Spatially-Coupled Code Design for Partial-Response Channels: Optimal Object-Minimization Approach

Ahmed Hareedy, Homa Esfahanizadeh, Andrew Tan, and Lara Dolecek
Electrical and Computer Engineering Department, University of California, Los Angeles, Los Angeles, CA 90095 USA
{ahareedy, hesfahanizadeh, andrewyt}@ucla.edu, and dolecek@ee.ucla.edu

***Abstract*—Because of their capacity-approaching performance and their complexity/latency advantages, spatially-coupled (SC) codes are among the most attractive error-correcting codes for use in modern dense storage devices. SC codes are constructed by partitioning an underlying block code and coupling the partitioned components. Here, we focus on circulant-based SC codes. Recently, the optimal overlap (OO)-circulant power optimizer (CPO) approach was introduced to construct high performance SC codes for additive white Gaussian noise (AWGN) and Flash channels. The OO stage operates on the protograph of the SC code to derive the optimal partitioning that minimizes the number of detrimental objects. Then, the CPO optimizes the circulant powers to further reduce this number. Since the nature of detrimental objects in the graph of a code critically depends on the characteristics of the channel of interest, extending the OO-CPO approach to construct SC codes for channels with intrinsic memory is not a straightforward task. In this paper, we tackle one relevant extension; we construct high performance SC codes for practical 1-D magnetic recording channels, i.e., partial-response (PR) channels. Via combinatorial techniques, we carefully build and solve the optimization problem of the OO partitioning, focusing on the objects of interest in the case of PR channels. Then, we customize the CPO to further reduce the number of these objects in the graph of the code. SC codes designed using the proposed OO-CPO approach for PR channels outperform prior state-of-the-art SC codes by around 3 orders of magnitude in frame error rate (FER) and 1.1 dB in signal-to-noise ratio (SNR), and more intriguingly, outperform structured block codes of the same length by around 1.6 orders of magnitude in FER and 0.4 dB in SNR.**

I. Introduction

Similar to other data storage systems, magnetic recording (MR) systems operate at very low frame error rate (FER) levels [1], [2]. Consequently, to guarantee high error correction capability in such systems, binary [2] and non-binary (NB) [3], [4] graph-based codes are used. The objects that dominate the error floor region of low-density parity-check (LDPC) codes simulated in partial-response (PR) and additive white Gaussian noise (AWGN) systems are different in their combinatorial nature because of the detector-decoder looping and the intrinsic memory in PR systems [5]. In particular, the authors in [5] introduced balanced absorbing sets (BASs) to characterize the detrimental objects in the case of PR (1-D MR) channels. Moreover, the weight consistency matrix (WCM) framework was introduced to systematically remove any type of absorbing sets from the graph of an NB-LDPC code [6], [7].

Spatially-coupled (SC) codes [8], [9] are graph-based codes constructed by partitioning an underlying block code into components of the same size, then rewiring these components multiple times [10]. In this work, the underlying block codes, and consequently our constructed SC codes, are circulant-based (CB) codes. SC codes offer not only complexity/latency gains (if windowed decoding [11] is used), but also an additional degree of freedom in the code design; this added flexibility is achieved via partitioning of the parity check matrix of the underlying block code. This observation makes SC codes receive an increasing level of attention in multiple applications. Contiguous [10] and non-contiguous [12], [13] partitioning schemes were introduced in the literature for various applications. Recently, the optimal overlap (OO)-circulant power optimizer (CPO) approach was introduced to design SC codes with superior performance for AWGN [14] and practical asymmetric Flash [15] channels. The OO partitioning operates on the protograph to compute the optimal set of overlap parameters that characterizes the partitioning. The CPO operates on the unlabeled graph (weights are set to 1's) to adjust the circulant powers. The objective is to minimize the number of instances of a common substructure that exists in different detrimental objects. If the SC code is binary, the unlabeled graph is the final graph. If the SC code is non-binary, the WCM framework [6], [7] is used to optimize the edge weights after applying the OO-CPO approach.

In this paper, we propose an approach based on tools from combinatorics, optimization, and graph theory, to construct high performance SC codes for PR channels. Unlike the case of AWGN and Flash channels (see [14] and [15]), the common substructure, whose number of instances we seek to minimize, in the case of PR channels can appear in different ways in the protograph of the SC code, making the optimization problem considerably more challenging. For that reason, we introduce the concept of the *pattern*, which is a configuration in the protograph that can result in instances of the common substructure in the unlabeled graph of the SC code after lifting. We derive an optimization problem, in which we express the weighted sum of the counts (numbers of instances) of all patterns in terms of the overlap parameters. Then, we compute the optimal set of overlap parameters (OO) that minimizes this sum. Moreover, we propose the necessary modifications to the CPO algorithm presented in [14] and [15] to make it suitable for the common substructure in the case of PR channels. We demonstrate the gains achieved by our OO-CPO (-WCM for NB SC codes) approach through tables and performance plots that compare our codes not only with SC codes, but also with CB block codes of the same length and the same rate.

The rest of the paper is organized as follows. Section II introduces the necessary preliminaries. Different patterns of the common substructure are discussed in Section III. The analysis of the optimization problem is presented in Section IV. The needed modifications over the baseline CPO are

detailed in Section V. We present our experimental results in Section VI. Finally, the work is concluded in Section VII.

II. Preliminaries

In this section, we review the construction of SC codes and the definitions of the objects of interest. Here, each row (resp., column) in a parity-check matrix corresponds to a check node (CN) (resp., variable node (VN)) in the equivalent Tanner graph of the matrix. Additionally, each non-zero entry in a parity-check matrix corresponds to an edge in the equivalent Tanner graph of the matrix.

Since the contribution of this work (the OO-CPO) is to optimize the topology of the underlying graph, we will focus on the unlabeled graphs and binary matrices. Labeled graphs and non-binary matrices will be discussed as needed. Let $\mathbf{H}$ be the binary parity-check matrix of the underlying regular CB code that has column weight (VN degree) γ and row weight (CN degree) κ. This matrix consists of $\gamma\kappa$ circulants. Each circulant is of the form $\sigma^{f_{i,j}}$, where $0 \leq i \leq \gamma-1$, $0 \leq j \leq \kappa-1$, and σ is the $z \times z$ identity matrix cyclically shifted one unit to the left. Circulant powers are $f_{i,j}$, $\forall i,j$, and they are defined, in addition to z, as the lifting parameters. Separable CB (SCB) codes have $f_{i,j} = f(i)f(j)$, $z \geq \kappa$, and z prime. The underlying block codes we use to design SC codes in this work are CB codes with no zero circulants.

The binary SC code is constructed as follows. First, $\mathbf{H}$ is partitioned into $(m+1)$ disjoint components (they all have the same size as $\mathbf{H}$): $\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_m$, where m is defined as the memory of the SC code. Each component $\mathbf{H}_y$, $0 \leq y \leq m$, contains some of the $\gamma\kappa$ circulants of $\mathbf{H}$ and zero circulants elsewhere such that $\mathbf{H} = \sum_{y=0}^{m} \mathbf{H}_y$. While we have completed the framework for any m and $\gamma \geq 3$, for simplicity, we focus on $m = 1$ in this paper, i.e., $\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_1$. Then, $\mathbf{H}_0$ and $\mathbf{H}_1$ are coupled L times (see [14, Fig. 1]) to construct the binary parity-check matrix of the SC code, $\mathbf{H}_{\text{SC}}$, which is of size $\gamma z(L+1) \times \kappa z L$. A replica is any $\gamma z(L+1) \times \kappa z$ submatrix of $\mathbf{H}_{\text{SC}}$ that contains $\left[\mathbf{H}_0^{\text{T}}\ \mathbf{H}_1^{\text{T}}\right]^{\text{T}}$ and zero circulants elsewhere. Replicas are denoted by $\mathbf{R}_\rho$, $1 \leq \rho \leq L$.

The protograph matrix (PM) of a binary CB matrix is the matrix resulting from replacing each $z \times z$ non-zero circulant with 1, and each $z \times z$ zero circulant with 0. The PMs of $\mathbf{H}$, $\mathbf{H}_0$, and $\mathbf{H}_1$ are $\mathbf{H}^{\text{p}}$, $\mathbf{H}_0^{\text{p}}$, and $\mathbf{H}_1^{\text{p}}$, respectively, and they are all of size $\gamma \times \kappa$. The PM of $\mathbf{H}_{\text{SC}}$ is $\mathbf{H}_{\text{SC}}^{\text{p}}$, and it is of size $\gamma(L+1) \times \kappa L$. This $\mathbf{H}_{\text{SC}}^{\text{p}}$ also has L replicas, $\mathbf{R}_\rho$, $1 \leq \rho \leq L$, but with 1×1 circulants. Non-binary SC (NB-SC) codes can be constructed from binary SC codes as described in [15] and guided by [7]. The NB codes we use in this work have parity-check matrices with their elements in GF(q), and $q = 2^\lambda$, where $\lambda \in \{2, 3, \ldots\}$ (in the binary case, $q = 2$).

A contiguous technique for partitioning $\mathbf{H}$ to construct $\mathbf{H}_{\text{SC}}$, namely cutting vector (CV) partitioning, was investigated aiming to generate SC codes for PR channels [10]. Multiple non-contiguous partitioning techniques were recently introduced in the literature, e.g., minimum overlap (MO) partitioning [12], general edge spreading [13], in addition to OO partitioning [14], [15]. These non-contiguous partitioning techniques significantly outperform contiguous ones [12], [14], [15]. However, as far as we know, no prior work has proposed non-contiguous techniques in the context of PR channels. The goal of this work is to derive the effective OO-CPO approach for partitioning and lifting to construct high performance SC codes optimized for PR channels.

The intrinsic memory along with detector-decoder iterations (global iterations) result in changing the combinatorial properties of detrimental objects in LDPC codes simulated over PR channels compared with the case of canonical channels [5]. In particular, these detrimental objects were shown to be absorbing sets that can have unsatisfied CNs with degrees > 1, while having a fewer number of unsatisfied (particularly degree-1) CNs. These objects were named BASs. We now present the definitions of different objects of interest. Examples on these objects of interest are in Fig. 1. Let $g = \lfloor \frac{\gamma-1}{2} \rfloor$.

Definition 1. *Consider a subgraph induced by a subset $\mathcal{V}$ of VNs in the Tanner graph of a code. Set all the VNs in $\mathcal{V}$ to values $\in$ GF(q)$\backslash\{0\}$ and set all other VNs to 0. The set $\mathcal{V}$ is said to be an (a, b, d_1, d_2, d_3)* ***balanced absorbing set of type two (BAST)*** *over GF(q) if the size of $\mathcal{V}$ is a, the number of unsatisfied CNs connected to $\mathcal{V}$ is b, $0 \leq b \leq \lfloor \frac{ag}{2} \rfloor$, the number of degree-1 (resp., 2 and > 2) CNs connected to $\mathcal{V}$ is d_1 (resp., d_2 and d_3), $d_2 > d_3$, all the unsatisfied CNs connected to $\mathcal{V}$ (if any) have either degree 1 or degree 2, and each VN in $\mathcal{V}$ is connected to strictly more satisfied than unsatisfied neighboring CNs (for some set of VN values).*

While the above definition was introduced in the context of non-binary codes [5], it is valid in the binary case as well (set $q = 2$). An (a, d_1, d_2, d_3) unlabeled BAST (UBS) is a BAST with the weights of all edges of its graph replaced by 1's. All our abbreviations are short-handed for simplicity.

