mathrm{H}(x) = 1 - R$ (where $\mathrm{H}(x)$ denotes the binary entropy), it states that for every $n \geq 1$ there exist binary codes of block-length n, rate at least R, and minimum Hamming distance at least $n\delta_2(R)$. [1] Its asymptotic tightness is still considered one of longest-standing unproved conjectures in coding theory [23], [37]. [2] 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 [18], [31], [32], [5], [39]. Although both the classical GV bound and expurgated bound are mere existence results, for binary symmetric memoryless channels it is known that the typical binary-linear code achieves both the GV bound and the expurgated exponent [17], [33], [2]. It is also known that the same does not hold true [2] for the typical random code, whose performance is bounded away from the GV bound, as well as (at low rates) from the expurgated error

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

[1] More precisely, using a basic sphere-covering argument, Gilbert [22] proved that for every positive integers n and d, there exist binary codes of block-length n, minimum Hamming distance d, and cardinality not smaller than $2^n / \sum_{0 \leq k < d} \binom{n}{k}$. Varshamov [38] improved on this bound, for finite lengths. Together with the upper bound on the volume of a discrete sphere $\sum_{0 \leq k < d} \binom{n}{k} < 2^{n\,\mathrm{H}(d/n)}$, their results imply the stated bound.

[2] Here, tightness is meant up to factors sublinear in n, whereas improvements on such factors is an active field of research, see e.g. [28].

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 (7) and (13)). 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 the typical group code achieves the GV bound (Theorem 1), while the typical binary-coset code is bounded away from it (Theorem 2). Similarly, the typical group code achieves the expurgated exponent (Corollary 1), while the typical binary-coset code does not (Corollary 2).

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 $\mathbb{Z}_2^3$, the direct product of three copies of the binary group, 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.

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 [18, pagg.206-209]. Recent works [24], [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 Corollary 2 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. To the best of the author's knowledge, such a limitation of the performance of binary-coset codes had not been proved before.

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 [35], [15], [29], [7], [14], [16] and references therein. It is well known [13] 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 one 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 Corollary 1 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], [36], [20], [9], [21]. It was empirically observed in [36] that LDPC codes over $\mathbb{Z}_8$ perform better than their binary-coset counterparts on the 8-PSK AWGNC: the results of the present paper point out to an analytical explanation for this phenomenon.

We observe that, in spite of the fact that the cyclic group $\mathbb{Z}_8$ matches the 8-PSK constellation, the average error exponent of the GCE has been shown [8] 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 [13]). 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 explained by the fact that the average error exponent (an *annealed* average in the statistical physics language [30, Ch. 5.4]), provides only a lower bound to the typical error exponent (a *quenched* quantity), 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. In fact, at low rates, the error probability of the average group code is dominated by the error probability of its binary subcode, i.e. the set of its codewords whose entries belong to the binary subgroup $4\mathbb{Z}_8$ [8]. Therefore, the error exponent of the average group code coincides with the random coding exponent of the binary-input channel obtained by restricting the input from the whole 8-PSK to a pair of its opposite elements. This is strictly smaller than the random coding exponent of the 8–PSK AWGNC, which is achieved by the uniform distribution over the whole 8-PSK constellation. On the other hand, at low rates, the typical error event is made between the two closest codewords in the code, and the error exponent coincides with the minimum distance. As it will be shown in the present paper, the typical group code has larger minimum distance than the typical binary-coset code, hence it also has better error exponent.

The remainder of the paper is organized as follows. In Sect. II, we formally introduce the GCE and the BCE (Sect. II-A), and state the main results of the paper (Sect.s II-B and II-C). In Sect. III 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. IV by applying the second-moment method (Sect. IV-A) and some convex optimization techniques (Sect. IV-B). Finally, Sect. V presents some concluding remarks and points out to generalizations of the results to

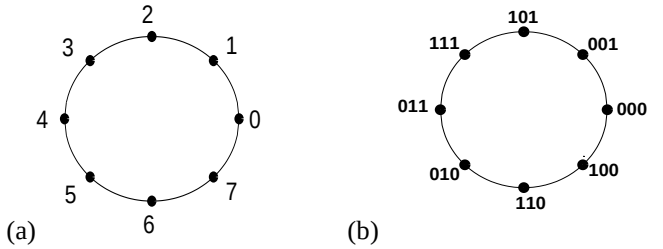

Fig. 1. The 8-PSK constellation with: (a) the isometric labeling $\mu : \mathbb{Z}_8 \to \mathcal{X}$; (b) a binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$. The latter is a so-called Gray labeling: neighbor signals are assigned labels differing in one digit only.

balanced DMCs. Sect. A is of a technical nature and discusses some continuity issues.

Before proceeding, let us establish some basic notation. The i-th entry of a vector $\boldsymbol{x}$ will be denoted by x_i. The scalar product of two functions $f, g : \mathcal{A} \to \mathbb{R}$, where $\mathcal{A}$ is some finite alphabet, and $\mathbb{R}$ is the set of real numbers, will be denoted by $\langle f, g \rangle := \sum_i f(i)g(i)$. Throughout, $\log$ will denote the logarithm in base 2, and $\mathrm{H}(\theta) := -\sum_i \theta(i) \log \theta(i)$ will denote the binary entropy of a probability distribution θ. With a slight, and common, abuse of notation, for $x \in [0, 1]$, $\mathrm{H}(x)$ will denote the entropy of a Bernoulli distribution with parameter x. Finally, $\mathbb{1}_A$ will denote the indicator function of a set A.

II. Problem statement and main results

A. Two capacity-achieving code ensembles for the 8-PSK Gaussian channel

We shall consider transmission over a memoryless AWGNC with input constrained on the 8-PSK signal constellation $\mathcal{X} := \{e^{i\frac{2\pi}{8}k} : 0 \leq k < 8\}$ and output space $\mathcal{Y} = \mathbb{R}^2$. The Bhattacharyya distance function associated to the 8-PSK AWGNC is

$$D : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+, \qquad D(x_1, x_2) := \frac{\log e}{8\sigma^2} ||x_1 - x_2||^2, \quad (1)$$

σ^2 being the noise variance. The symmetry group Π, i.e. the subgroup of permutations of $\mathcal{X}$ leaving D invariant, is isomorphic to the dihedral group D_8 with 16 elements [15], [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. The constellation $\mathcal{X}$ is said to be geometrically uniform [15], meaning that for every $x_1, x_2 \in \mathcal{X}$ there exists $\pi \in \Pi$ such that $\pi(x_1) = x_2$. Moreover, the cyclic group $\mathbb{Z}_8$ is a generating group of $\mathcal{X}$ [29], i.e. Π has a subgroup G isomorphic to $\mathbb{Z}_8$ such that for all $x_1, x_2 \in \mathcal{X}$ there exists a unique $\pi \in \Pi$ such that $\pi(x_1) = x_2$. In particular, let $\mu(z) := e^{i\frac{2\pi}{8}z}$ be the standard isometric labeling, and consider the function $D_\mu(z_1, z_2) := D(\mu(z_1 + z_2), \mu(z_2))$. Then, all the columns $D_\mu(\cdot, z)$ coincide with the distance profile

$$d : \mathbb{Z}_8 \to \mathbb{R}, \qquad d(z) := D(\mu(0), \mu(z)). \quad (2)$$

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

$$D_\eta : \mathbb{Z}_2^3 \times \mathbb{Z}_2^3 \to \mathbb{R}^+, \qquad D_\eta(z_1, z_2) := D(\eta(z_2), \eta(z_2 + z_1)), \quad (3)$$

coincide.

We shall consider block-codes $\mathcal{C} \subseteq \mathcal{X}^n$, and denote their rate by $R(\mathcal{C}) := n^{-1} \log |\mathcal{C}|$, their minimum distance by $d_{\min}(\mathcal{C}) := \min\{\sum_{i=1}^n D(x_i, z_i) : \boldsymbol{x} \neq \boldsymbol{z} \in \mathcal{C}\}$, and their maximum-likelihood error probability by $p_e(\mathcal{C})$. 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. II-A, 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. $\mathrm{d}_{\min}(\mathcal{G}) = \min\{\sum_{1 \leq j \leq n} d(x_j) | \boldsymbol{x} \neq \boldsymbol{0} \in K\}$. Similarly, group codes are known to enjoy the uniform error property. 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 does 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 $R \in [0, 3]$, and a blocklength $n \geq 1$, set $\overline{R} := 3 - R$, and $l := \lfloor \overline{R} n / 3 \rfloor$. We shall consider the two following code ensembles:

Group code ensemble For $n \geq 1$, let Φ_n^R be a random matrix uniformly distributed over $\mathbb{Z}_8^{l \times n}$. Define the random group code

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

Binary coset ensemble Let $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$ be an arbitrary labeling.For $n \geq 1$, consider a random matrix Ψ_n^R, uniformly distributed over $\mathbb{Z}_2^{3l \times 3n}$, and $\boldsymbol{W}_n$ be an independent random vector, uniformly distributed over $\mathbb{Z}_2^{3n}$. Define the random binary-coset code as

$$\mathcal{B}_n^R := \eta_n \left(\ker \Psi_n^R + \boldsymbol{W}_n \right). \quad (4)$$

Throughout the paper, we shall say that the *typical group code* (respectively, the *typical binary-coset code*) satisfies a certain (in)equality if, for all $\varepsilon > 0$, the probability that $\mathcal{G}_n^R$ (resp. by $\mathcal{B}_n^R$) violates such (in)equality by more than ε vanishes as the block-length n grows large, and the design rate R is kept constant. We observe that the group code ensemble and the binary-coset ensemble are sometimes defined as images of random generating matrices rather than kernels of random syndrome matrices, as above. However, while leading to different properties for finite lengths, it can be shown that such alternative definitions do not alter the asymptotic properties of the typical group, and binary-coset, code.