Definition 2. *Let $\mathcal{V}$ be a subset of VNs in the unlabeled (binary) Tanner graph of a code. Let $\mathcal{O}$ (resp., $\mathcal{T}$ and $\mathcal{H}$) be the set of degree-1 (resp., 2 and > 2) CNs connected to $\mathcal{V}$. This graphical configuration is an (a, d_1)* ***unlabeled elementary trapping set (UTS)*** *if $|\mathcal{V}| = a$, $|\mathcal{O}| = d_1$, and $|\mathcal{H}| = 0$. A UTS is an* ***unlabeled elementary absorbing set (UAS)*** *if each VN in $\mathcal{V}$ is connected to strictly more neighbors in $\mathcal{T}$ than in $\mathcal{O}$.*

A binary protograph configuration is also defined by (a, d_1) for simplicity. The WCM framework removes a BAST from the graph of an NB code by careful processing of its edge weights (see [5], [6], and [7] for details).

III. The Common Substructure and Its Patterns

The idea of focusing on a common substructure in the design of the unlabeled graph of an SC code simplifies the optimization procedure. Additionally, minimizing the number of instances of the common substructure significantly reduces the multiplicity of several different types of detrimental objects simultaneously [10], [14], which is a lot more feasible compared with operating on all these detrimental objects separately (especially for partitioning). It was shown in [10] that the $(4, 4(\gamma-2))$ UAS/UTS, $\gamma \geq 3$, is the common substructure of interest for PR channels (unlike the case for AWGN [13], [14] and Flash channels [15], where the substructure of interest is

the $(3, 3(\gamma-2)))$. Fig. 1 shows UBSs of multiple detrimental BASTs for codes with $\gamma \in \{3, 4\}$ simulated over PR channels, demonstrating that the common substructure of interest is the $(4, 4(\gamma-2))$ UAS/UTS.

We note that the $(4, 4(\gamma-2))$ UAS/UTS is a cycle of length 8 ***with no internal connections*** (ignore degree-1 CNs). From [16] (see also [15]), it is known that each cycle in the unlabeled graph (the graph of $\mathbf{H}_{\text{SC}}$) is derived from a configuration in the protograph (the graph of the PM $\mathbf{H}^{\text{p}}_{\text{SC}}$) under specific conditions on the powers of the circulants involved in that cycle. Thus, in the OO stage, we operate on the protograph. Then, in the CPO stage, we operate on the circulant powers.

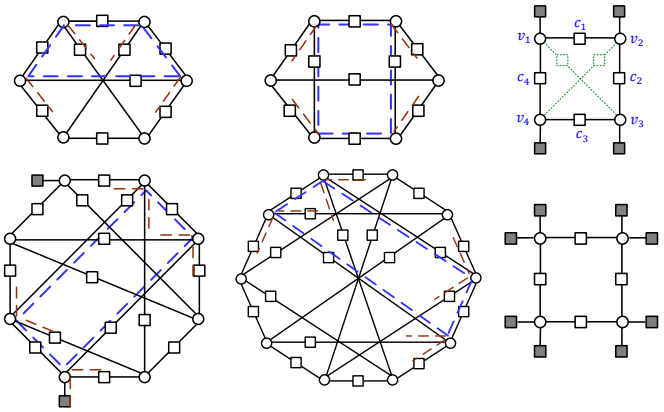


Fig. 1. The UBSs of multiple detrimental BASTs and the associated common substructures. Upper panel ($\gamma = 3$): two non-isomorphic $(6, 0, 9, 0)$ UBSs, and the common substructure is the $(4, 4)$ UAS. Lower panel ($\gamma = 4$): an $(8, 2, 15, 0)$ UBS and a $(10, 0, 20, 0)$ UBS, and the common substructure is the $(4, 8)$ UTS. Common substructures are marked with dashed lines. Internal connections in a cycle of length 8 are shown in dotted green lines.

Remark 1. *Let x^- be an integer s.t. $0 \leq x^- \leq x$. Note that a $(4^-, 4(\gamma-2)^-)$ configuration in the protograph of the code can result in $(4, 4(\gamma-2))$ UASs/UTSs in the unlabeled graph depending on the circulant power arrangement. Thus, in the OO stage, we operate on all protograph configurations that can result in $(4, 4(\gamma-2))$ UASs/UTSs (cycles of length 8 with no internal connections) in the unlabeled graph, including the protograph configurations that do have internal connections. Then in the CPO stage, we treat the $(4, 4(\gamma-2))$ UASs/UTSs and the $(4, 4(\gamma-2)-2\delta)$ UASs/UTSs differently, where $\delta \in \{1, 2\}$ is the number of existing internal connections in the configuration after lifting.*

The major difference between the $(4, 4(\gamma-2))$ UAS/UTS and the $(3, 3(\gamma-2))$ UAS/UTS is that there are multiple configurations in the protograph that can generate the former object in the unlabeled graph. We call these different configurations ***patterns***. A pattern is defined by the dimensions of the matrix of its subgraph. The following lemma investigates the number and nature of these patterns.

Lemma 1. *The number of distinct patterns (with different dimensions) in the protograph of a code that can result in $(4, 4(\gamma-2))$ UASs/UTSs in the unlabeled graph of the code after lifting is 9, in the case of $\gamma \geq 4$. The numbers of CNs and VNs in these 9 patterns are both in $\{2, 3, 4\}$. This number of distinct patterns reduces to 7 in the case of $\gamma = 3$.*

Proof. Since the objects of interest in the unlabeled graph are cycles of length 8 with 4 CNs and 4 VNs, a protograph pattern that can generate some of them must have at most 4 CNs and 4 VNs. Moreover, to result in cycles of length 8 after lifting, the pattern must have at least 2 CNs and 2 VNs. Combining these two statements yields that the numbers of CNs and VNs of a protograph pattern that can result in $(4, 4(\gamma-2))$ UASs/UTSs in the unlabeled graph must be in $\{2, 3, 4\}$.

Consequently, in order to have 9 patterns for the case of $\gamma \geq 4$, we show that selecting any number of CNs in $\{2, 3, 4\}$ and any number of VNs in $\{2, 3, 4\}$ can result in a pattern (or more) that is capable of generating cycles of length 8 in the unlabeled graph. Fig. 2 demonstrates the above statement, focusing on the matrix representation of patterns and cycles. In the case of $\gamma = 3$, a pattern cannot have 4 ones in a column, which reduces the number of patterns to 7. ■

We define the 9 patterns according to the dimensions of their submatrices as follows. Pattern P_1 is 2×2, Pattern P_2 is 2×3, Pattern P_3 is 3×2, Pattern P_4 is 2×4, Pattern P_5 is 4×2, Pattern P_6 is 3×3, Pattern P_7 is 3×4, Pattern P_8 is 4×3, and Pattern P_9 is 4×4 (all illustrated in Fig. 2).

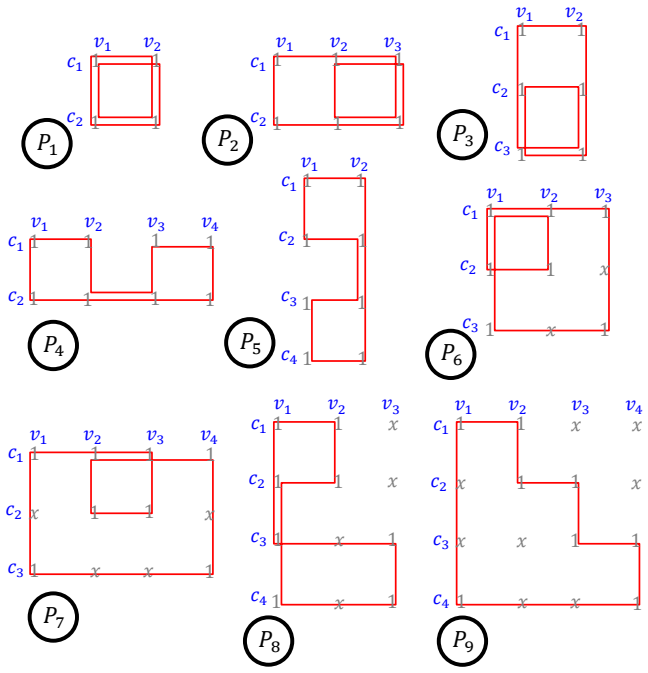


Fig. 2. The 9 protograph patterns that can result in cycles of length 8 in the unlabeled graph after lifting. One way of traversing each pattern to generate cycles of length 8 is depicted in red. Note that only Pattern P_9 represents a cycle of length 8 in the protograph.

Remark 2. *Following the same logic we used in Lemma 1 and its proof for the $(3, 3(\gamma-2))$ UAS/UTS, leads to a possibility to also have patterns for this case, with the number of CNs and VNs in $\{2, 3\}$. However, a careful analysis guides to the fact that only one protograph pattern can result in $(3, 3(\gamma-2))$ UASs/UTSs (cycles of length 6) after lifting, which is the 3×3 pattern, and it is itself a cycle of length 6 [14], [15].*

The following lemma discusses the relation between different protograph patterns and the resulting cycles after lifting. Define a ***cycle-8 candidate*** of Pattern P_ℓ as a way to traverse

P_ℓ in order to reach a cycle of length 8 in the unlabeled graph of the code after lifting. Some candidates are shown in Fig. 2.

Lemma 2. *Let ζ_{P_ℓ} be the number of distinct cycle-8 candidates of Pattern P_ℓ. Then,*

$$\zeta_{P_\ell} = \begin{cases} 1, & \ell \in \{1,6,9\}, \\ 2, & \ell \in \{7,8\}, \\ 3, & \ell \in \{2,3\}, \\ 6, & \ell \in \{4,5\}. \end{cases} \tag{1}$$

Proof. We define a cycle-8 candidate according to the connectivity as follows: $c_1-v_1-c_2-v_2-c_3-v_3-c_4-v_4$ (each CN connects the next two VNs in a circular fashion, see Fig. 1). From Fig. 2, there is only one cycle-8 candidate for Pattern P_1, which is $c_1-v_1-c_2-v_2-c_1-v_1-c_2-v_2$, and this is the case for all square patterns. Thus, $\zeta_{P_\ell} = 1$ for $\ell \in \{1,6,9\}$. It can be understood from Fig. 2 that $\zeta_{P_\ell} \neq 1$ for all the remaining patterns. In particular, we have two cycle-8 candidates for Pattern P_7, that are: $c_1-v_1-c_2-v_2-c_1-v_3-c_3-v_4$ and $c_1-v_1-c_2-v_3-c_1-v_2-c_3-v_4$ (which is the red cycle on P_7 in Fig. 2). The situation is the same for Pattern P_8 because it is the transpose of P_7. Thus, $\zeta_{P_\ell} = 2$ for $\ell \in \{7,8\}$. The rest of the cases can be derived similarly. ∎

Pattern P_1 has $\zeta_{P_\ell} = 1$ (see (1)), and it results in $z/2$ or 0 cycles of length 8 after lifting (since P_1 is only 2×2), while all the remaining patterns result in z or 0 cycles of length 8 after lifting [15], [16]. Thus, we define the ***pattern weight***, β_{P_ℓ}, which plays an important role in the discrete optimization problem of the OO, as follows:

$$\beta_{P_\ell} = \begin{cases} 1/2, & \ell = 1, \\ \zeta_{P_\ell}, & \ell \in \{2,3,4,5,6,7,8,9\}. \end{cases} \tag{2}$$

IV. OO: Building and Solving the Optimization Problem

Now, we are ready to build the optimization problem. Consider the protograph of an SC code. The **weighted sum** of the total number of instances of all patterns is given by:

$$F_{\text{tot}} = \sum_{\ell=1}^{9} \beta_{P_\ell} F_{P_\ell}, \tag{3}$$

where F_{P_ℓ} is the total number of instances of Pattern P_ℓ. The goal is to express F_{tot}, through F_{P_ℓ}, $\forall \ell$, as a function of the overlap parameters, then finding the optimal set of overlap parameters that minimize F_{tot} for OO partitioning. We first recall the definition and the properties of overalp parameters. More details on that part can be found in [14].