An immediate consequence of the symmetry properties discussed in Sect. II-A 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 [18] $E_8^r(R)$. It is not hard to show that binary-coset codes achieve capacity and

$$\mathbb{E}\left[p_e(\mathcal{B}_n^R)\right] \leq 2^{-nE_8^r(R)}\,,$$

$$\lim_{n\to+\infty} -\frac{1}{n}\log \mathbb{E}\left[p_e(\mathcal{B}_n^R)\right] = E_8^r(R)\,.$$

In fact, the standard random coding averaging arguments of [18, pagg.206-207] as well as the tightness considerations of [19] apply, upon observing that, for every $\boldsymbol{z} \in \mathbb{Z}_2^{3n}$, the event $A_{\boldsymbol{z}} := \left\{\Psi_n^R \boldsymbol{z} = \Psi_n^R \boldsymbol{W}_n\right\}$ has probability 8^{-l}, and that $A_{\boldsymbol{z}_1}$, $A_{\boldsymbol{z}_2}$ and $A_{\boldsymbol{z}_3}$ are mutually independent for linearly independent $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3 \in \mathbb{Z}_2^{3n}$.

As far as group codes are concerned, 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 group codes achieve capacity, and that their average error probability can be upper-bounded by a term exponentially decreasing in the block-length n

$$\mathbb{E}\left[p_e(\mathcal{G}_n^R)\right] \leq 2^{-nE_{\mathbb{Z}_8}^r(R)}\,. \tag{5}$$

The exponent appearing in the righthand side of (5) is given by

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

with $E_4^r(\frac{2}{3}R)$ and $E_2^r(\frac{1}{3}R)$ denoting the random coding error exponents of the AWGNCs with input restricted over the 4-PSK and the 2-PSK constellation, respectively. As shown in [8], the bound (5) 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 error exponent of the average binary-coset code of design rate R (i.e. the exponential decay rate of $\mathbb{E}[p_e(\mathcal{B}_n^R)]$) coincides with the random coding exponent $E_8^r(R)$, while the error exponent of the average group code (i.e. the exponential decay rate of $\mathbb{E}[p_e(\mathcal{G}_n^R)]$) is strictly smaller than $E_8^r(R)$ for low R. 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 error exponent achievable of the average group code. 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. In contrast, the results stated in the two following subsections show that the typical group code outperforms the typical binary-coset code, thus reversing the hierarchies outlined by the average-code analysis.

B. Gilbert-Varshamov bound and typical minimum distances

Let Ω be the space of probability vectors over $\mathbb{Z}_8$, and, for $0 \leq R \leq 3$, define

$$\Omega_R := \left\{\omega \in \Omega : \mathrm{H}(\omega) \geq \overline{R}\right\} \tag{6}$$

$$\delta_8(R) := \min\left\{\langle \omega, d\rangle : \omega \in \Omega_R\right\}\,, \tag{7}$$

where d is the squared Euclidean weight function defined in (2). In Sect. A, $\delta_8(R)$ is proved to be continuous and non-increasing as a function of the rate R. The GV bound for the 8-PSK AWGNC [6, Th. 10.5.1] states that, for every $0 \leq R \leq 3$, and any $n \geq 1$, there exists a block code $\mathcal{C}_n$ of length n and rate not smaller than R.[3] While the aforementioned is a mere existence result, the question we want to address here is whether $\delta_8(R)$ is achieved by either the typical group code or the typical binary-coset code. In fact, using arguments analogous to those in [2], it is not difficult to see that the the typical random code sampled from the random coding ensemble does not achieve the GV bound. This is because the minimum distance of the RCE of design rate R turns out to be the minimum of the relative distance between all possible $\binom{\lceil 2^{Rn}\rceil}{2}$ choices of pairs of distinct codewords. Since the differences between such pairs of codewords are pairwise independent random variables, uniformly distributed over $\mathcal{X}^n$, the normalized minimum distance of the typical random code can be shown to coincide with $\delta_8(2R)$.

We shall therefore concentrate on the performance of the group coding ensemble, and the binary-coset ensemble. Here, the algebraic structure prevents the differences between different pairs of codewords to be pairwise independent, and this will be proven to lead to higher typical minimum distances. In particular, the following result concerns the GCE:

Theorem 1 (Minimum distance of the typical group code) *For all $0 \leq R \leq 3$, the normalized minimum distance of the typical group code of design rate R coincides with $\delta_8(R)$.*

Proof: See Sect. III 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 Θ be the set of joint probability vectors over $\mathbb{Z}_2^3 \times \mathbb{Z}_2^3$. For $0 \leq R \leq 3$, define the sets

$$\underline{\Theta}_R := \left\{\theta \in \Theta :\ \mathrm{H}(\theta) \geq 2\overline{R},\ \mathrm{H}(\upsilon) \geq \overline{R}\right\}\,, \tag{8}$$

$$\overline{\Theta}_R := \left\{\theta \in \Theta :\ \mathrm{H}(\theta) - \mathrm{H}(\upsilon) \geq \overline{R},\ \mathrm{H}(\upsilon) \geq \overline{R}\right\}\,, \tag{9}$$

where $\upsilon(\,\cdot\,) = \sum_z \theta(\,\cdot\,, z)$ is the first-component marginal of θ. Define the functions

$$\underline{\delta}_\eta(R) = \min\left\{\langle \theta, D_\eta\rangle :\ \theta \in \underline{\Theta}_R\right\}\,, \tag{10}$$

$$\overline{\delta}_\eta(R) := \min\left\{\langle \theta, D_\eta\rangle :\ \theta \in \overline{\Theta}_R\right\}\,, \tag{11}$$

Theorem 2 (Minimum distance of the typical binary-coset code) *For every $0 < R < 3$, the normalized minimum distance of the typical binary-coset code is lower-bounded by $\underline{\delta}_\eta(R)$ and upper-bounded by $\overline{\delta}_\eta(R)$. Furthermore,*

$$\underline{\delta}_\eta(R) \leq \overline{\delta}_\eta(R) < \delta_8(R)\,. \tag{12}$$

Proof: See Sect IV. ■

[3] The definition (7) can be shown to be equivalent to that of $E_L(R)$ defined in [6, pag. 399], upon observing that, in the case of the 8-PSK, the optimizing distribution in the definition of $E_L(R)$ has to be symmetric with respect to rotations.

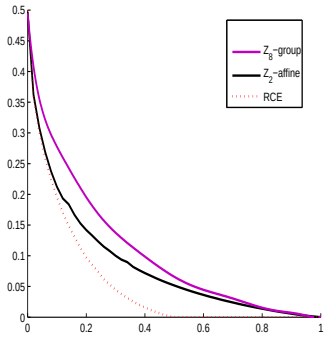

Fig. 2. A comparison of $\delta_8(R)$ (purple line) and $\overline{\delta}_\eta(R)$ where $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$ is the Gray labeling described in Fig.1. As a reference $\delta_8(2R)$ (which is the typical normalized minimum distance of the RCE) is plotted in dotted red. For the specific choice of the binary labeling $\eta : \mathbb{Z}_2^3 \to \mathcal{X}$, and the chosen resolution, it seems that $\overline{\delta}_\eta(R) = \underline{\delta}_\eta(R)$.

In Fig.2 the normalized minimum distances of the typical group code and of the typical binary-coset code are plotted as a function of the design rate R, together with the normalized minimum distance of the typical random code.

C. Expurgated bound and typical error exponents

For every rate $0 \le R \le 3$ the *expurgated exponent* of the 8-PSK AWGNC is

$$E_8^x(R) := \min\left\{\langle \omega, d\rangle + \overline{R} - \mathrm{H}(\omega) : \omega \in \Omega_R\right\}. \qquad (13)$$

The expurgated exponent $E_8^x(R)$ and the GV distance $\delta_8(R)$ coincide at small rates. Indeed, let $\omega^x := e^{-d}/\sum_z e^{-d(z)}$ be the minimizer of the map $\omega \mapsto \langle \omega, d\rangle + \overline{R} - \mathrm{H}(\omega)$ over the whole type space Ω, $R_8^x := \mathrm{H}(\omega^x) > 0$ be the minimum rate R for which $\omega^x \in \Omega_R$, and $R_8^0 := \log \sum_z \frac{1}{8} e^{-d(z)}$ denote the so-called cut-off rate. We have that:

- for rates $R_8^x \le R \le R_8^0$, the minimum in (13) is achieved by ω^x, and $E_8^x(R) = R_8^0 - R$;
- for rates $0 \le R \le R_8^x$, Lemma 8 implies that the minimum in (13) is achieved by some type ω such that $\mathrm{H}(\omega) = \overline{R}$, so that

$$E_8^x(R) = \delta_8(R), \qquad \forall 0 \le R \le R_8^x. \qquad (14)$$

The expurgated bound (see [18, pagg.153-157], and [11, pagg.185-186,192-195]) guarantees, for all rates $0 < R < 3$, and $n \ge 1$, the existence of a code $\mathcal{C}_n \subseteq \mathcal{X}^n$ with rate not smaller than R, and error probability not exceeding $2^{-nE_8^x(R)}$. 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 typical random code. Therefore, we shall be concerned with the error exponents of the typical group code, and of the typical binary-coset code. The following results will be proven as consequences of Theorem 1, and Theorem 2, respectively.

Corollary 1 (Error exponent of the typical group code) *For every $0 < R < R_8^x$, the error exponent of the typical group code of design rate R coincides with $E_8^x(R)$.*

Proof: See Sect. III. ■

Corollary 2 (Error exponent of the typical binary-coset code) *There exists some $R_\eta^x > 0$ such that, for every $0 < R < R_\eta^x$, the error exponent of the typical binary-coset code of design rate R is strictly smaller than $E_8^x(R)$.*

Proof: See Sect. IV. ■

III. Performance of the typical group code

In this section we shall show that the typical group code has normalized minimum distance, and error exponent, bounded from below by the GV distance, and the expurgated exponent, respectively. The proof of tightness of these bounds will instead be given in Sect. B, thus completing the proof of Theorem 1. Throughout the section, Ω will denote the space of all probability vectors over $\mathbb{Z}_8$, $\Omega_n \subseteq \Omega$ will denote the set of all types (i.e. empirical frequencies, see [12]) of length-n strings with entries in $\mathbb{Z}_8$, and $(\mathbb{Z}_8)_\omega^n \subseteq \mathbb{Z}_8^n$ the set of length-n strings of type ω.

We shall apply the first-moment method [1, Ch. 2] to the type-enumerator function

$$G_n^R(\omega) := \left|(\mathbb{Z}_8)_\omega^n \cap \ker \Phi_n^R\right|,$$

counting the number of codewords of type ω in the random group code of rate R and length n. As a first step in our analysis, we evaluate the expected value $\mathbb{E}[G_n^R(\omega)]$. It will prove convenient to denote by $2^{\zeta(\omega)}$ the order of the smallest subgroup of $\mathbb{Z}_8$ supporting ω.

We have the following result:

Lemma 1 *For every design rate $0 < R < 3$ and $\mathbb{Z}_8$-type ω in $\mathcal{P}_n(\mathbb{Z}_8)$ such that $\omega(0) < 1$,*

$$\mathbb{E}\left[G_n^R(\omega)\right] \le 2^{n\left(\mathrm{H}(\omega) - \frac{\overline{R}}{3}\zeta(\omega)\right)}.$$

Proof: Let $\boldsymbol{x}$ be an n-tuple of type ω, and let $h := 2^{3-\zeta(\omega)}$ be the largest power of two dividing all the nonzero entries of $\boldsymbol{x}$. Then, every entry x_i belongs to $h\mathbb{Z}_8$, and there exists some $1 \le i^* \le n$ such that x_{i^*} is not divisible by $2h$. For $1 \le i \le n$, let us denote by $\boldsymbol{Y}_i$ the i-th column of Φ_n^R, which is a r.v. uniformly distributed over $\mathbb{Z}_8^l$. Then, one has that $H := x_{i^*}\boldsymbol{Y}_{i^*}$, and that $K := \sum_{i \ne i^*} x_i \boldsymbol{Y}_i$ takes values in $h\mathbb{Z}_8^l$. It follows that

$$\begin{aligned}\mathbb{P}(\Phi_n^R \boldsymbol{x} = \boldsymbol{0}) &= \textstyle\sum_{\boldsymbol{k} \in h\mathbb{Z}_8^l} \mathbb{P}\left(H = -\boldsymbol{k}, K = \boldsymbol{k}\right) \\ &= \textstyle\sum_{\boldsymbol{k} \in h\mathbb{Z}_8^l} 2^{-l\zeta(\omega)} \mathbb{P}\left(K = \boldsymbol{k} | H = -\boldsymbol{k}\right) \\ &= 2^{-l\zeta(\omega)}.\end{aligned}$$

Now, observe that $\mathbb{E}\left[G_n^R(\omega)\right] = |(\mathbb{Z}_8)_\omega^n|\,\mathbb{P}\left(\Phi_n^R \boldsymbol{x} = \boldsymbol{0}\right)$. Then, the claim follows from the standard estimation for the binomial $|(\mathbb{Z}_8)_\omega^n| = \binom{n}{n\omega} \le 2^{n\,\mathrm{H}(\theta)}$ (see e.g. [12]). ■

For $0 \le R \le 3$, consider the sets

$$\begin{aligned} \Omega''_R &:= \left\{\omega : \omega(j) = 0,\ \forall j \notin 2\mathbb{Z}_8,\ \mathrm{H}(\omega) \ge \tfrac{2}{3}\overline{R}\right\}, \\ \Omega'_R &:= \left\{\omega : \omega(j) = 0,\ \forall j \notin 4\mathbb{Z}_8,\ \mathrm{H}(\omega) \ge \tfrac{1}{3}\overline{R}\right\}. \end{aligned} \tag{15}$$

Let

$$\begin{aligned} \delta_4\left(\tfrac{2}{3}R\right) &:= \min\{\langle\omega, d\rangle : \omega \in \Omega''_R\} \\ \delta_2\left(\tfrac{1}{3}R\right) &:= \min\{\langle\omega, d\rangle : \omega \in \Omega'_R\} \end{aligned}$$

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

$$\delta_{\mathbb{Z}_8} := \min\left\{\delta_8(R), \delta_4\left(\tfrac{2}{3}R\right), \delta_2\left(\tfrac{1}{3}R\right)\right\}.$$

For $R^* > R$, and a blocklength n, consider the event

$$F := \bigcup_{\omega}\left\{G_n^R(\omega) \ge 1\right\} \cap \left\{\mathrm{H}(\omega) < \tfrac{\overline{R}^*}{3}\zeta(\omega)\right\}.$$

Observe that, since the set $\left\{\omega : \mathrm{H}(\omega) \ge \overline{R}^*\zeta(\omega)/3\right\}$ is contained in the union $\Omega_{R^*} \cup \Omega''_{R^*} \cup \Omega'_{R^*}$, one has that the inequality $\mathrm{d}_{\min}(\mathcal{G}_n^R) \ge n\delta_{\mathbb{Z}_8}(R^*)$ holds whenever F does not occur. Then, by subsequently using the union bound, Markov's inequality, and Lemma 1, one has

$$\begin{aligned} \mathbb{P}\left(\mathrm{d}_{\min}(\mathcal{G}_n^R) \ge \delta_{\mathbb{Z}_8}(R^*)\right) &\ge 1 - \mathbb{P}(F) \\ &\ge 1 - \textstyle\sum_\omega \mathbb{E}\left[G_n^R(\omega)\right] \\ &\ge 1 - \textstyle\sum_\omega 2^{n\left(\mathrm{H}(\omega) - \frac{\overline{R}}{3}\zeta(\omega)\right)} \\ &\ge 1 - |\Omega_n| 2^{-n(R^* - R)} \\ &\xrightarrow{n\to+\infty} 1, \end{aligned} \tag{16}$$

the last step following from the fact that the number of $\mathbb{Z}_8$-types, $|\Omega_n| = \binom{n+7}{7}$, grows only polynomially fast with n (see e.g. [12]). From the continuity of $\delta_{\mathbb{Z}_8}(R)$, and the arbitrariness of $R^* > R$, it thus follows that the typical group code has normalized minimum distance not smaller than $\delta_{\mathbb{Z}_8}(R)$.

Clearly, $\delta_{\mathbb{Z}_8}(R) \le \delta_8(R)$. We shall now prove that, in fact, the equality holds. Observe that our arguments have relied only on the algebraic structure of the group $\mathbb{Z}_8$, while the geometric properties of the 8-PSK constellation have not played any role so far. In fact, counterexamples can be constructed as in [8] showing that Lemma 2 fails to hold true for other DMCs with the same symmetry structure of the 8-PSK AWGNC. The geometry of the 8-PSK constellation allows us to prove the following result:

Lemma 2 *For every design rate $0 \le R \le 3$,*

$$\delta_8(R) = \delta_{\mathbb{Z}_8}(R).$$

Proof: For $R = 3$, trivially $\delta_2(\frac{1}{3}R) = \delta_4(\frac{2}{3}R) = \delta_8(R) = 0$, and then $\delta_8(R) = \delta_{\mathbb{Z}_8}(R) = 0$.

Now, let us assume that $R < 3$. Since the entropy function is concave and the unique minimum of the map $\omega \mapsto \langle\omega, d\rangle$ on Ω is achieved with $\omega(0) = 1$, we can apply Lemma 8 and claim that a minimizer $\omega \in \Omega_R$ in the definition (7) of $\delta_8(R)$ necessarily satisfies $\mathrm{H}(\omega) = \overline{R}$. Then, using Lagrangian multipliers, we obtain

$$\delta_8(R) = Z_8(\lambda_8)^{-1} \sum_{x\in\mathbb{Z}_8} d(x) e^{-\lambda_8 d(x)},$$

where $Z_8(\lambda_8) := \sum_{x\in\mathbb{Z}_8} e^{-\lambda_8 d(x)}$, and $\lambda_8 > 0$ solves the equation $\mathrm{H}\left(Z_8(\lambda_8)e^{-\lambda_8 d}\right) = \overline{R}$. Analogously,

$$\delta_4\left(\tfrac{2}{3}R\right) = Z_4(\lambda_4)^{-1} \sum_{x\in 2\mathbb{Z}_8} e^{-\lambda_4 d(x)} d(x).$$

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

$$\delta_2\left(\tfrac{1}{3}R\right) = d(4)\alpha, \qquad \alpha := Z_2(\lambda_2)^{-1}e^{-\lambda_2 d(4)},$$

where $Z_2(\lambda_2) := 1 + e^{-\lambda_2 d(4)}$, and $\lambda_2 > 0$ solves $\mathrm{H}\left(Z_2(\lambda_2)^{-1}e^{-\lambda_2 d(4)}\right) = \frac{1}{3}\overline{R}$. Observe that $\alpha \in (0, 1/2)$.