Definition 3. *For $m = 1$, let $\mathbf{\Pi} = \left[\mathbf{H}_0^{\mathrm{T}}\ \mathbf{H}_1^{\mathrm{T}}\right]^{\mathrm{T}}$, and let $\mathbf{\Pi}^{\mathrm{p}}$ be its PM (of size $2\gamma \times \kappa$). A **degree-μ overlap** among μ rows (or CNs) of $\mathbf{\Pi}^{\mathrm{p}}$ indexed by $\{i_1, \ldots, i_\mu\}$, $1 \leq \mu \leq \gamma$, $0 \leq i_1, \ldots, i_\mu \leq 2\gamma - 1$, is defined as a position (column) in which all these rows have 1's simultaneously. A **degree-μ overlap parameter**, $t_{\{i_1,\ldots,i_\mu\}}$, is defined as the number of degree-μ overlaps among the rows indexed by $\{i_1, \ldots, i_\mu\}$ in $\mathbf{\Pi}^{\mathrm{p}}$. A degree-1 overlap parameter t_{i_1}, $0 \leq i_1 \leq 2\gamma - 1$, is defined as the number of 1's in row i_1 of $\mathbf{\Pi}^{\mathrm{p}}$.*

Note that a degree-μ overlap parameter, if $\mu > 1$, is always zero if in the set $\{i_1, \ldots, i_\mu\}$ there exists at least one pair of distinct row indices, say (i_{τ_1}, i_{τ_2}), with the property that $i_{\tau_1} \equiv i_{\tau_2} \pmod{\gamma}$ [14]. Define the set of non-zero overlap parameters as $\mathcal{O}$. The parameters in $\mathcal{O}$ are not all independent. The set of independent non-zero overlap parameters, $\mathcal{O}_{\text{ind}}$, is:

$$\begin{aligned} \mathcal{O}_{\text{ind}} = \{ & t_{\{i_1,\ldots,i_\mu\}} \text{ s.t. } 1 \leq \mu \leq \gamma,\ 0 \leq i_1, \ldots, i_\mu \leq \gamma - 1, \\ & \forall \{i_{\tau_1}, i_{\tau_2}\} \subset \{i_1, \ldots, i_\mu\}, i_{\tau_1} \not\equiv i_{\tau_2} \pmod{\gamma}\}. \end{aligned} \tag{4}$$

The other non-zero overlap parameters in $\mathcal{O} \backslash \mathcal{O}_{\text{ind}}$ are obtained from the parameters in $\mathcal{O}_{\text{ind}}$ according to [14, Lemma 3]. Since we are focusing on $m = 1$, the cardinality of the set $\mathcal{O}_{\text{ind}}$, which determines the complexity of the discrete optimization problem of the OO stage, is given by:

$$\mathcal{N}_{\text{ind}} = |\mathcal{O}_{\text{ind}}| = \sum_{\mu=1}^{\gamma} \binom{\gamma}{\mu} \tag{5}$$

As demonstrated in Fig. 2, for all the patterns of interest, the highest overlap degree is $\mu = 4$ (a pattern has at most 4 CNs). Note that while the overlap parameters themselves must be restricted to $\mathbf{\Pi}^{\mathrm{p}}$, the concept of the degree-μ overlap can be generalized from $\mathbf{\Pi}^{\mathrm{p}}$ to the PM of the SC code, $\mathbf{H}_{\text{SC}}^{\mathrm{p}}$. We will use this generalization in the analysis of patterns.

We aim at expressing F_{P_ℓ}, $\forall \ell$, in terms of the parameters in $\mathcal{O}_{\text{ind}}$. Let $\mathbf{R}_r$ be a replica in which at least one VN of the pattern being studied exists. We call $\mathbf{R}_r$ the reference replica. Moreover, let the CNs (or rows) of the pattern be of the form $c_k = (r-1)\gamma + i_k$, $1 \leq k \leq 4$. For two replicas $\mathbf{R}_\rho$ and $\mathbf{R}_\nu$, define $\theta_{\rho,\nu} = \rho - \nu$. In the following, we consider the protograph of an SC code with parameters $\gamma \geq 3$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$. We define $[x]^+ = \max\{x, 0\}$, and $F_{P_\ell,1}^k$ as the number of instances of Pattern P_ℓ that starts at replica $\mathbf{R}_1$ and spans k consecutive replicas. Note that each VN in a pattern corresponds to an overlap (see Fig. 2).

The counts of different existence possibilities of the nine patterns in addition to the final formulas of F_{P_ℓ}, $\forall \ell$, are presented in the forthcoming subsections.

A. Analysis of Pattern P_1 (size 2×2)

This pattern has two VNs that are *adjacent* (connected via at least one path of only one CN). Thus, Pattern P_1 has its VNs located in at most two replicas, and the pattern spans (i.e., its VNs span) at most $m + 1 = 2$ consecutive replicas (see [14, Lemma 1]). Suppose P_1 has the CNs c_1 and c_2. The two overlaps forming the pattern are of degree-2, and they are both $c_1 - c_2$ overlaps (among c_1 and c_2).

Lemma 3. *Case 1.1: The number of instances of P_1 with CNs c_1 and c_2, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\mathcal{A}_{P_1}\left(t_{\{i_1,i_2\}}\right) = \binom{t_{\{i_1,i_2\}}}{2}. \tag{6}$$

Case 1.2: The number of instances of P_1 with CNs c_1 and c_2, and all overlaps in two replicas, $\mathbf{R}_r$ and $\mathbf{R}_e$, $r < e$, is:

$$\begin{aligned} &\mathcal{B}_{P_1}\left(t_{\{i_1,i_2\}}, t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}\right) \\ &\quad = t_{\{i_1,i_2\}} t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}. \end{aligned} \tag{7}$$

Proof. In Case 1.1, the number we are after is the number of ways to choose 2 overlaps out of $t_{\{i_1,i_2\}}$ overlaps, which is given by (6). In Case 1.2, the number we are after is the number of ways to choose 1 overlap out of $t_{\{i_1,i_2\}}$ and 1 overlap out of $t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma\}}$, which is given by (7). ■

Theorem 1. *The total number of instances of Pattern P_1 in the binary protograph of an SC code that has parameters $\gamma \geq 3$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_1} = \sum_{k=1}^{2} (L-k+1) F_{P_1,1}^{k}, \tag{8}$$

where $F_{P_1,1}^{k}$, $k \in \{1,2\}$, are given by:

$$\begin{aligned} F_{P_1,1}^{1} &= \sum_{\{i_1,i_2\} \subset \{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_1}\left(t_{\{i_1,i_2\}}\right), \\ F_{P_2,1}^{2} &= \sum_{\{i_1,i_2\} \subset \{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_1}\left(t_{\{i_1,i_2\}}, t_{\{i_1-\gamma,i_2-\gamma\}}\right), \end{aligned} \tag{9}$$

with $\overline{i_1} \neq \overline{i_2}$, where $\overline{i} = (i \bmod \gamma)$.

Proof. To compute F_{P_1}, we use the formula in [14, Theorem 1], with χ, which is the maximum number of replicas the pattern can span, equals 2. Then, $F_{P_1,1}^{1}$ (resp., $F_{P_1,1}^{2}$) is the sum of function $\mathcal{A}_{P_1}$ (resp., $\mathcal{B}_{P_1}$) over all possible values of $\{i_1, i_2\}$ (the degree-2 overlap indices in $\mathbf{\Pi}^{\mathrm{p}}$). ■

B. Analysis of Pattern P_2 (size 2×3)

This pattern has three VNs, with each pair of them being adjacent. Thus, P_2 spans at most 2 consecutive replicas, which also means its VNs are located in at most 2 replicas. Suppose P_2 has the CNs c_1 and c_2. The three overlaps forming P_2 are of degree-2, and they are all $c_1 - c_2$ overlaps.

Lemma 4. *Case 2.1: The number of instances of P_2 with CNs c_1 and c_2, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\mathcal{A}_{P_2}\left(t_{\{i_1,i_2\}}\right) = \binom{t_{\{i_1,i_2\}}}{3}. \tag{10}$$

Case 2.2: The number of instances of P_2 with CNs c_1 and c_2, and all overlaps in two replicas, s.t. two overlaps are in $\mathbf{R}_r$, and one overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned} &\mathcal{B}_{P_2}\left(t_{\{i_1,i_2\}}, t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma\}}\right) \\ &\quad = \binom{t_{\{i_1,i_2\}}}{2} t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma\}}. \end{aligned} \tag{11}$$

Proof. In Case 2.1, the number we are after is the number of ways to choose 3 overlaps out of $t_{\{i_1,i_2\}}$ overlaps, which is given by (10). In Case 2.2, the number we are after is the number of ways to choose 2 overlap out of $t_{\{i_1,i_2\}}$ and 1 overlap out of $t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma\}}$, which is given by (11). ■

Theorem 2. *The total number of instances of Pattern P_2 in the binary protograph of an SC code that has parameters $\gamma \geq 3$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_2} = \sum_{k=1}^{2} (L-k+1) F_{P_2,1}^{k}, \tag{12}$$

where $F_{P_2,1}^{k}$, $k \in \{1,2\}$, are given by:

$$\begin{aligned} F_{P_2,1}^{1} &= \sum_{\{i_1,i_2\} \subset \{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_2}\left(t_{\{i_1,i_2\}}\right), \\ F_{P_2,1}^{2} &= \sum_{\{i_1,i_2\} \subset \{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_2}\left(t_{\{i_1,i_2\}}, t_{\{i_1-\gamma,i_2-\gamma\}}\right) \\ &\quad + \sum_{\{i_1,i_2\} \subset \{0,\ldots,\gamma-1\}} \mathcal{B}_{P_2}\left(t_{\{i_1,i_2\}}, t_{\{i_1+\gamma,i_2+\gamma\}}\right), \end{aligned} \tag{13}$$

with $\overline{i_1} \neq \overline{i_2}$.

Proof. To compute F_{P_2}, we use the formula in [14, Theorem 1], with $\chi = 2$. Then, $F_{P_2,1}^{1}$ is the sum of function $\mathcal{A}_{P_2}$ over all possible values of $\{i_1, i_2\}$. Regarding $F_{P_2,1}^{2}$, we need to distinguish between two situations; when $r < e$ (i.e., replica $\mathbf{R}_r$, which has two overlaps, comes before replica $\mathbf{R}_e$), and when $r > e$ (i.e., replica $\mathbf{R}_r$ comes after replica $\mathbf{R}_e$). This distinction gives the two summations of function $\mathcal{B}_{P_2}$ in (13). ■

C. Analysis of Pattern P_3 (size 3×2)

This pattern has two VNs that are adjacent. Thus, Pattern P_3 spans at most 2 consecutive replicas. Suppose P_3 has the CNs c_1, c_2, and c_3. The two overlaps forming P_3 are of degree-3, and they are all $c_1 - c_2 - c_3$ overlaps.