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

$$d(4) = 2d(2) = 2d(6) \tag{17}$$

$$d(1) = d(7), \quad d(3) = d(5), \quad d(1) = d(4) - d(3) < \frac{d(4)}{4}. \tag{18}$$

It follows from (17) that

$$Z_4(2s) = \left(1 + e^{-sd(4)}\right)^2 = Z_2(s)^2,$$

for all $s \ge 0$. Then, (17) implies that

$$\begin{aligned} Z_4(2\lambda_2)^{-1}e^{-2\lambda_2 d(0)} &= Z_2(\lambda_2)^{-2} = (1-\alpha)^2, \\ Z_4(2\lambda_2)^{-1}e^{-2\lambda_2 d(2)} &= Z_4(2\lambda_2)^{-1}e^{-2\lambda_2 d(6)} = \alpha(1-\alpha), \\ Z_4(2\lambda_2)^{-1}e^{-2\lambda_2 d(4)} &= \alpha^2. \end{aligned}$$

Therefore,

$$\mathrm{H}\left(Z_4(2\lambda_2)^{-1}e^{-2\lambda_2 d|_{2\mathbb{Z}_8}}\right) = 2\mathrm{H}(\alpha) = 2\mathrm{H}\left(Z_2(\lambda_2)^{-1}e^{-\lambda_2 d|_{4\mathbb{Z}_8}}\right),$$

so that $2\lambda_2 = \lambda_4$. Hence,

$$\begin{aligned} \delta_4\left(\tfrac{2}{3}R\right) &= Z_4(\lambda_4)^{-1}\left\langle e^{-\lambda_4 d}\mathbb{1}_{2\mathbb{Z}_8}, d\right\rangle \\ &= \alpha^2 d(4) + 2\alpha(1-\alpha)d(2) \\ &= \alpha d(4) \\ &= Z_2(\lambda_4/2)^{-1} d(4) e^{-\frac{\lambda_4}{2}d(4)} \\ &= \delta_2(R/3). \end{aligned}$$

Since $\delta_8(R)$ is defined in (7) as the minimum of $\langle\omega, d\rangle$ over Ω_R, in order to estimate it from above it is sufficient to estimate $\langle\hat{\omega}, d\rangle$ for some $\hat{\omega} \in \Omega_R$. We do so for $\hat{\omega}$ defined by

$$\begin{aligned} \hat{\omega}(0) &:= (1-\alpha)^3, \qquad \hat{\omega}(1) := \hat{\omega}(2) := \hat{\omega}(7) := \alpha(1-\alpha)^2, \\ \hat{\omega}(4) &:= \alpha^3, \qquad \hat{\omega}(6) := \hat{\omega}(5) := \hat{\omega}(3) := \alpha^2(1-\alpha), \end{aligned} \tag{19}$$

It is straightforward to verify that $\mathrm{H}(\hat{\omega}) = 3\,\mathrm{H}(\alpha) = \overline{R}$, so that $\hat{\omega} \in \Omega_R$. Moreover, it follows from (17) and (18) that

$$\begin{aligned} \langle\hat{\omega}, d\rangle &= \textstyle\sum_x d(x)\hat{\omega}(x) \\ &= \alpha^3 d(4) + 2\alpha^2(1-\alpha)\left(d(4) - d(1)\right) \\ &\quad + \alpha(1-\alpha)\tfrac{1}{2}d(4) + 2\alpha(1-\alpha)^2 d(1) \\ &= 2\alpha d(1)\left(2\alpha^2 - 3\alpha + 1\right) - \tfrac{\alpha}{2}d(4)\left(2\alpha^2 - 3\alpha - 1\right) \\ &= \alpha d(4) + \alpha d(4)\left(2d(1) - \tfrac{1}{2}d(4)\right)\left(2\alpha^2 - 3\alpha + 1\right) \\ &< \alpha d(4), \end{aligned}$$

last inequality following from (18) and the fact that $2\alpha^2 - 3\alpha + 1 > 0$ for every $\alpha \in (0, 1/2)$. It follows that

$$\delta_8(R) \leq \langle \hat{\omega}, d \rangle < \alpha d(4) = \delta_2\left(R/3\right) ,$$

thus concluding the proof. ∎

As a consequence of Lemma 2, and our previous arguments, we have proven that the typical group code achieves the GV bound. In order to complete the proof of Theorem 1, it remains to prove that the normalized minimum distance of the typical group code does not exceed the GV distance. This is technically more involved, and will be the object of Sect. B.

We conclude this section by showing that the typical group code achieves the expurgated exponent $E_8^x(R)$. For this, we shall use the union-Bhattacharyya bound [39], in order to estimate of the error probability of the GCE in terms of its type-enumerating functions:

$$p_e(\mathcal{G}_n^R) \leq \sum_{\omega} G_n^R(\omega) 2^{-n\langle \omega, d \rangle} . \tag{20}$$

Similarly to what we have seen for the analysis of the minimum distance, it is natural to consider the expurgated exponents of the 4-PSK and 2-PSK AWGNC, given by

$$\begin{aligned} E_4^x(\tfrac{2}{3}R) &:= \min\left\{ \langle \omega, d \rangle - \mathrm{H}(\omega) + \tfrac{2}{3}\overline{R} : \omega \in \Omega_R'' \right\} , \\ E_2^x(\tfrac{1}{3}R) &:= \min\left\{ \langle \omega, d \rangle - \mathrm{H}(\omega) + \tfrac{1}{3}\overline{R} : \omega \in \Omega_R' \right\} , \end{aligned}$$

where Ω_R' and Ω_R'' have been defined in (15). Then, based on Lemma 1 and (20), a first-moment argument as in (16) allows one to show that

$$\begin{aligned} p_e(\mathcal{G}_n^R) \quad \leq \quad & 2^{-nE_8^x(R')+o(n)} + 2^{-nE_4^x(\frac{2}{3}(R'))+o(n)} \\ & + 2^{-nE_2^x(\frac{1}{3}(R'))+o(n)} , \end{aligned} \tag{21}$$

for every $R' > R$. On the other hand, arguing as in the proof of Lemma 2, one can show that

$$E_8^x(R) \leq E_4^x(\tfrac{2}{3}R) \leq E_2^x(\tfrac{1}{3}R) . \tag{22}$$

Hence, (21), (22), and the continuity of $E_8^x(R)$ as a function of the rate R show that the typical group code achieves the expurgated exponent.

IV. Performance of the typical binary-coset code

In the present section, we shall prove that the typical binary-coset code is bounded away both from the GV distance and the expurgated exponent. We shall proceed in two steps. First, in Sect. IV-A, we shall prove that the normalized minimum distance of the typical binary-coset code of design rate R is between $\underline{\delta}_\eta(R)$ and $\overline{\delta}_\eta(R)$. This will involve the use of the fist moment, and the second moment method, respectively. Then, in Sect. IV-B, we shall prove the rightmost inequality in (12): 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 Θ and Θ_n for the spaces of joint probability vectors, and of joint types, respectively, over $\mathbb{Z}_2^3 \times \mathbb{Z}_2^3$. Also Υ, and Υ_n, respectively, will denote the spaces of probability vectors, and of types, over $\mathbb{Z}_2^3$.

It will prove convenient to consider a slightly different version of the binary-coset ensemble, as explained below. Observe that, since the rows of Φ_n^R are mutually independent and uniformly distributed over $\mathbb{Z}_2^{3n}$, the probability that the j-th row of Φ_n^R is linear dependent on the other $(3l-1)$ rows is bounded from above by $2^{-3n}2^{3l-1} \leq 2^{-nR}$. Then a standard union-bounding technique implies that the probability of the event $A := \{\Psi_n^R \text{ is surjective}\}$ is at least $1 - n2^{-nR}$, and therefore converges to 1 as n grows. Now, consider a random vector $\boldsymbol{Z}_n$ uniformly distributed over $\mathbb{Z}_2^{3l}$, and independent from Ψ_n^R. Notice that, given A_n, the conditioned probability measures of the random cosets $\ker \Psi_n^R + \boldsymbol{W}_n$, and $(\Psi_n^R)^{-1}\boldsymbol{Z}_n$, both coincide with the uniform distribution on the set of affine spaces of $\mathbb{Z}_2^{3n}$ of dimension $3(n-l)$. Therefore, every statement concerning properties of the typical binary-coset ensemble is not altered if one replaces definition (4) with

$$\mathcal{B}_n^R := \eta_n\left((\Psi_n^R)^{-1}\boldsymbol{Z}_n\right) . \tag{23}$$

Therefore, from now on, we shall consider (23) to be the definition of the random binary-coset code.