Lemma 5. *Case 3.1: The number of instances of P_3 with CNs c_1, c_2, and c_3, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\mathcal{A}_{P_3}\left(t_{\{i_1,i_2,i_3\}}\right) = \binom{t_{\{i_1,i_2,i_3\}}}{2}. \tag{14}$$

Case 3.2: The number of instances of P_3 with CNs c_1, c_2, and c_3, and all overlaps in two replicas, $\mathbf{R}_r$ and $\mathbf{R}_e$, $r < e$, is:

$$\begin{aligned} &\mathcal{B}_{P_3}\left(t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}\right) \\ &\quad = t_{\{i_1,i_2,i_3\}} t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}. \end{aligned} \tag{15}$$

Proof. In Case 3.1, $\mathbf{R}_r$, the number we are after is the number of ways to choose 2 overlaps out of $t_{\{i_1,i_2,i_3\}}$, which is given by (14). In Case 3.2, the number we are after is the number of ways to choose 1 overlap out of $t_{\{i_1,i_2,i_3\}}$ and 1 overlap out of $t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}$, which is given by (15). ■

Theorem 3. *The total number of instances of Pattern P_3 in the binary protograph of an SC code that has parameters $\gamma \geq 3$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_3} = \sum_{k=1}^{2} (L-k+1) F_{P_3,1}^{k}, \tag{16}$$

where $F_{P_3,1}^{k}$, $k \in \{1,2\}$, are given by:

$$\begin{aligned} F_{P_3,1}^{1} &= \sum_{\{i_1,i_2,i_3\} \subset \{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_3}\left(t_{\{i_1,i_2,i_3\}}\right), \\ F_{P_3,1}^{2} &= \sum_{\{i_1,i_2,i_3\} \subseteq \{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_3}\left(t_{\{i_1,i_2,i_3\}}, t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma\}}\right), \end{aligned} \tag{17}$$

with $\overline{i_1} \neq \overline{i_2}$, $\overline{i_1} \neq \overline{i_3}$, and $\overline{i_2} \neq \overline{i_3}$.

Proof. To compute F_{P_3}, we use the formula in [14, Theorem 1], with $\chi = 2$. Then, $F^1_{P_3,1}$ (resp., $F^2_{P_3,1}$) is the sum of function $\mathcal{A}_{P_3}$ (resp., $\mathcal{B}_{P_3}$) over all possible values of $\{i_1, i_2, i_3\}$ (the degree-3 overlap indices in $\mathbf{\Pi}^{\text{p}}$). ■

D. Analysis of Pattern P_4 (size 2×4)

This pattern has four VNs, with each pair of them being adjacent. Consequently, P_4 spans at most 2 consecutive replicas. Suppose P_4 has the CNs c_1 and c_2. The four overlaps forming P_4 are of degree-2, and they are all $c_1 - c_2$ overlaps.

Lemma 6. *Case 4.1: The number of instances of P_4 with CNs c_1 and c_2, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\mathcal{A}_{P_4}\left(t_{\{i_1,i_2\}}\right) = \binom{t_{\{i_1,i_2\}}}{4}. \tag{18}$$

Case 4.2: The number of instances of P_4 with CNs c_1 and c_2, and all overlaps in two replicas, s.t. three overlaps are in $\mathbf{R}_r$, and one overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}\mathcal{B}_{P_4}&\left(t_{\{i_1,i_2\}}, t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}\right)\\ &= \binom{t_{\{i_1,i_2\}}}{3} t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}.\end{aligned} \tag{19}$$

Case 4.3: The number of instances of P_4 with CNs c_1 and c_2, and all overlaps in two replicas, s.t. two overlaps are in $\mathbf{R}_r$, and two overlaps are in $\mathbf{R}_e$, $r < e$, is:

$$\begin{aligned}\mathcal{C}_{P_4}&\left(t_{\{i_1,i_2\}}, t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}\right)\\ &= \binom{t_{\{i_1,i_2\}}}{2}\binom{t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}}{2}.\end{aligned} \tag{20}$$

Proof. In Case 4.1, the number we are after is the number of ways to choose 4 overlaps out of $t_{\{i_1,i_2\}}$, which is given by (18). In Case 4.2, the number we are after is the number of ways to choose 3 overlaps out of $t_{\{i_1,i_2\}}$ and 1 overlap out of $t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}$, which is given by (19). In Case 4.3, the number we are after is the number of ways to choose 2 overlaps out of $t_{\{i_1,i_2\}}$ and 2 overlaps out of $t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}}$, which is given by (20). ■

Theorem 4. *The total number of instances of Pattern P_4 in the binary protograph of an SC code that has parameters $\gamma \geq 3$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_4} = \sum_{k=1}^{2} (L - k + 1) F^k_{P_4,1}, \tag{21}$$

where $F^k_{P_4,1}$, $k \in \{1, 2\}$, are given by:

$$\begin{aligned}F^1_{P_4,1} &= \sum_{\{i_1,i_2\}\subset\{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_4}\left(t_{\{i_1,i_2\}}\right),\\ F^2_{P_4,1} &= \sum_{\{i_1,i_2\}\subset\{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_4}\left(t_{\{i_1,i_2\}}, t_{\{i_1-\gamma,i_2-\gamma\}}\right)\\ &\quad + \sum_{\{i_1,i_2\}\subset\{0,\ldots,\gamma-1\}} \mathcal{B}_{P_4}\left(t_{\{i_1,i_2\}}, t_{\{i_1+\gamma,i_2+\gamma\}}\right)\\ &\quad + \sum_{\{i_1,i_2\}\subset\{\gamma,\ldots,2\gamma-1\}} \mathcal{C}_{P_4}\left(t_{\{i_1,i_2\}}, t_{\{i_1-\gamma,i_2-\gamma\}}\right),\end{aligned} \tag{22}$$

with $\overline{i_1} \neq \overline{i_2}$.

Proof. To compute F_{P_4}, we use the formula in [14, Theorem 1], with $\chi = 2$. Then, $F^1_{P_4,1}$ is the sum of function $\mathcal{A}_{P_4}$ over all possible values of $\{i_1, i_2\}$. Regarding $F^2_{P_4,1}$, we need to account for Case 4.2 and Case 4.3. For Case 4.2, we need to distinguish between two situations; when $r < e$ (i.e., replica $\mathbf{R}_r$, which has three overlaps, comes before replica $\mathbf{R}_e$), and when $r > e$ (i.e., replica $\mathbf{R}_r$ comes after replica $\mathbf{R}_e$). This distinction gives the two summations of function $\mathcal{B}_{P_4}$ in (21). This distinction is not needed for Case 4.3 since the two replicas have the same number of degree-2 overlaps. ■

E. Analysis of Pattern P_5 (size 4×2)

This pattern has two adjacent VNs. Thus, Pattern P_5 spans at most 2 consecutive replicas. Pattern P_5 does not exist in the case of $\gamma = 3$. Suppose P_5 has the CNs c_1, c_2, c_3, and c_4. The two overlaps forming P_5 are of degree-4, and they are all $c_1 - c_2 - c_3 - c_4$ overlaps.

Lemma 7. *Case 5.1: The number of instances of P_5 with CNs c_1, c_2, c_3, and c_4, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\mathcal{A}_{P_5}\left(t_{\{i_1,i_2,i_3,i_4\}}\right) = \binom{t_{\{i_1,i_2,i_3,i_4\}}}{2}. \tag{23}$$

Case 5.2: The number of instances of P_5 with c_1, c_2, c_3, and c_4, and all overlaps in two replicas, $\mathbf{R}_r$ and $\mathbf{R}_e$, $r < e$, is:

$$\begin{aligned}\mathcal{B}_{P_5}&\left(t_{\{i_1,i_2,i_3,i_4\}}, t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma, i_4+\theta_{r,e}\gamma\}}\right)\\ &= t_{\{i_1,i_2,i_3,i_4\}} t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma, i_4+\theta_{r,e}\gamma\}}.\end{aligned} \tag{24}$$

Proof. In Case 5.1, $\mathbf{R}_r$, the number we are after is the number of ways to choose 2 overlaps out of $t_{\{i_1,i_2,i_3,i_4\}}$, which is given by (23). In Case 5.2, the number we are after is the number of ways to choose 1 overlap out of $t_{\{i_1,i_2,i_3,i_4\}}$ and 1 overlap out of $t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma, i_4+\theta_{r,e}\gamma\}}$, given in (24). ■

Theorem 5. *The total number of instances of Pattern P_5 in the binary protograph of an SC code that has parameters $\gamma \geq 4$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_5} = \sum_{k=1}^{2} (L - k + 1) F^k_{P_5,1}, \tag{25}$$

where $F^k_{P_5,1}$, $k \in \{1, 2\}$, are given by:

$$\begin{aligned}F^1_{P_5,1} &= \sum_{\{i_1,i_2,i_3,i_4\}\subset\{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_5}\left(t_{\{i_1,i_2,i_3,i_4\}}\right),\\ F^2_{P_5,1} &= \sum_{\{i_1,i_2,i_3,i_4\}\subseteq\{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_5}\left(t_{\{i_1,i_2,i_3,i_4\}}, t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma,i_4-\gamma\}}\right),\end{aligned} \tag{26}$$

with $\overline{i_1} \neq \overline{i_2}$, $\overline{i_1} \neq \overline{i_3}$, $\overline{i_1} \neq \overline{i_4}$, $\overline{i_2} \neq \overline{i_3}$, $\overline{i_2} \neq \overline{i_4}$, and $\overline{i_3} \neq \overline{i_4}$.

Proof. To compute F_{P_5}, we use the formula in [14, Theorem 1], with $\chi = 2$. Then, $F^1_{P_5,1}$ (resp., $F^2_{P_5,1}$) is the sum of function $\mathcal{A}_{P_5}$ (resp., $\mathcal{B}_{P_5}$) over all possible values of $\{i_1, i_2, i_3, i_4\}$ (the degree-4 overlap indices in $\mathbf{\Pi}^{\text{p}}$). ■

F. Analysis of Pattern P_6 (size 3×3)

This pattern has three VNs, with each pair of them being adjacent. Thus, P_6 spans at most 2 consecutive replicas. Suppose P_6 has the CNs c_1, c_2, and c_3. Define ***distinct overlaps*** to be overlaps from different families, i.e., overlaps between different sets of CNs. Pattern P_6 is formed of three overlaps; two (distinct) of degree-2 and one of degree-3. Define c_1 as the CN connecting the three VNs. Thus, the overlaps are $c_1 - c_2$, $c_1 - c_3$, and $c_1 - c_2 - c_3$ (see P_6 in Fig. 2). Again, each VN corresponds to an overlap.

Lemma 8. *Case 6.1: The number of instances of P_6 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\begin{aligned}
&\mathcal{A}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}\right) \\
&\quad = t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_2,i_3\}}-1\right)^+\left(t_{\{i_1,i_2,i_3\}}-2\right)^+ \\
&\quad + t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_3\}}-t_{\{i_1,i_2,i_3\}}\right)\left(t_{\{i_1,i_2,i_3\}}-1\right)^+ \\
&\quad + \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_2,i_3\}}-1\right)^+ \\
&\quad + \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)\left(t_{\{i_1,i_3\}}-t_{\{i_1,i_2,i_3\}}\right)t_{\{i_1,i_2,i_3\}}.
\end{aligned} \tag{27}$$

Case 6.2: The number of instances of P_6 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in two replicas, s.t. the two degree-2 overlaps are in $\mathbf{R}_r$, and the degree-3 overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}
&\mathcal{B}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}\right) \\
&= \left[t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_3\}}-1\right)^+ + \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)t_{\{i_1,i_3\}}\right] \\
&\quad \cdot t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}.
\end{aligned} \tag{28}$$

Case 6.3: The number of instances of P_6 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in two replicas, s.t. the degree-3 overlap and the $c_1 - c_2$ overlap are in $\mathbf{R}_r$, and the $c_1 - c_3$ overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}
&\mathcal{C}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}\right) \\
&\quad = t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_2\}}-1\right)^+ t_{\{i_1+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}.
\end{aligned} \tag{29}$$

Proof. In Case 6.1, the number we are after is the number of ways to choose 1 overlap from each family (there exist three different families for P_6). In order to avoid over-counting, it is required to distinguish between the two situations when a degree-2 overlap ($c_1 - c_2$ or $c_1 - c_3$) is part of a $c_1 - c_2 - c_3$ degree-3 overlap, and when this is not the case. Taking this requirement into account yields the four added terms in (27). The same applies for Case 6.2, with the exception that here the degree-3 overlap is chosen from $t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}$ overlaps, which reduces the number of added terms to the two in (28). Following the same logic gives (29). ■

Theorem 6. *The total number of instances of Pattern P_6 in the binary protograph of an SC code that has parameters $\gamma \geq 3$, κ, $m = 1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_6} = \sum_{k=1}^{2}(L-k+1)F_{P_6,1}^k, \tag{30}$$

where $F_{P_6,1}^k$, $k \in \{1,2\}$, are given by:

$$\begin{aligned}
F_{P_6,1}^1 &= \sum_{i_1\in\{0,\ldots,2\gamma-1\},\{i_2,i_3\}\subset\{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}\right), \\
F_{P_6,1}^2 &= \sum_{i_1\in\{\gamma,\ldots,2\gamma-1\},\{i_2,i_3\}\subset\{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma\}}\right) \\
&\quad + \sum_{i_1\in\{0,\ldots,\gamma-1\},\{i_2,i_3\}\subset\{0,\ldots,\gamma-1\}} \mathcal{B}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\gamma,i_2+\gamma,i_3+\gamma\}}\right) \\
&\quad + \sum_{i_1\in\{\gamma,\ldots,2\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{\gamma,\ldots,2\gamma-1\}} \mathcal{C}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1-\gamma,i_3-\gamma\}}\right) \\
&\quad + \sum_{i_1\in\{0,\ldots,\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{0,\ldots,\gamma-1\}} \mathcal{C}_{P_6}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\gamma,i_3+\gamma\}}\right),
\end{aligned} \tag{31}$$

with $\overline{i_1} \neq \overline{i_2}$, $\overline{i_1} \neq \overline{i_3}$, and $\overline{i_2} \neq \overline{i_3}$.