A. Upper and lower bounds on the minimum distance of the typical binary-coset code

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 function

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

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

$$V_n^R(\upsilon) := \left|\left\{\boldsymbol{x} \in (\mathbb{Z}_2^3)_\upsilon^n : \Psi_n^R\boldsymbol{x} = \boldsymbol{0}\right\}\right| ,$$

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

$$\min\left\{\langle \theta, D_\eta \rangle : \theta \in \Theta, \textstyle\sum_x \theta(0,x) < 1, U_n^R(\theta) \geq 1\right\} .$$

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

Lemma 3 *For every $\theta \in \Theta_n$, let $\upsilon(\,\cdot\,) = \sum_x \theta(\,\cdot\,, x) \in \Upsilon$ be its first-component marginal of θ. If $\upsilon(0) < 1$, then*

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

Proof: For every $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{Z}_2^{3n}$ such that $\boldsymbol{x} \neq \boldsymbol{0}$ 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

$$\mathbb{E}\left[U_n^R(\theta)\right] = \sum_{(\boldsymbol{x},\boldsymbol{y})} \mathbb{P}\left(\Psi_n^R\boldsymbol{x} = \boldsymbol{0}\,,\ \Psi_n^R\boldsymbol{y} = \boldsymbol{Z}_n\right) = \binom{n}{n\theta} 8^{-2l} ,$$

where the summation above is extended to all pairs $(\boldsymbol{x},\boldsymbol{y})$ of joint type θ. The expectation $\mathbb{E}[V_n^R(\upsilon)]$ is computed analogously. ■

Fix some $R^* > R$. Using Lemma 3, an argument based on a first-moment method, and analogous to the one applied in Sect. III, proves that the probability that $U_n^R(\theta) \geq 1$ for some joint type θ with either $\mathrm{H}(\theta) \leq 2\overline{R}^*$, or $\mathrm{H}(\theta_1) \leq \overline{R}^*$ goes to zero as n grows to infinity. Thanks to the continuity of $\underline{\delta}_\eta(R)$ and the arbitrariness of $R^* > R$, this proves that the normalized minimum distance of the typical binary-coset code is bounded from below by $\underline{\delta}_\eta(R)$.

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

Lemma 4 *For all $n \geq 1$, and every joint type θ,*

$$\mathrm{Var}\left[U_n^R(\theta)\right] \leq \binom{n}{n\theta}\binom{n}{n\upsilon}\frac{16}{8^{3l}} + \frac{\binom{n}{n\theta}^2}{\binom{n}{n\upsilon}}\frac{1}{8^{3l}} + \binom{n}{n\theta}\frac{8}{8^{2l}}, \tag{24}$$

where υ is the first-component marginal of θ.

Proof: We have

$$\mathrm{Var}\left[U_n^R(\theta)\right] = \sum_{(\boldsymbol{x}_1,\boldsymbol{y}_1)}\sum_{(\boldsymbol{x}_2,\boldsymbol{y}_2)} c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big),$$

where the summations are extended to all pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1)$ and $(\boldsymbol{x}_2,\boldsymbol{y}_2)$ of joint type θ, and $c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big)$ is the covariance of $\mathbb{1}_{\{\Psi_n^R\boldsymbol{x}_1=\boldsymbol{0},\Psi_n^R\boldsymbol{y}_1=\boldsymbol{Z}_n\}}$ and $\mathbb{1}_{\{\Psi_n^R\boldsymbol{x}_2=\boldsymbol{0},\Psi_n^R\boldsymbol{y}_2=\boldsymbol{Z}_n\}}$.

We are now going to estimate the covariance terms $c\big(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1,\boldsymbol{y}_2\big)$, by 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 $\upsilon(0) < 1$, $\boldsymbol{x}_1$ 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 θ. First, suppose that $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are linear independent. Then, the r.v.s $\Psi_n^R\boldsymbol{x}_1,\Psi_n^R\boldsymbol{x}_2,\Psi_n^R\boldsymbol{y}_1$ and $\Psi_n^R\boldsymbol{y}_2$ 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 so. In this case we have that the random variables $\Psi_n^R\boldsymbol{x}_1,\Psi_n^R\boldsymbol{x}_2$ and $\Psi_n^R\boldsymbol{y}_1-\boldsymbol{Z}_n$ are independent, so that

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

Since there are at most $16\binom{n}{n\theta}\binom{n}{n\upsilon}$ possible choices of such pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)$ of joint type θ, their contribution is estimated by the first addend in the righthand side of (24).

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 $\Psi_n^R\boldsymbol{x}_1$, $\Psi_n^R\boldsymbol{y}_1$ and $\Psi_n^R\boldsymbol{y}_2$ are independent so that

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

Since there are at most $\binom{n}{n\theta}^2\binom{n}{n\upsilon}^{-1}$ possible choices of such pairs $(\boldsymbol{x}_1,\boldsymbol{y}_1),(\boldsymbol{x}_2,\boldsymbol{y}_2)$ of joint type θ, their contribution can be estimated by the second addend in the right-hand side of (24).

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

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

Therefore, their contribution can be estimated by the third addend in the righthand side of (24). ■

Let us now fix some $R^* > R$, and some $\theta^* \in \overline{\Omega}_{R^*}$ such that $\overline{\delta}_\eta(R^*) = \langle\theta^*, D_\eta\rangle$. Denote by υ^* the first-component marginal of θ^*. Consider a sequence of joint types (θ_n) converging to θ^*, with θ_n in Θ_n for every $n \geq 1$, and let (υ_n) be the corresponding sequence of first-component marginals. Define the event $A_n := \{U_n^R(\theta_n) = 0\}$. We can apply Chebyshev's inequality and use Lemma 3 and Lemma 4 obtaining

$$\begin{aligned}\mathbb{P}(A_n) &\leq \mathrm{Var}\left[U_n^R(\theta_n)\right]\mathbb{E}\left[U_n^R(\theta_n)\right]^{-2}\\ &\leq 16\tbinom{n}{n\upsilon_n}\tbinom{n}{n\theta_n}^{-1}8^l + \tbinom{n}{n\upsilon_n}^{-1}8^l + 8\tbinom{n}{n\theta_n}^{-1}8^{2l}\\ &= 2^{n(\overline{R}+\mathrm{H}(\upsilon_n)-\mathrm{H}(\theta_n))+o(n)} + 2^{n(\overline{R}-\mathrm{H}(\upsilon_n))+o(n)}\\ &\quad + 2^{n(2\overline{R}-2\,\mathrm{H}(\theta_n))+o(n)}.\end{aligned}$$

Then, since $\lim_n\theta_n = \theta^* \in \overline{\Omega}_{R^*}$, with $R^* > R$, one has that $\lim_n\mathbb{P}(A_n) = 0$. From this, it follows that the typical binary-coset code has normalized minimum distance not exceeding $\langle\theta^*, D_\eta\rangle = \underline{\delta}_\eta(R^*)$. Finally, from the arbitrariness of $R^* > R$, a standard continuity argument allows one to conclude that the normalized minimum distance of the typical binary-coset code is upper-bounded by $\underline{\delta}_\eta(R)$.

B. Comparing $\overline{\delta}_\eta(R)$ and $\delta_8(R)$

We now want to compare the distance bounds $\underline{\delta}_\eta(R)$, $\overline{\delta}_\eta(R)$, and $\delta_8(R)$ defined in (10), (11) and (7) respectively. First, observe that any joint type $\theta \in \overline{\Theta}_R$ trivially satisfies $\mathrm{H}(\theta) \geq 2\overline{R}$, so that $\overline{\Theta}_R \subseteq \underline{\Theta}_R$. From this, it immediately follows that $\overline{\delta}_\eta(R) \geq \underline{\delta}_\eta(R)$. Notice also that the inequality above holds as an equality whenever $\underline{\delta}_\eta(R) = \langle\theta, D_\eta\rangle$ for some joint type θ 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 $\overline{\delta}_\eta(R)$ and $\underline{\delta}_\eta(R)$ do coincide. However, we will now concentrate on comparing $\overline{\delta}_\eta(R)$ with the GV-distance $\delta_8(R)$, in particular showing that the former is strictly below the latter.

In order to do that, for a given $0 < R < 3$, we consider the $\mathbb{Z}_8$-type ω^* in Ω_R giving the GV-distance, i.e. such that $\delta_8(R) = \langle\omega^*, d\rangle$. Since the entropy function is concave and the map $\omega \mapsto \langle\omega, d\rangle$ is linear and it achieves its global minimum in $\omega(0) = 1$, Lemma 8 can be applied to guarantee that $\mathrm{H}(\omega^*) = \overline{R}$. Hence, using Lagrangian multipliers we may express it as

$$\omega^*(x) = Z(\lambda)^{-1}e^{-\lambda d(x)}, \tag{25}$$

where $Z(\lambda) := \sum_x e^{-\lambda d(x)}$ and $\lambda \in (0,+\infty)$ is the unique solution of the equation $\mathrm{H}\left(Z(\lambda)^{-1}e^{-\lambda d}\right) = \overline{R}$. From $\omega^* \in \Omega$, we may define a joint type θ^* in Θ as follows. For every z in $\mathbb{Z}_2^3$, consider the bijection

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

where the $+$ sign refers to addition in $\mathbb{Z}_8$, while the $-$ refers to difference in $\mathbb{Z}_2^3$. Now define

$$\theta^*(x,z) := \tfrac{1}{8}\omega^*\left(\sigma_z(x)\right)\,, \qquad x,z\in\mathbb{Z}_2^3\,, \tag{26}$$

and let υ^* be its first-component marginal. A few simple properties of θ^* and υ^* are gathered in the following:

Lemma 5 *For all $0<R<3$,*

$$\theta^*(x,z)>0\,, \qquad \upsilon^*(x)>0\,, \qquad \forall\, x,\, z\in\mathbb{Z}_2^3\,, \tag{27}$$

$$\langle\theta^*, D_\eta\rangle = \delta_8(R)\,. \tag{28}$$

$$\mathrm{H}(\theta^*) = 3+\overline{R}\,. \tag{29}$$

$$\mathrm{H}(\upsilon^*) > \overline{R}\,. \tag{30}$$

Proof: The inequality (27) follows immediately from (25).