Proof. To compute F_{P_6}, we use the formula in [14, Theorem 1], with $\chi = 2$. Then, $F_{P_6,1}^1$ is the sum of function $\mathcal{A}_{P_6}$ over all possible values of i_1 and $\{i_2, i_3\}$. In Pattern P_6, CN c_1, which connects all three VNs, is different from the other two CNs. Moreover, in a group of three CNs that can form P_6, c_1 can be any one of these three CNs. These facts are the reason why i_1 of c_1 has to be separated from $\{i_2, i_3\}$, despite having the same range, in the expression of $F_{P_6,1}^1$. Regarding $F_{P_6,1}^2$, we need to account for Case 6.2 and Case 6.3. For each case of the two, we need to distinguish between two situations; when $r < e$ and when $r > e$. This distinction gives the four summations of $F_{P_6,1}^2$ in (31). Note that the ranges of i_2 and i_3 are different in Case 6.3, unlike Case 6.2. ■

G. Analysis of Pattern P_7 (size 3×4)

This pattern has four VNs, with each pair of them being adjacent. Consequently, P_7 spans at most 2 consecutive replicas. Suppose P_7 has the CNs c_1, c_2, and c_3. The pattern is formed of four degree-2 overlaps that are evenly distributed over two different families. Define c_1 as the CN connecting the four VNs. Thus, the overlaps are two $c_1 - c_2$ and two $c_1 - c_3$ overlaps (see P_7 in Fig. 2 for clarification).

Lemma 9. *Case 7.1: The number of instances of P_7 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\begin{aligned}
&\mathcal{A}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}\right) \\
&\quad = \binom{t_{\{i_1,i_2,i_3\}}}{2}\binom{\left(t_{\{i_1,i_3\}}-2\right)^+}{2} \\
&\quad + t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)\binom{\left(t_{\{i_1,i_3\}}-1\right)^+}{2} \\
&\quad + \binom{t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}}{2}\binom{t_{\{i_1,i_3\}}}{2}.
\end{aligned} \tag{32}$$

Case 7.2: The number of instances of P_7 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in two replicas, s.t. three overlaps are in $\mathbf{R}_r$, and one $c_1 - c_3$ overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}
&\mathcal{B}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}\right) \\
&\quad = \Bigg[\binom{t_{\{i_1,i_2,i_3\}}}{2}\left(t_{\{i_1,i_3\}}-2\right)^+ \\
&\quad + t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)\left(t_{\{i_1,i_3\}}-1\right)^+ \\
&\quad + \binom{t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}}{2} t_{\{i_1,i_3\}}\Bigg] t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}.
\end{aligned} \tag{33}$$

Case 7.3: The number of instances of P_7 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in two replicas, s.t. the two c_1-c_2 overlaps are in $\mathbf{R}_r$, and the two c_1-c_3 overlaps are in $\mathbf{R}_e$, $r<e$, is:

$$\begin{aligned}
&\mathcal{C}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}\right) \\
&\quad = \binom{t_{\{i_1,i_2\}}}{2}\binom{t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}}{2}.
\end{aligned} \tag{34}$$

Case 7.4: The number of instances of P_7 with CNs c_1, c_2, and c_3 as defined in the previous paragraph, and all overlaps in two replicas, s.t. two distinct overlaps (from different families) are in $\mathbf{R}_r$, and two distinct overlaps are in $\mathbf{R}_e$, $r<e$, is:

$$\begin{aligned}
&\mathcal{D}_{P_7}\big(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}} \\
&\quad , t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}, t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}\big) \\
&= \left[t_{\{i_1,i_2,i_3\}}\left(t_{\{i_1,i_3\}}-1\right)^+ + \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)t_{\{i_1,i_3\}}\right] \\
&\quad \cdot \Big[t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}\left(t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}-1\right)^+ \\
&\quad + \left(t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma\}} - t_{\{i_1+\theta_{r,e}\gamma, i_2+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}\right) \\
&\quad\quad \cdot t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}\Big].
\end{aligned} \tag{35}$$

Proof. In Case 7.1, the number we are after is the number of ways to choose 2 overlaps from each family (the pattern has two c_1-c_2 overlaps and two c_1-c_3 overlaps). In order to avoid over-counting, it is required to distinguish between the three situations when the two c_1-c_2 overlaps are each part of a $c_1-c_2-c_3$ degree-3 overlap, when only one c_1-c_2 overlap is part of a $c_1-c_2-c_3$ overlap, and when neither of them is. Taking this requirement into account yields the three added terms in (32). The same applies for Case 7.2, with the exception that here one c_1-c_3 overlap is chosen from $t_{\{i_1+\theta_{r,e}\gamma, i_3+\theta_{r,e}\gamma\}}$ overlaps. In Case 7.3, there is no need to make this distinction. Finally, in Case 7.4, the distinction is applied separately on the c_1-c_2 overlap in $\mathbf{R}_r$ and the c_1-c_2 overlap in $\mathbf{R}_e$ to give (35). ∎

Theorem 7. *The total number of instances of Pattern P_7 in the binary protograph of an SC code that has parameters $\gamma \geq 3$, κ, $m=1$, $L \geq 3$, and $\mathcal{O}$, is:*

$$F_{P_7} = \sum_{k=1}^{2}(L-k+1)F_{P_7,1}^k, \tag{36}$$

where $F_{P_7,1}^k$, $k \in \{1,2\}$, are given by:

$$\begin{aligned}
F_{P_7,1}^1 &= \sum_{i_1\in\{0,\ldots,2\gamma-1\},\{i_2,i_3\}\subset\{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}\right), \\
F_{P_7,1}^2 &= \sum_{i_1\in\{\gamma,\ldots,2\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{\gamma,\ldots,2\gamma-1\}} \mathcal{B}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1-\gamma,i_3-\gamma\}}\right) \\
&\quad + \sum_{i_1\in\{0,\ldots,\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{0,\ldots,\gamma-1\}} \mathcal{B}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1+\gamma,i_3+\gamma\}}\right) \\
&\quad + \sum_{i_1\in\{\gamma,\ldots,2\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{\gamma,\ldots,3\gamma-1\}} \mathcal{C}_{P_7}\left(t_{\{i_1,i_2\}}, t_{\{i_1-\gamma,i_3-\gamma\}}\right) \\
&\quad + \sum_{i_1\in\{\gamma,\ldots,2\gamma-1\},\{i_2,i_3\}\subset\{\gamma,\ldots,2\gamma-1\}} \mathcal{D}_{P_7}\big(t_{\{i_1,i_2\}}, t_{\{i_1,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_1-\gamma,i_2-\gamma\}} \\
&\qquad\qquad , t_{\{i_1-\gamma,i_3-\gamma\}}, t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma\}}\big),
\end{aligned} \tag{37}$$

with $\overline{i_1} \neq \overline{i_2}$, $\overline{i_1} \neq \overline{i_3}$, and $i_2 \neq i_3$.

Proof. To compute F_{P_7}, we use the formula in [14, Theorem 1], with $\chi=2$. Then, $F_{P_7,1}^1$ is the sum of function $\mathcal{A}_{P_7}$ over all possible values of i_1 and $\{i_1,i_2\}$. In Pattern P_7, CN c_1, which connects all four VNs, is different from the other two CNs. Moreover, in a group of three CNs that can form P_7, c_1 can be any one of these three CNs. These facts are again the reason why i_1 of c_1 has to be separated from $\{i_2,i_3\}$ in the expression of $F_{P_7,1}^1$. Regarding $F_{P_7,1}^2$, we need to account for Case 7.2, Case 7.3, and Case 7.4. For Case 7.2, we need to distinguish between two situations; when $r<e$ and when $r>e$, which gives the two summations of $\mathcal{B}_{P_7}$ in (37). This distinction is not needed for neither Case 7.3 nor Case 7.4 since the two replicas have the same number and connectivity of degree-2 overlaps. Note that c_2 and c_3 are not adjacent (no path of only one VN connects them) in P_7, which means it is possible to have $\overline{i_2}=\overline{i_3}$, but not $i_2=i_3$, for that pattern. ∎

H. Analysis of Pattern P_8 (size 4×3)

This pattern has three VNs, and the adjacent pairs are v_1-v_2 and v_1-v_3 (not all pairs) according to P_8 in Fig. 2. Thus, P_8 spans at most $2m+1=3$ consecutive replicas (see [14, Lemma 1]). Pattern P_8 does not exist in the case of $\gamma=3$. Suppose P_8 has the CNs c_1, c_2, c_3, and c_4. The pattern is formed of three overlaps, two of degree-2 and one of degree-4. The degree-2 overlaps are not only distinct, but also mutually exclusive (i.e., they do not share any CNs). Define the CNs such that c_1 and c_2 are directly connected twice, which is the same for c_3 and c_4. Thus, the overlaps are c_1-c_2, c_3-c_4, and $c_1-c_2-c_3-c_4$ (see also P_8 in Fig. 2).

Lemma 10. *Case 8.1: The number of instances of P_8 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\begin{aligned}
&\mathcal{A}_{P_8}\left(t_{\{i_1,i_2\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}}\right) \\
&= t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_1,i_2,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_2,i_3,i_4\}}-2\right)^+ \\
&+ t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_1,i_2,i_3,i_4\}}-1\right)^+ \\
&+ \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_1,i_2,i_3,i_4\}}-1\right)^+ \\
&+ \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_2,i_3,i_4\}}.
\end{aligned} \tag{38}$$

Case 8.2: The number of instances of P_8 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in two replicas, s.t. the two degree-2 overlaps are in $\mathbf{R}_r$, and the degree-4 overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}&\mathcal{B}_{P_8}\big(t_{\{i_1,i_2\}},t_{\{i_3,i_4\}},t_{\{i_1,i_2,i_3,i_4\}}\\&\quad,t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\big)\\&\quad=\Big[t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-1\right)^+\\&\qquad+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_3,i_4\}}\Big]\\&\qquad\cdot t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}.\end{aligned}\tag{39}$$

Case 8.3: The number of instances of P_8 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in two replicas, s.t. the degree-4 overlap and the c_1-c_2 overlap are in $\mathbf{R}_r$, and the c_3-c_4 overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}&\mathcal{C}_{P_8}\left(t_{\{i_1,i_2\}},t_{\{i_1,i_2,i_3,i_4\}},t_{\{i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\right)\\&\quad=t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_1,i_2\}}-1\right)^+t_{\{i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}.\end{aligned}\tag{40}$$

Case 8.4: The number of instances of P_8 with c_1, c_2, c_3, and c_4 as defined previously, and all overlaps in three replicas, s.t. the c_1-c_2 overlap is in $\mathbf{R}_r$, the c_3-c_4 overlap is in $\mathbf{R}_e$, and the degree-4 overlap is in $\mathbf{R}_s$, $r<e$, is:

$$\begin{aligned}&\mathcal{D}_{P_8}\big(t_{\{i_1,i_2\}},t_{\{i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\\&\quad,t_{\{i_1+\theta_{r,s}\gamma,i_2+\theta_{r,s}\gamma,i_3+\theta_{r,s}\gamma,i_4+\theta_{r,s}\gamma\}}\big)\\&\quad=t_{\{i_1,i_2\}}t_{\{i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\\&\qquad\cdot t_{\{i_1+\theta_{r,s}\gamma,i_2+\theta_{r,s}\gamma,i_3+\theta_{r,s}\gamma,i_4+\theta_{r,s}\gamma\}}.\end{aligned}\tag{41}$$

Proof. In Case 8.1, the number we are after is the number of ways to choose 1 overlap from each family (there exist three different families for P_8). In order to avoid over-counting, it is required to distinguish between the two situations when a degree-2 overlap (c_1-c_2 or c_3-c_4) is part of a $c_1-c_2-c_3-c_4$ degree-4 overlap, and when this is not the case. Taking this requirement into account yields the four added terms in (38). The same applies for Case 8.2, with the exception that here the degree-4 overlap is chosen from $t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}$ overlaps, which reduces the number of added terms to the two in (39). Following the same logic gives (40). This distinction is not needed for Case 8.4. ∎

Theorem 8. *The total number of instances of Pattern P_8 in the binary protograph of an SC code that has parameters $\gamma\geq 4$, κ, $m=1$, $L\geq 3$, and $\mathcal{O}$, is:*

$$F_{P_8}=\sum_{k=1}^{3}(L-k+1)F_{P_8,1}^k,\tag{42}$$

where $F_{P_8,1}^k$, $k\in\{1,2,3\}$, are given by:

$$\begin{aligned}F_{P_8,1}^1&=\frac{1}{2}\sum_{\{i_1,i_2\}\subset\{0,\ldots,2\gamma-1\},\{i_3,i_4\}\subset\{0,\ldots,2\gamma-1\}}\mathcal{A}_{P_8}\left(t_{\{i_1,i_2\}},t_{\{i_3,i_4\}},t_{\{i_1,i_2,i_3,i_4\}}\right),\\F_{P_8,1}^2&=\frac{1}{2}\sum_{\{i_1,i_2\}\subset\{\gamma,\ldots,2\gamma-1\},\{i_3,i_4\}\subset\{\gamma,\ldots,2\gamma-1\}}\mathcal{B}_{P_8}\big(t_{\{i_1,i_2\}},t_{\{i_3,i_4\}},t_{\{i_1,i_2,i_3,i_4\}}\\&\qquad,t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma,i_4-\gamma\}}\big)\\&\quad+\frac{1}{2}\sum_{\{i_1,i_2\}\subset\{0,\ldots,\gamma-1\},\{i_3,i_4\}\subset\{0,\ldots,\gamma-1\}}\mathcal{B}_{P_8}\big(t_{\{i_1,i_2\}},t_{\{i_3,i_4\}},t_{\{i_1,i_2,i_3,i_4\}}\\&\qquad,t_{\{i_1+\gamma,i_2+\gamma,i_3+\gamma,i_4+\gamma\}}\big)\\&\quad+\sum_{\{i_1,i_2\}\subset\{0,\ldots,2\gamma-1\},\{i_3,i_4\}\subset\{\gamma,\ldots,2\gamma-1\}}\mathcal{C}_{P_8}\left(t_{\{i_1,i_2\}},t_{\{i_1,i_2,i_3,i_4\}},t_{\{i_3-\gamma,i_4-\gamma\}}\right)\\&\quad+\sum_{\{i_1,i_2\}\subset\{0,\ldots,2\gamma-1\},\{i_3,i_4\}\subset\{0,\ldots,\gamma-1\}}\mathcal{C}_{P_8}\left(t_{\{i_1,i_2\}},t_{\{i_1,i_2,i_3,i_4\}},t_{\{i_3+\gamma,i_4+\gamma\}}\right)\\F_{P_8,1}^3&=\sum_{\{i_1,i_2\}\subset\{\gamma,\ldots,2\gamma-1\},\{i_3,i_4\}\subset\{2\gamma,\ldots,3\gamma-1\}}\mathcal{D}_{P_8}\big(t_{\{i_1,i_2\}},t_{\{i_3-2\gamma,i_4-2\gamma\}}\\&\qquad,t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma,i_4-\gamma\}}\big),\end{aligned}\tag{43}$$

with $\overline{i_1}\neq\overline{i_2}$, $\overline{i_1}\neq\overline{i_3}$, $\overline{i_1}\neq\overline{i_4}$, $\overline{i_2}\neq\overline{i_3}$, $\overline{i_2}\neq\overline{i_4}$, and $\overline{i_3}\neq\overline{i_4}$.

Proof. To compute F_{P_8}, we use the formula in [14, Theorem 1], with $\chi=3$. Then, $F_{P_8,1}^1$ is the sum of function $\mathcal{A}_{P_8}$ over all possible values of $\{i_1,i_2\}$ and $\{i_3,i_4\}$. In Pattern P_8, CNs c_1 and c_2 are directly connected twice, and CNs c_3 and c_4 are directly connected twice, which creates two separate groups of CNs. Consequently, the set $\{i_1,i_2\}$ has to be separated from the set $\{i_3,i_4\}$, despite having the same range, in the expression of $F_{P_8,1}^1$. Regarding $F_{P_8,1}^2$, we need to account for Case 8.2 and Case 8.3. For both cases, we need to distinguish between two situations; when $r<e$ and when $r>e$, which results in two summations for each case. Since it does not matter for the counts of $\mathcal{A}_{P_8}$ and $\mathcal{B}_{P_8}$ whether the set $\{i_1,i_2\}$ or the set $\{i_3,i_4\}$ is chosen first, we multiply by $\frac{1}{2}$ in (43) to account for repetitions (it does matter for the count of $\mathcal{C}_{P_8}$ because the degree-2 overlaps, c_1-c_2 and c_3-c_4, are in two different replicas). Regarding $F_{P_8,1}^3$, the only situation under which P_8 spans 3 consecutive replicas in the case of $m=1$ is what is described in Case 8.4, with the addition that the degree-4 overlap has to be in the middle replica (i.e., $r<s<e$). This situation is accounted for in the last line of (43). ∎

I. Analysis of Pattern P_9 (size 4×4)

This pattern has four VNs, and the adjacent pairs are v_1-v_2, v_2-v_3, v_3-v_4, and v_1-v_4 (not all pairs) according to P_9 in Fig. 2. Thus, P_9 also spans at most $2m+1=3$ consecutive replicas. Suppose P_9 has the CNs c_1, c_2, c_3, and c_4. The pattern is formed of four distinct degree-2 overlaps. Define the CNs such that the adjacent pairs (connected via at least one path of only one VN) are c_1-c_2, c_2-c_3, c_3-c_4, and c_1-c_4. This definition already implies what the overlaps are.

Lemma 11. *Case 9.1: The number of instances of P_9 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in one replica, $\mathbf{R}_r$, is:*

$$\mathcal{A}_{P_9}\big(t_{\{i_1,i_2\}},t_{\{i_2,i_3\}},t_{\{i_3,i_4\}},t_{\{i_1,i_4\}},t_{\{i_1,i_2,i_3\}},t_{\{i_1,i_2,i_4\}},t_{\{i_1,i_3,i_4\}},t_{\{i_2,i_3,i_4\}},t_{\{i_1,i_2,i_3,i_4\}}\big)$$
$$= \mathcal{A}_{P_9,1}+\mathcal{A}_{P_9,2}+\mathcal{A}_{P_9,3}+\mathcal{A}_{P_9,4}, \tag{44}$$

$$\begin{aligned}
\mathcal{A}_{P_9,1} &= t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_1,i_2,i_3,i_4\}}-1\right)^+\\
&\cdot\left(t_{\{i_1,i_3,i_4\}}-2\right)^+\left(t_{\{i_1,i_4\}}-3\right)^+\\
&+t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_1,i_2,i_3,i_4\}}-1\right)^+\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)\left(t_{\{i_1,i_4\}}-1\right)^+,
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{P_9,2} &= \left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_2,i_3,i_4\}}\\
&\cdot\left(t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_2,i_3,i_4\}}\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot t_{\{i_1,i_3,i_4\}}\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}-1\right)^+t_{\{i_1,i_4\}}\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}-1\right)^+\\
&\cdot t_{\{i_1,i_3,i_4\}}\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}-1\right)^+\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)t_{\{i_1,i_4\}},
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{P_9,3} &= \left(t_{\{i_1,i_2,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_2,i_3,i_4\}}\\
&\cdot\left(t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-3\right)^+\\
&+\left(t_{\{i_1,i_2,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_2,i_3,i_4\}}\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+\left(t_{\{i_1,i_2,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot t_{\{i_1,i_3,i_4\}}\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+\left(t_{\{i_1,i_2,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)\\
&\cdot t_{\{i_1,i_3,i_4\}}\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+\left(t_{\{i_1,i_2,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)\left(t_{\{i_1,i_4\}}-1\right)^+,
\end{aligned}$$

$$\begin{aligned}
\mathcal{A}_{P_9,4} &= \left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_4\}}+t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_1,i_3,i_4\}}-1\right)^+\left(t_{\{i_1,i_4\}}-2\right)^+\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_4\}}+t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_4\}}+t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_1,i_3,i_4\}}\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_4\}}+t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_2,i_3,i_4\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}-1\right)^+t_{\{i_1,i_4\}}\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_4\}}+t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)t_{\{i_1,i_3,i_4\}}\left(t_{\{i_1,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_4\}}+t_{\{i_1,i_2,i_3,i_4\}}\right)\\
&\cdot\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)\left(t_{\{i_3,i_4\}}-t_{\{i_1,i_3,i_4\}}\right)t_{\{i_1,i_4\}}.
\end{aligned} \tag{45}$$

Case 9.2: The number of instances of P_9 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in two replicas, s.t. three overlaps are in $\mathbf{R}_r$, and the c_1-c_4 overlap is in $\mathbf{R}_e$, is:

$$\begin{aligned}
&\mathcal{B}_{P_9}\big(t_{\{i_1,i_2\}},t_{\{i_2,i_3\}},t_{\{i_3,i_4\}},t_{\{i_1,i_2,i_3\}},t_{\{i_2,i_3,i_4\}},t_{\{i_1,i_2,i_3,i_4\}},t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\big)\\
&=\Big[t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_2,i_3,i_4\}}-1\right)^+\left(t_{\{i_3,i_4\}}-2\right)^+\\
&+t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)\left(t_{\{i_3,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)t_{\{i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2,i_3\}}-t_{\{i_1,i_2,i_3,i_4\}}\right)\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}-1\right)^+t_{\{i_3,i_4\}}\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)t_{\{i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-1\right)^+\\
&+\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)\left(t_{\{i_2,i_3\}}-t_{\{i_2,i_3,i_4\}}\right)t_{\{i_3,i_4\}}\Big]\\
&\cdot t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}.
\end{aligned} \tag{46}$$