It is easy to verify that

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

Then (28) follows, since

$$\begin{aligned} \langle\theta^*, D_\eta\rangle &= \textstyle\sum_{x,z}\theta^*(x,z)D_\eta(x,z) \\ &= \textstyle\sum_z \frac{1}{8}\omega^*\left(\sigma_z(x)\right)d\left(\sigma_z(x)\right) \\ &= \langle\omega^*, d\rangle \\ &= \delta_8(R)\,. \end{aligned}$$

From (26) we have $\sum_x\theta^*(x,z)=\frac{1}{8}\sum_x\omega^*(\sigma_z(x))=\frac{1}{8}$, so that the second-component marginal θ_2^* is the uniform measure over $\mathbb{Z}_2^3$. Again from (26) we have that the conditioned measure of θ on $\mathbb{Z}_2^3\times\{z\}$ coincides with $\omega^*\circ\sigma_z$, for every z in $\mathbb{Z}_2^3$. Then, one has

$$\begin{aligned} \mathrm{H}(\theta^*) &= \mathrm{H}\left(\theta_2^*\right)+\textstyle\sum_x\theta^*\left(\mathbb{Z}_2^3\times\{x\}\right)\mathrm{H}\left(\omega^*\circ\sigma_x\right) \\ &= 3+\mathrm{H}(\omega^*) \\ &= 3+\overline{R}\,, \end{aligned}$$

showing (29).

Finally, observe that $\upsilon^* = \frac{1}{8}\sum_x\omega^*\circ\sigma_x$ is a convex combination of permutations of the vector ω^*. As argued in Sect. II-A, for every labeling $\eta:\mathbb{Z}_2^3\to\mathcal{X}$ there exists at least a pair of nonequal columns of the matrix D_η, i.e.

$$D_\eta(\cdot,z_1)\neq D_\eta(\cdot,z_2)\,,$$

for some $z_1,z_2\in\mathbb{Z}_2^3$. As a consequence, $d\circ\sigma_{z_1}\neq d\circ\sigma_{z_2}$ which, together with (25), implies $\omega^*\circ\sigma_{z_1}\neq\omega_{(R)}\circ\sigma_{z_2}$. Hence, from the strict concavity and the permutation invariance of the entropy function H it follows that

$$\begin{aligned} \mathrm{H}(\upsilon^*) &= \mathrm{H}\left(\tfrac{1}{8}\textstyle\sum_x\omega^*\circ\sigma_x\right) \\ &> \tfrac{1}{8}\textstyle\sum_x\mathrm{H}\left(\omega^*\circ\sigma_x\right) \\ &= \mathrm{H}(\omega^*) \\ &= \overline{R}\,, \end{aligned}$$

showing (30). ∎

We are now ready to prove the rightmost inequality in (12).

Proposition 1 *For every labeling $\eta:\mathbb{Z}_2^3\to\mathbb{Z}_8$,*

$$\overline{\delta}_\eta(R)<\delta_8(R)\,,$$

for all $0<R<3$,

Proof: For $x\in\mathbb{Z}_2^3$, let $m_x:=\min\{D_\eta(x,z)\,:\,z\in\mathbb{Z}_2^3\}$ be the minimum element of the x-th row of D_η, and $M_x:=\{z\in\mathbb{Z}_2^3\,:\,D_\eta(x,z)=m_x\}$ the set of minimizers. Observe that $D_\eta(0,z)=D(\eta(z),\eta(z))=0$ for every binary labeling η and any $z\in\mathbb{Z}_2^3$. Therefore, one has $m_0=0$ and $|M_0|=8$. On the other hand, since no binary labeling η is isometric, there necessarily exist some non-costant row of D_η, so that in particular

$$\exists\, x,z\in\mathbb{Z}_2^3:\ m_x<D_\eta(x,z)\,. \tag{31}$$

For $\upsilon\in\Upsilon$, consider the set $\Theta_\upsilon\subseteq\Theta$ of joint measures with first-component marginal υ, and define $f(\upsilon):=\min\left\{\langle\theta,D_\eta\rangle\,:\,\theta\in\Theta_\upsilon,\,\mathrm{H}(\theta)-\mathrm{H}(\upsilon)\geq\overline{R}\right\}$. As an immediate consequence of (30), one has that $\overline{\delta}_\eta(R)\leq f(\upsilon^*)$. Therefore, in order to prove the claim, it is sufficient to show that $f(\upsilon^*)<\delta_8(R)$.

First, suppose that $\sum_x\upsilon^*(x)\log n_x\geq\overline{R}$. Then, it is easy to check that $f(\upsilon^*)=\sum_x\upsilon^*(x)m_x$. Hence, it follows from, (27), (31) and (28) that

$$\begin{aligned} f(\upsilon^*) &= \textstyle\sum_x m_x\upsilon^*(x) \\ &= \textstyle\sum_x\sum_z\frac{1}{8}\theta^*(\sigma_z(x))m_x \\ &< \textstyle\sum_x\sum_z\frac{1}{8}\theta^*(\sigma_z(x))D_\eta(x,z) \\ &= \delta_8(R)\,, \end{aligned}$$

thus proving the claim.

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

$$\begin{aligned} \upsilon^*(0) &= \textstyle\sum_z\theta^*(0,z) \\ &= Z(\lambda)^{-1} \\ &> Z(\lambda)^{-1}\tfrac{1}{8}\textstyle\sum_z e^{-\lambda d(\sigma_z(x))} \\ &= \textstyle\sum_z\theta^*(x,z) \\ &= \upsilon^*(x)\,. \end{aligned}$$

Hence, υ^* is not the uniform measure over $\mathbb{Z}_2^3$ and, as a consequence $\mathrm{H}\left(\upsilon^*\right)<3$. Therefore, from (29) and (30), $\mathrm{H}(\theta^*)=3+\overline{R}>\mathrm{H}(\upsilon^*)+\overline{R}$. Then, thanks to the concavity of the entropy function, we can apply Lemma 8, obtaining that

$$f(\upsilon^*)<\langle\theta^*,D_\eta\rangle=\delta_8(R)\,,$$

the last equality following from (28). ∎

It immediately follows from Proposition 1, and the results of Sect. IV-B, that the typical binary-coset code is bounded away from the GV distance. In fact, the analogous statement holds for the expurgated exponent as well. To see that, arguing as in [6, p. 413], one can show that

$$p_e(\mathcal{B}_n^R)\geq 2^{-\,\mathrm{d}_{\min}(\mathcal{B}_n^R)+o(n)}\geq 2^{-n\overline{\delta}_\eta(R)+o(n)}\,.$$

Since $E_8^x(R)=\delta_8(R)>\overline{\delta}_\eta(R)$, at low rates, this shows that the typical binary-coset code is does not achieve the expurgated exponent.

V. Extensions and concluding remarks

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 2 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. IV 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

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

Assume that the DMC has has zero-error capacity equal to zero, so that $D(x_1, x_2)$ is finite for every $x_1, x_2 \in \mathcal{X}$, and further that it is balanced (see [32]), i.e. that, for all $x, z \in \mathcal{X}$,

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

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

$$\delta(R) := \min\{\langle \omega, d\rangle : \omega \in \Omega_R\}\,,$$

$$E^x(R) := \min\{\langle \omega, d\rangle - \mathrm{H}(\omega) + R : \omega \in \Omega_R\}\,,$$

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

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

Now consider the automorphism group, i.e. the subgroup of distance-preserving permutations of $\mathcal{X}$,

$$\Pi := \{\pi : D(\pi(x_1), \pi(x_2)) = D(x_1, x_2)\,, \forall x_1, x_2 \in \mathcal{X}\}\,.$$

Assume that Π does not have any subgroup isomorphic to $\mathbb{Z}_p^r$. Then, for any labeling $\eta : \mathbb{Z}_p^r \to \mathcal{X}$, the matrix D_η defined as in (3) has at least two distinct columns. Then, it follows that both Theorem 2 and Corollary 2 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 Π 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 or 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 work. Some of the proofs, especially those concerning the arguments on the expurgated exponent, have been considerably simplified thanks to the valuable suggestions of one of the anonymous referees.

Appendix 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 $\delta_8(R)$ and the expurgated error exponent $E_8^x(R)$, as well as the bounds $\overline{\delta}_\eta(R)$ and $\underline{\delta}_\eta(R)$. We shall obtain these results as a consequence of the general lemmas presented below.

For some fixed $d \ge 1$, 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

$$f : \mathbb{R} \to \overline{\mathbb{R}}\,, \qquad f(y) := \inf\left\{g(\xi)\middle|\, \xi \in \Xi : h(\xi) \le y\right\}\,. \tag{32}$$

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

Lemma 6 *If g and h are both l.s.c., then f defined in (32) 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 7 *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 (32) is continuous on $[h^*, +\infty)$ where $h^* := \min\{h(\xi)\,|\,\xi \in \Xi\}$.*

Proof: Since f is nonincreasing and l.s.c. by Lemma 6, it remains to show that

$$\lim_n f(y_n) \le f(y) \tag{33}$$

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 (33) 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 $h(\xi) \leq y$. If $h(\xi) < y$, then $h(\xi) \leq y_n$ for sufficiently large n, so that $f(y_n) \leq g(\xi) = f(y)$ definitively in n and (33) 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 $h(\xi_{n_k}) \leq y_k$ for every k. Therefore we have $f(y_k) \leq g(\xi_{n_k})$ and so

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

thus concluding the proof. ∎

By considering $\Xi = \Omega$, $h(\omega) = -\mathrm{H}(\omega)$ and $g(\omega) = \langle \omega, d\rangle$ (respectively $g(\omega) = \langle \omega, d\rangle + \overline{R} - \mathrm{H}(\omega)$), Lemma 7 implies the continuity of $\delta_8(R)$ ($E_8^x(R)$). Indeed, observe that $-\mathrm{H}(\,\cdot\,)$ is convex and therefore does not admit local minima which are not global minima. Similarly, the continuity of $\delta_\eta(R)$ follows by taking $\Xi = \Theta$, $g(\theta) = \langle \theta, D_\eta\rangle$, and $h(\theta) = \max\{-\frac{1}{2}\mathrm{H}(\theta), -\mathrm{H}(\upsilon)\}$, where υ is the first component marginal of θ. Observe that h is convex, as it is the maximum of two convex functions.