Case 9.3: The number of instances of P_9 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in two replicas, s.t. c_1-c_2 and c_2-c_3 overlaps are in $\mathbf{R}_r$, and c_3-c_4 and c_1-c_4 overlaps are in $\mathbf{R}_e$, $r<e$, is:

$$\begin{aligned}
&\mathcal{C}_{P_9}\big(t_{\{i_1,i_2\}},t_{\{i_2,i_3\}},t_{\{i_1,i_2,i_3\}},t_{\{i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}},t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}},t_{\{i_1+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\big)\\
&=\Big[t_{\{i_1,i_2,i_3\}}\left(t_{\{i_2,i_3\}}-1\right)^++\left(t_{\{i_1,i_2\}}-t_{\{i_1,i_2,i_3\}}\right)t_{\{i_2,i_3\}}\Big]\\
&\cdot\Big[t_{\{i_1+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\left(t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}-1\right)^+\\
&+\left(t_{\{i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}-t_{\{i_1+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\right)\\
&\cdot t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\Big].
\end{aligned} \tag{47}$$

Case 9.4: The number of instances of P_9 with CNs c_1, c_2, c_3, and c_4 as defined in the previous paragraph, and all overlaps in two replicas, s.t. c_1-c_2 and c_3-c_4 overlaps are in $\mathbf{R}_r$, and c_2-c_3 and c_1-c_4 overlaps are in $\mathbf{R}_e$, $r<e$, is:

$$
\begin{aligned}
&\mathcal{D}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}}, t_{\{i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}} \\
&\quad , t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}, t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\big) \\
&= \Big[t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-1\right)^+ \\
&\quad + \left(t_{\{i_1,i_2\}} - t_{\{i_1,i_2,i_3,i_4\}}\right) t_{\{i_3,i_4\}}\Big] \\
&\cdot \Big[t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\left(t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}-1\right)^+ \\
&\quad + \left(t_{\{i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}} - t_{\{i_1+\theta_{r,e}\gamma,i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\right) \\
&\qquad \cdot t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}\Big]. \qquad (48)
\end{aligned}
$$

Case 9.5: The number of instances of P_9 with CNs c_1, c_2, c_3, and c_4 as defined previously, and all overlaps in three replicas, s.t. c_1-c_2 and c_3-c_4 overlaps are in $\mathbf{R}_r$, the c_2-c_3 overlap is in $\mathbf{R}_e$, and the c_1-c_4 overlap is in $\mathbf{R}_s$, $e<s$, is:

$$
\begin{aligned}
&\mathcal{G}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}} \\
&\quad , t_{\{i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}}, t_{\{i_1+\theta_{r,s}\gamma,i_4+\theta_{r,s}\gamma\}}\big) \\
&\quad = \Big[t_{\{i_1,i_2,i_3,i_4\}}\left(t_{\{i_3,i_4\}}-1\right)^+ \\
&\qquad + \left(t_{\{i_1,i_2\}} - t_{\{i_1,i_2,i_3,i_4\}}\right) t_{\{i_3,i_4\}}\Big] \\
&\qquad \cdot t_{\{i_2+\theta_{r,e}\gamma,i_3+\theta_{r,e}\gamma\}} t_{\{i_1+\theta_{r,s}\gamma,i_4+\theta_{r,s}\gamma\}}. \qquad (49)
\end{aligned}
$$

Proof. In Case 9.1, the number we are after is the number of ways to choose 1 overlap from each family (there exist four different families for P_9). In order to avoid over-counting, multiple distinctions need to be performed. For the degree-2 overlap c_1-c_2, it is required to distinguish between the four situations when that overlap is part of a $c_1-c_2-c_3-c_4$ degree-4 overlap, when that overlap is part of a $c_1-c_2-c_3$ degree-3 overlap that is not itself part of a $c_1-c_2-c_3-c_4$ degree-4 overlap, when that overlap is part of a $c_1-c_2-c_4$ degree-3 overlap that is not itself part of a $c_1-c_2-c_3-c_4$ degree-4 overlap, and when neither of these previous three situations holds. This particular distinction results in having four functions, $\mathcal{A}_{P_9,1}$, $\mathcal{A}_{P_9,2}$, $\mathcal{A}_{P_9,3}$, and $\mathcal{A}_{P_9,4}$. Next, only the degree-2 overlaps c_2-c_3, c_3-c_4, and c_1-c_4 need to be chosen. Consequently, for for the degree-2 overlap c_2-c_3, it is required to distinguish between only three situations; when that overlap is part of a $c_1-c_2-c_3-c_4$ degree-4 overlap, when that overlap is part of a $c_2-c_3-c_4$ degree-3 overlap that is not itself part of a $c_1-c_2-c_3-c_4$ degree-4 overlap, and when neither of these previous two situations holds. A logically similar distinction needs to be performed for the next degree-2 overlap, c_3-c_4 (only two situations to distinguish between). Addressing all these distinctions results in (44). The same applies for Case 9.2, with the exception that here the degree-2 overlap c_1-c_4 is chosen from $t_{\{i_1+\theta_{r,e}\gamma,i_4+\theta_{r,e}\gamma\}}$ overlaps, which divides the number of added terms in (45) by four to reach (46). Following the same logic used in Case 7.4 gives (47) and (48). Case 9.5 is similar to Case 8.2, with the exception that here there are two degree-2 overlaps outside $\mathbf{R}_r$, and they are distributed over $\mathbf{R}_e$ (for the c_2-c_3 overlap) and $\mathbf{R}_s$ (for the c_1-c_4 overlap). ■

Theorem 9. *The total number of instances of Pattern P_9 in the binary protograph of an SC code that has parameters $\gamma\geq 3$, κ, $m=1$, $L\geq 3$, and $\mathcal{O}$, is:*

$$
F_{P_9} = \sum_{k=1}^{3}(L-k+1)F^k_{P_9,1}, \qquad (50)
$$

where $F^k_{P_9,1}$, $k\in\{1,2,3\}$, are given by:

$$
\begin{aligned}
F^1_{P_9,1} &= \frac{1}{4}\sum_{i_1\in\{0,\ldots,2\gamma-1\},\{i_2,i_4\}\subset\{0,\ldots,2\gamma-1\},i_3\in\{0,\ldots,2\gamma-1\}} \mathcal{A}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_2,i_3\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_4\}}, t_{\{i_1,i_2,i_3\}} \\
&\qquad , t_{\{i_1,i_2,i_4\}}, t_{\{i_1,i_3,i_4\}}, t_{\{i_2,i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}}\big), \\
F^2_{P_9,1} &= \sum_{\{i_1,i_4\}\subset\{\gamma,\ldots,2\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{0,\ldots,2\gamma-1\}} \mathcal{B}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_2,i_3\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3\}} \\
&\qquad , t_{\{i_2,i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}}, t_{\{i_1-\gamma,i_4-\gamma\}}\big) \\
&\quad + \sum_{\{i_1,i_4\}\subset\{0,\ldots,\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_3\in\{0,\ldots,2\gamma-1\}} \mathcal{B}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_2,i_3\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3\}} \\
&\qquad , t_{\{i_2,i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}}, t_{\{i_1+\gamma,i_4+\gamma\}}\big) \\
&\quad + \sum_{\{i_1,i_3\}\subset\{\gamma,\ldots,2\gamma-1\},i_2\in\{0,\ldots,2\gamma-1\},i_4\in\{\gamma,\ldots,3\gamma-1\}} \mathcal{C}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_2,i_3\}}, t_{\{i_1,i_2,i_3\}}, t_{\{i_3-\gamma,i_4-\gamma\}} \\
&\qquad , t_{\{i_1-\gamma,i_4-\gamma\}}, t_{\{i_1-\gamma,i_3-\gamma,i_4-\gamma\}}\big) \\
&\quad + \frac{1}{2}\sum_{\{i_1,i_4\}\subset\{\gamma,\ldots,2\gamma-1\},i_2\in\{\gamma,\ldots,2\gamma-1\},i_3\in\{\gamma,\ldots,2\gamma-1\}} \mathcal{D}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}}, t_{\{i_2-\gamma,i_3-\gamma\}} \\
&\qquad , t_{\{i_1-\gamma,i_4-\gamma\}}, t_{\{i_1-\gamma,i_2-\gamma,i_3-\gamma,i_4-\gamma\}}\big), \\
F^3_{P_9,1} &= \sum_{\{i_1,i_4\}\subset\{\gamma,\ldots,2\gamma-1\},i_2\in\{0,\ldots,\gamma-1\},i_3\in\{0,\ldots,\gamma-1\}} \mathcal{G}_{P_9}\big(t_{\{i_1,i_2\}}, t_{\{i_3,i_4\}}, t_{\{i_1,i_2,i_3,i_4\}} \\
&\qquad , t_{\{i_2+\gamma,i_3+\gamma\}}, t_{\{i_1-\gamma,i_4-\gamma\}}\big), \qquad (51)
\end{aligned}
$$

with $\overline{i_1}\neq\overline{i_2}$, $i_1\neq i_3$, $\overline{i_1}\neq\overline{i_4}$, $\overline{i_2}\neq\overline{i_3}$, $i_2\neq i_4$, and $\overline{i_3}\neq\overline{i_4}$.

Proof. To compute F_{P_9}, we use the formula in [14, Theorem 1], with $\chi=3$. Then, $F^1_{P_9,1}$ is the sum of function $\mathcal{A}_{P_9}$ over all possible values of i_1, $\{i_2,i_4\}$, and i_3,. In a group of four CNs, there exists 3 unique ways to construct P_9 (which is a cycle of length 8) among them. Since the above separation gives 12 options, we multiply by $\frac{1}{4}$ in the expression of $F^1_{P_9,1}$ to account for repetitions (similar logic applies for $\mathcal{D}_{P_9}$ in $F^2_{P_9,1}$). Regarding $F^2_{P_9,1}$, we need to account for Case 9.2, Case 9.3, and Case 9.4. For Case 9.2, we need to distinguish between two situations; when $r<e$ and when $r>e$, which is not needed for Cases 9.3 and 9.4 since the two replicas in them have the same number and connectivity of degree-2 overlaps. Regarding $F^3_{P_9,1}$, the only situation under which P_9 spans 3 consecutive replicas in the case of $m=1$ is what is described in Case 9.5, with the addition that the the replica with two degree-2 overlaps has to be the middle replica (i.e., $s<r<e$). This situation is accounted for in the last line of (51). Note that c_1 and c_3 are not adjacent in P_9, and the same applies for c_2 and c_4. Thus, it is possible to have $\overline{i_1}=\overline{i_3}$ and $\overline{i_2}=\overline{i_4}$, but not $i_1=i_3$ nor $i_2=i_4$, for that pattern. ■

After deriving the expressions of F_{P_ℓ}, $\forall\ell$, as functions of the overlap parameters in $\mathcal{O}$, we use (3), (4), and [14, Lemma 3] to express F_{tot} as a function of the parameters in $\mathcal{O}_{\text{ind}}$ (which

is the set of independent, non-zero overlap parameters). Thus, our ***discrete optimization problem*** is:

$$F_{\text{tot}}^{*} = \min_{\mathcal{O}_{\text{ind}}} F_{\text{tot}}. \tag{52}$$

The constraints of the optimization problem in (52) are linear constraints capturing interval constraints in addition to the balanced partitioning constraint [15]. These are the constraints under which the partitioning becomes valid. Similar to the set $\mathcal{O}_{\text{ind}}$, the optimization constraints depend only on code parameters, and not on the common substructure of interest (which depends on the channel). For the case of $\gamma = 3$, $m = 1$, and any κ, $\mathcal{O}_{\text{ind}} = \{t_0, t_1, t_2, t_{\{0,1\}}, t_{\{0,2\}}, t_{\{1,2\}}, t_{\{0,1,2\}}\}$, and the optimization constraints are (see also [14] and [15]):

$$\begin{aligned}
&0 \le t_0 \le \kappa,\ 0 \le t_{\{0,1\}} \le t_0,\ t_{\{0,1\}} \le t_1 \le \kappa - t_0 + t_{\{0,1\}},\\
&0 \le t_{\{0,1,2\}} \le t_{\{0,1\}},\ t_{\{0,1,2\}} \le t_{\{0,2\}} \le t_0 - t_{\{0,1\}} + t_{\{0,1,2\}},\\
&t_{\{0,1,2\}} \le t_{\{1,2\}} \le t_1 - t_{\{0,1\}} + t_{\{0,1,2\}},\\
&t_{\{0,2\}} + t_{\{1,2\}} - t_{\{0,1,2\}} \le t_2\\
&\qquad \le \kappa - t_0 - t_1 + t_{\{0,1\}} + t_{\{0,2\}} + t_{\{1,2\}} - t_{\{0,1,2\}},\\
&\textit{and}\ \ \lfloor 3\kappa/2 \rfloor \le t_0 + t_1 + t_2 \le \lceil 3\kappa/2 \rceil .
\end{aligned} \tag{53}$$

The solution of this optimization problem is not unique. However, since all the solutions have the same performance (e.g., they all achieve F_{tot}^{*}, see also [15]), we work with one of these solutions, and call it an optimal vector, $\mathbf{t}^{*}$.