Finally, the continuity of $\overline{\delta}_\eta(R)$ follows again by applying Lemma 7 with the choices $\Xi = \Theta$, $g(\theta) = \langle \theta, D_\eta\rangle$, and $h(\theta) = \max\{-\mathrm{H}(\theta) + \mathrm{H}(\upsilon), -\mathrm{H}(\upsilon)\}$. In this last case, the absence of strictly local minima of h can be verified directly as follows. If $\theta \in \Theta$ is a local minimum for $h(\,\cdot\,)$, and $h(\theta) = -\mathrm{H}(\theta) + \mathrm{H}(\upsilon)$, then, for every x such that $\upsilon(x) > 0$, necessarily the conditional measure of θ on $\{x\}\times\mathbb{Z}_2^3$ coincides with the uniform distribution over $\mathbb{Z}_2^3$. It follows that $h(\theta) = -3$, and therefore θ 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 8 *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*

$$y < h^*\,, \qquad h^* := \min_{\xi\in\Xi^*} h(\xi)\,,$$

any minimizer ξ_y for the convex optimization problem

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

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

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

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

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

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

Then, $f(y) = \min_{h(\xi)\leq y} g(\xi) \leq g(\xi_{\lambda^*}) < g(\xi)$, so that ξ cannot be a minimizer. ∎

Appendix B
An upper bound on the normalized minimum distance of the typical group code

In this section we shall show that the normalized minimum distance of the typical group code of design rate R does not exceed the GV distance $\delta_8(R)$, thus completing the proof of Theorem 1. Our arguments involve an application of the second moment method [1, pagg.43-63].

The key point consists in estimating the covariance of the type-enumerating function $G_n^R(\omega)$, for every type $\omega \in \Omega_n$. For this, one has to compute the joint probabilities

$$\mathbb{P}(\Phi_n^R \boldsymbol{x} = \mathbf{0}, \Phi_n^R \boldsymbol{y} = \mathbf{0})\,,$$

for all pairs $(\boldsymbol{x}, \boldsymbol{z}) \in (\mathbb{Z}_8)_\omega^n \times (\mathbb{Z}_8)_\omega^n$. Such a joint probability does not depend on the type ω only, but on the specific choices of $\boldsymbol{x}$ and $\boldsymbol{z}$ as well. In particular, let $m = 2^{\zeta(\omega)}$ be the order of the smallest subgroup of $\mathbb{Z}_8$ supporting ω, and observe that the subgroup of $\mathbb{Z}_8^n$ generated by $\boldsymbol{x}$ and $\boldsymbol{z}$ is necessarily 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 $\boldsymbol{x} = \boldsymbol{w}$). In other words, it is possible to partition the set of ordered pairs of n-tuples of type ω as follows:[4]

$$(\mathbb{Z}_8)_\omega^n \times (\mathbb{Z}_8)_\omega^n = \bigcup_{h|m} A_{n,\omega,h}\,, \tag{34}$$

with $A_{n,\omega,h}$ denoting the set of all pairs $(\boldsymbol{x}, \boldsymbol{z})$ 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 $A_{n,\omega,h}$. For every $h \mid 8$, consider the probability measure τ_h over $\{0, \ldots, 8/h - 1\}$, defined by

$$\tau_h(i) := \omega(i + \tfrac{8}{h}\mathbb{Z}_8)\,.$$

Also, for all $0 \leq i < h$, let ω_h^i be the conditional distribution of ω over the coset $i + \frac{8}{h}\mathbb{Z}_8$, i.e.

$$\omega_h^i(j) := \tau_h(i)^{-1}\omega(j)\,, \qquad j \in i + \tfrac{8}{h}\mathbb{Z}_8\,.$$

Lemma 9 *For every n, ω in $\mathcal{P}_n(\mathbb{Z}_8)$, and $h \mid m$, one has that*

$$|A_{n,\omega,h}| \leq 4\binom{n}{n\omega} \prod_{\substack{1\leq i\leq 8/h:\\ \tau_h(i)>0}} \binom{n_i}{n_i\omega_h^i}\,, \tag{35}$$

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

Proof: Let $\boldsymbol{x}$ and $\boldsymbol{z}$ be in $(\mathbb{Z}_8)_\omega^n$. A necessary condition for the subgroup of $\mathbb{Z}_8^n$ generated by $\boldsymbol{x}$ and $\boldsymbol{z}$ to be isomorphic

[4] For two naturals a and b, $a \mid b$ stays for "a divides b".

to $\frac{8}{m}\mathbb{Z}_8 \oplus \frac{8}{h}\mathbb{Z}_8$ is the existence of some α in the set $\mathbb{Z}_8^*$ of invertible elements of Z_8, such that

$$-h\alpha\boldsymbol{x} + h\boldsymbol{z} = \boldsymbol{0}\,. \tag{36}$$

For (36) to hold, necessarily $\boldsymbol{z}$ has to belong to the coset $\alpha\boldsymbol{x} + \frac{8}{h}\mathbb{Z}_8^n$. Thus, whenever (36) 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. Notice that since both $\boldsymbol{x}$ and $\boldsymbol{z}$ are assumed to be of type ω, this implies that

$$\tau_h(i) = \tau_h(\alpha i)\,, \qquad i = 0, \ldots, 8/h - 1\,. \tag{37}$$

For those α for which (37) is not satisfied there exists no pair $(\boldsymbol{x}, \boldsymbol{z})$ satisfying (36). Thus, with no loss of generality we can restrict ourselves to considering values of α such that (37) 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 $(\mathbb{Z}_8)^n_\omega$ is the existence of an index permutation $\sigma : \{1, 2. \ldots, n\} \to \{1, 2. \ldots, n\}$ such that $\sigma\boldsymbol{x} := \boldsymbol{x} \circ \sigma^{-1} = \boldsymbol{z}$. Furthermore, (37) imposes an additional constraint on the structure of σ, which has necessarily to be of the form $\sigma = \sigma^1 \circ \sigma^2 \circ \ldots \circ \sigma^{8/h} \circ \tilde{\sigma}$, where:

- $\tilde{\sigma}$ is some permutation mapping, for all i, all the indices corresponding to entries of $\boldsymbol{x}$ in the coset $i + \frac{8}{h}\mathbb{Z}_8$, into the indices corresponding to the entries of $\boldsymbol{z}$ in the same coset;
- for every i, σ^i permutes only the indices corresponding to the entries of $\boldsymbol{x}$ in $i + \frac{8}{h}\mathbb{Z}_8$.

Thus, for a given $\boldsymbol{x}$ in $(\mathbb{Z}_8)^n_\omega$ and α in $\mathbb{Z}_8^*$ such that (37) is satisfied, we have that the number of $\boldsymbol{z}$ such that $(\boldsymbol{x}, \boldsymbol{z}) \in A_{n,\omega,h}$ equals the cardinality of the orbit of $\tilde{\sigma}\boldsymbol{x}$ under the action of the subgroup of index permutations

$$S := \{\sigma = \sigma^1 \circ \sigma^2 \circ \ldots \circ \sigma^{8/h}\}\,,$$

where, for every i, σ^i permutes only the indices corresponding to the entries of $\boldsymbol{x}$ in $i + \frac{8}{h}\mathbb{Z}_8$. Clearly the order of this group is $|S| = \prod_{i=1}^{8/h} n_i!$, while the cardinality of the stabilizer of $\tilde{\sigma}_{\alpha,\boldsymbol{x}}\boldsymbol{x}$ in S is $\prod_{i=1}^{8} (n\omega(i))!$, so that the orbit of $\tilde{\sigma}\boldsymbol{x}$ in S has cardinality

$$\prod_{i=1}^{8/h} n_i! \Big/ \prod_{i=1}^{8} (n\omega(i))! = \prod_{i=1}^{8/h} \binom{n_i}{n_i\omega_h^i}\,.$$

This allows one to conclude the proof. ∎

Lemma 10 *For every $n \geq 1$, and $\omega \in \Omega_n$,*

$$\mathrm{Var}\left[G_n^R(\omega)\right] \leq 4\binom{n}{n\omega} m^{-l} \sum_{\substack{h|m \\ h<m}} h^{-l} \prod_{i=1}^{8/h} \binom{n_i}{n_i\omega_h^i}\,, \tag{38}$$

for n_i defined as in Lemma 9.

Proof: Assume that $\boldsymbol{x}, \boldsymbol{z} \in A_{n,\omega,h}$ for some $h \mid m$. Notice that, for every $1 \leq j \leq l$, the pair $\left((\Phi_n^R\boldsymbol{x})_j, (\Phi_n^R\boldsymbol{z})_j\right)$ is uniformly distributed over the subgroup of $\mathbb{Z}_8^2$ generated by $\{(x_i, z_i) : 1 \leq i \leq n\}$, which is isomorphic to K, the subgroup of $\mathbb{Z}_8^n$ generated by $\boldsymbol{x}$ and $\boldsymbol{z}$. As K is in turn is isomorphic to a group of type $\frac{8}{m}\mathbb{Z}_8 \oplus \frac{8}{h}\mathbb{Z}_8$, it follows that

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

Observe that the r.v.s $\left\{(\Phi_n^R x)_j, (\Phi_n^R z)_j : 1 \leq j \leq l\right\}$ are mutually independent, since they correspond to different rows of the random matrix Φ_n^R. Then, one has

$$\mathbb{P}\left(\Phi_n^R\boldsymbol{x} = \boldsymbol{0}, \Phi_n^R\boldsymbol{z} = \boldsymbol{0}\right) = (hm)^{-l}\,. \tag{39}$$

It follows from (34), (35) and (39) that

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

and the claim follows immediately from Lemma 9. ∎

We are now ready to state the main result of this section, whose proof will involve geometric considerations on the 8-PSK constellation:

Proposition 2 *For every $0 < R^* < R < 3$, the minimum distance of the typical group code of design rate R is upper-bounded by $\delta_8(R^*)$.*

Proof: Let $\omega \in \Omega_{R^*}$ be such that $\langle \omega, d\rangle = \delta_8(R^*)$. As an immediate consequence of Lemma 1 and Lemma 10, one has

$$\begin{aligned}\frac{\mathrm{Var}\left[G_n^R(\omega)\right]}{\mathbb{E}\left[G_n^R(\omega)\right]^2} &\leq \binom{n}{n\omega}^{-1} \sum_h \left(\frac{m}{h}\right)^l \prod_{i=1}^{8/h} \binom{n_i}{n_i\omega_h^i} \\ &= 2^{n \max_h \left\{\frac{R}{3}\log\frac{m}{h} - \mathrm{H}(\tau_h)\right\} + o(n)}\,,\end{aligned} \tag{40}$$

with the index h, in both the summation and maximization above, running over all divisors of m, excluding m itself.

Observe that (17) and (18) imply that the entries of ω satisfy the following ordering

$$\omega(0) > \omega(1) = \omega(7) > \omega(2) = \omega(6) > \omega(3) = \omega(5) > \omega(4). \tag{41}$$

Define the sets $A_0 := \{0, 1, 7, 2\}$, $B_0 := \{0, 1, 6, 3\}$, and $C_0 := \{0, 5, 6, 7\}$. Let A_1, B_1 and C_1 be the complements, in $\mathbb{Z}_8$, of A_0, B_0, and C_0, respectively. It follows from (41) that

$$\begin{aligned}&\omega(A_0) \geq \omega(2\mathbb{Z}_8)\,, && \omega(A_0) \geq \omega(2\mathbb{Z}_8 + 1)\,, \\ &\omega(B_0) \geq \omega(2\mathbb{Z}_8)\,, && \omega(B_0) \geq \omega(2\mathbb{Z}_8 + 1)\,, \\ &\omega(C_0) \geq \omega(2\mathbb{Z}_8)\,, && \omega(C_0) \geq \omega(2\mathbb{Z}_8 + 1)\,.\end{aligned} \tag{42}$$

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, it follows from (42) that

$$\mathrm{H}(\omega) \geq \mathrm{H}\left(\omega(A_0)\right) + \mathrm{H}\left(\omega(B_0)\right) + \mathrm{H}\left(\omega(C_0)\right) = 3\,\mathrm{H}(\tau_4). \tag{43}$$

Let us now introduce the sets $D := \{0, 2\}$ and $E := \{1, 7\}$. We have from (41) that

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

It thus follows that

$$\begin{aligned}\mathrm{H}(\tau_2) &= \mathrm{H}(\tau_4)+\tau_4(0)\,\mathrm{H}\left(\omega_4^0(4\mathbb{Z}_8)\right)+\tau_4(1)\,\mathrm{H}\left(\omega_4^1(4\mathbb{Z}_8+1)\right)\\ &\geq \mathrm{H}(\tau_4)+\tau_4(0)\,\mathrm{H}\left(\omega_4^0(D)\right)+\tau_4(1)\,\mathrm{H}\left(\omega_4^1(E)\right)\,.\end{aligned} \tag{44}$$

Observe that

$$\begin{aligned}\omega_4^0(D) &= \tau_4(0)^{-1}\omega(0)+\tau_4(0)^{-1}\omega(2)\\ &= \omega_4^0(4\mathbb{Z}_8)\omega_2^0(0)+\omega_4^0(4\mathbb{Z}_8+2)\omega_2^2(2)\,.\end{aligned}$$

By the concavity of the entropy function, one has that

$$\mathrm{H}\left(\omega_4^0(D)\right)\geq\omega_4^0(4\mathbb{Z}_8)\,\mathrm{H}\left(\omega_2^0(0)\right)+\omega_4^0(4\mathbb{Z}_8+2)\,\mathrm{H}\left(\omega_2^0(2)\right).$$

An analogous reasoning leads to

$$\mathrm{H}\left(\omega_4^1(E)\right)\geq\omega_4^1(4\mathbb{Z}_8+1)\,\mathrm{H}\left(\omega_2^1(1)\right)+\omega_4^1(4\mathbb{Z}_8+3)\,\mathrm{H}\left(\omega_2^3(3)\right).$$

Upon substituting the two inequalities above in (44), one gets

$$\begin{aligned}\mathrm{H}(\tau_2) &\geq \mathrm{H}(\tau_4)+\textstyle\sum_{i=0}^3\omega(4\mathbb{Z}_8+i)\,\mathrm{H}(\omega_2^i(i))\\ &= \mathrm{H}(\tau_4)+\mathrm{H}(\omega)-\mathrm{H}(\boldsymbol{\tau}_2)\\ &\geq \tfrac{4}{3}\,\mathrm{H}(\omega)-\mathrm{H}(\tau_2)\,,\end{aligned}$$

last inequality following from (43). Then

$$\mathrm{H}\left(\tau_2\right)\geq\frac{2}{3}\,\mathrm{H}(\omega)\,. \tag{45}$$

Now let (ω_n) be a sequence converging to ω, with $\omega_n\in\Omega_n$ for every n. By successively applying Chebyshev's inequality, (40), (43) and (45), one gets

$$\begin{aligned}\mathbb{P}\left(G_n^R(\omega_n)=0\right) &\leq \mathrm{Var}\left[G_n^R(\omega_n)\right]\mathbb{E}[G_n^R(\omega_n)]^{-2}\\ &\leq 2^{n\max\left\{\frac{\overline{R}}{3}-\mathrm{H}(\tau_4),\frac{2}{3}\overline{R}-\mathrm{H}(\tau_2),\overline{R}-\mathrm{H}(\omega)\right\}+o(n)}\\ &\leq 2^{n\left(\overline{R}-\mathrm{H}(\omega)\right)/3+o(n)}\\ &\leq 2^{-n(R-R^*)/3+o(n)}\,,\end{aligned}$$

the last inequality following from the fact that $\omega\in\Omega_{R^*}$. ■

Finally, observe that, from Proposition 2 and the continuity of $\delta_8(R)$, it follows that the normalized minimum distance of the typical group code of design rate R is upper-bounded by $\delta_8(R)$, thus completing the proof of Theorem 1. From the bound

$$p_e(\mathcal{G}_n^R)\geq 2^{-\,\mathrm{d}_{\min}(\mathcal{G}_n^R)+o(n)}\,,$$

it also follows that the error exponent of the typical group code does not exceed $E^x(R)$ for every design rate $R\leq R_8^x$.

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] R. Blahut, *Principles and practice of information theory*, Addison-Wesley, 1987.

[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*, vol. 55, pp. 2037-2054, 2009.

[9] G. Como, and F. Fagnani, "Average spectra and minimum distances of low-density parity-check codes over Abelian groups", *SIAM J. Discr. Math.*, vol. 23, pp. 19-53, 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] R.L. Dobrushin, "Asymptotic optimality of group and systematic codes for some channels", *Theor. Probab. Appl.*, vol. 8, pp. 47-59, 1963.

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

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

[16] 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.

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

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

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

[20] 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.

[21] F. Garin, and F. Fagnani, "Analysis of serial turbo codes over Abelian groups for symmetric channels", *SIAM J. Discr. Math.*, vol. 22, pp. 1488-1526, 2008.

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

[23] V.D. Goppa, "Bounds for codes", *Dokl. Acad. Nauk.*, vol. 333, p. 423, 1993.

[24] 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.

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

[26] 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.

[27] 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. Asymptotic Improvement of the Gilbert-Varshamov

[28] T. Jiang, and A. Vardy, "Asymptotic Improvement of the Gilbert-Varshamov Bound on the Size of Binary Codes", *IEEE Trans. Inf. Theory*, vol. 50, pp. 1655-1664, 2004.

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

[30] M. Mézard, and A. Montanari, *Information, Physics, and computation*, Oxford University Press, Oxford, 2009.

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

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

[33] J.N. Pierce, "Limit distribution of the minimum distance of random linear codes", *IEEE Trans. Inf. Theory*, vol. 13, pp. 595-599, 1967.

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

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

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

[37] A. Vardy, "What's new and exciting in algebraic and combinatorial coding theory?", *Plenary Lecture at ISIT 2006*, [online] av. at http://media.itsoc.org/isit2006/vardy/

[38] R.R. Varshamov, "Estimate of the number of signals in error correcting codes", *Dokl. Acad. Nauk.*, vol. 117, pp. 739-741, 1957.
[39] A.J. Viterbi, and J. Omura, *Principles of digital communication and coding*, McGraw-Hill, New York, 1979.

Giacomo Como Giacomo Como received the BSc, MS and PhD degrees in Applied Mathematics from Politecnico di Torino in 2002, 2004 and 2008, respectively. In 2006-07 he was Visiting Assistant in Research at the Department of Electrical Engineering, Yale University. He is currently a Postdoctoral Associate at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA. His current research interests include information theory, coding theory, and distributed estimation and control.