V. CPO: Customization for PR Systems

Using the optimal vector $\mathbf{t}^{*}$, computed as described in the previous section, $\mathbf{H}^{\text{p}}$ is partitioned and the protograph matrix of the SC code, $\mathbf{H}_{\text{SC}}^{\text{p}}$, is constructed. The next step is preventing as many objects in the protograph as possible from being reflected in the unlabeled graph of the SC code, via optimizing the circulant powers using the CPO. Here, the CPO is customized for the $(4, 4(\gamma-2))$ object, which is the common substructure for detrimental configurations in the case of PR systems (see also Fig. 1).

From the previous analysis, a Pattern P_ℓ spans at most either $m+1 = 2$ or $2m+1 = 3$ consecutive replicas, depending on the value of ℓ. Thus, in the CPO, it suffices to operate on the PM $\mathbf{\Pi}_1^{3,\text{p}}$, which is the non-zero part of the first 3 replicas in $\mathbf{H}_{\text{SC}}^{\text{p}}$, and has the size $4\gamma \times 3\kappa$. Circulant powers associated with the 1's in $\mathbf{H}^{\text{p}}$ are defined as $f_{i,j}$, where $0 \le i \le \gamma-1$ and $0 \le j \le \kappa-1$. Let the circulant powers associated with the 1's in $\mathbf{\Pi}_1^{3,\text{p}}$ be $f'_{i',j'}$, where $0 \le i' \le 4\gamma-1$ and $0 \le j' \le 3\kappa-1$. From the repetitive nature of the PM $\mathbf{\Pi}_1^{3,\text{p}}$, $f'_{i',j'} = f_{\overline{i'},\widetilde{j'}}$, where $\overline{i'} = (i' \bmod \gamma)$ and $\widetilde{j'} = (j' \bmod \kappa)$. Define our cycle-8 candidate in the graph of $\mathbf{\Pi}_1^{3,\text{p}}$ as $c_1 - v_1 - c_2 - v_2 - c_3 - v_3 - c_4 - v_4$, which is a particular way of traversing a pattern and not necessarily a protograph cycle (see also Figures 1 and 2). This candidate results in z (or $z/2$ in the case of P_1 only) cycles of length 8 after lifting if and only if [16]:

$$\begin{aligned}
f'_{c_1,v_1} + f'_{c_2,v_2} + f'_{c_3,v_3} + f'_{c_4,v_4} \equiv\\
f'_{c_1,v_2} + f'_{c_2,v_3} + f'_{c_3,v_4} + f'_{c_4,v_1} \pmod{z}.
\end{aligned} \tag{54}$$

The goal is to prevent as many cycle-8 candidates in the graph of $\mathbf{H}_{\text{SC}}^{\text{p}}$ as possible from being converted into z (or $z/2$ in the case of P_1) $(4, 4(\gamma-2))$ UASs/UTSs in the graph of $\mathbf{H}_{\text{SC}}$, which is the unlabeled graph of the SC code. In other words, a cycle-8 candidate in the graph of $\mathbf{H}_{\text{SC}}^{\text{p}}$ is allowed to be converted into multiple $(4, 4(\gamma-2)-2\delta)$ UASs/UTSs with $\delta \in \{1, 2\}$, as long as they are not $(4, 0)$ UASs, in the unlabeled graph since these are not instances of the common substructure of interest. These $(4, 4(\gamma-2)-2\delta)$ UASs/UTSs, $\delta \in \{1, 2\}$, are cycles of length 8 with ***internal connections***, which means v_1 and v_3 are adjacent or/and v_2 and v_4 are adjacent (see Fig. 1). For the cycle-8 candidate in the graph of $\mathbf{\Pi}_1^{3,\text{p}}$ that is described in the previous paragraph and has a CN, say c_5, connecting v_1 and v_3, in order to have this internal connection in the lifted cycles, the following condition for a cycle of length 6 must be satisfied in addition to (54):

$$f'_{c_1,v_1} + f'_{c_2,v_2} + f'_{c_5,v_3} \equiv f'_{c_1,v_2} + f'_{c_2,v_3} + f'_{c_5,v_1} \pmod{z}. \tag{55}$$

Similarly, for that cycle-8 candidate in the graph of $\mathbf{\Pi}_1^{3,\text{p}}$ that has a CN, say c_6, connecting v_2 and v_4, in order to have this internal connection in the lifted cycles, the following condition for a cycle of length 6 must be satisfied in addition to (54):

$$f'_{c_1,v_1} + f'_{c_6,v_2} + f'_{c_4,v_4} \equiv f'_{c_1,v_2} + f'_{c_6,v_4} + f'_{c_4,v_1} \pmod{z}. \tag{56}$$

Note that the two CNs, c_5 and c_6, have to be different from the CNs of the pattern itself in order that we consider them in the CPO algorithm as internal connections. The reason is that the final unlabeled graphs of our codes must have no cycles of length 4 (which is also why (54) is applied for P_1 since $f'_{c_1,v_1} + f'_{c_2,v_2} \equiv f'_{c_1,v_2} + f'_{c_2,v_1} \pmod{z}$ is not allowed for any protograph cycle of length 4, $c_1 - v_1 - c_2 - v_2$).

The following lemma discusses the internal connections for different patterns in the protograph.

Lemma 12. *Let η_{P_ℓ} be the maximum number of internal connections Pattern P_ℓ can have (multiple internal connections between the same two VNs are only counted once). Then,*

$$\eta_{P_\ell} = \begin{cases} 0, & \ell \in \{1, 3, 5\},\\ 1, & \ell \in \{2, 6, 8\},\\ 2, & \ell \in \{4, 7, 9\}. \end{cases} \tag{57}$$

Proof. A protograph pattern, P_ℓ, with only two variable nodes cannot have any internal connections ($\ell \in \{1, 3, 5\}$). A protograph pattern with three variable nodes can have at most one internal connection ($\ell \in \{2, 6, 8\}$). A protograph pattern with four variable nodes can have up to two internal connection ($\ell \in \{4, 7, 9\}$), which completes the proof. ■

The case of multiple internal connections between the same two VNs is addressed in the CPO algorithm.

The steps of the customized CPO algorithm for SC codes that have parameters $\gamma \ge 3$, κ, $m = 1$, and $L \ge 3$, are:

1) Assign initial circulant powers to all the $\gamma\kappa$ 1's in $\mathbf{H}^{\text{p}}$. In this work, our initial powers are as in SCB codes. In particular, $f_{i,j} = (i^2)(2j)$, $0 \le i \le \gamma-1$ and $0 \le j \le \kappa-1$ (initially no cycles of length 4 in $\mathbf{H}_{\text{SC}}$).

2) Construct $\mathbf{\Pi}_1^{3,\mathrm{p}}$, via $\mathbf{H}^{\mathrm{p}}$ and $\mathbf{t}^*$. Circulant powers of the 1's in $\mathbf{\Pi}_1^{3,\mathrm{p}}$, $f'_{i',j'}$, are obtained from the 1's in $\mathbf{H}^{\mathrm{p}}$.
3) Define a counting variable $\psi_{i,j}$, $0 \leq i \leq \gamma - 1$ and $0 \leq j \leq \kappa - 1$, for each of the 1's in $\mathbf{H}^{\mathrm{p}}$. Define another counting variable $\psi'_{i',j'}$, $0 \leq i' \leq 4\gamma - 1$ and $0 \leq j' \leq 3\kappa - 1$, for each of the elements in $\mathbf{\Pi}_1^{3,\mathrm{p}}$. Initialize all the variables in this step with zeros. Only $3\gamma\kappa$ counting variables of the form $\psi'_{i',j'}$ are associated with 1's in $\mathbf{\Pi}_1^{3,\mathrm{p}}$. The other variables remain zeros.
4) Locate all instances of the nine patterns in $\mathbf{\Pi}_1^{3,\mathrm{p}}$. Note that locating P_1 means also locating all cycles of length 4 in $\mathbf{\Pi}_1^{3,\mathrm{p}}$, which is needed.
5) Determine the ζ_{P_ℓ} ways to traverse each instance of P_ℓ, $\forall \ell$, to reach a $(4, 4(\gamma - 2))$ UAS/UTS in the unlabeled graph, which are the cycle-8 candidates.
6) Specify all internal connections (CNs) in each candidate determined in Step 5 if they can exist.
7) For each cycle-8 candidate in $\mathbf{\Pi}_1^{3,\mathrm{p}}$, check whether (54) is satisfied for its circulant powers or not.
8) If (54) is satisfied, and the candidate has no internal connections, or (54) is satisfied and the candidate has internal connection(s) but neither (55) nor (56) is satisfied for any internal connection, mark this cycle-8 candidate as an ***active candidate***.
9) Let $F^{k,\mathrm{a}}_{P_\ell,1}$, where $k \in \{1, 2, 3\}$, be the number of active candidates of P_ℓ starting at the first replica and spanning k consecutive replicas in $\mathbf{\Pi}_1^{3,\mathrm{p}}$. Thus, the number of active candidates of P_ℓ spanning k consecutive replicas in $\mathbf{\Pi}_1^{3,\mathrm{p}}$ is $(4-k)F^{k,\mathrm{a}}_{P_\ell,1}$. *(For example, for $k = 1$, $3F^{1,\mathrm{a}}_{P_\ell,1}$ is the number of active candidates of P_ℓ, for any value of ℓ, spanning one replica in $\mathbf{\Pi}_1^{3,\mathrm{p}}$.)*
10) Compute the number of $(4, 4(\gamma - 2))$ UASs/UTSs in $\mathbf{H}_{\mathrm{SC}}$ using the following formula (see also [14]):

$$F_{\mathrm{SC}} = \sum_{\ell=1}^{9} \sum_{k=1}^{3} \left((L - k + 1) F^{k,\mathrm{a}}_{P_\ell,1} \right) z_{P_\ell}, \tag{58}$$

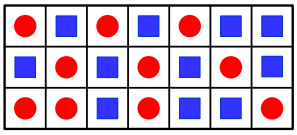
where $z_{P_\ell} = z/2$ if $\ell = 1$, and $z_{P_\ell} = z$ otherwise.
11) Count the number of active candidates each 1 in $\mathbf{\Pi}_1^{3,\mathrm{p}}$ is involved in. Assign weight $w_k = (L-k+1)/(4-k)$ to the number of active candidates spanning k consecutive replicas in $\mathbf{\Pi}_1^{3,\mathrm{p}}$ (see also [14]). Multiply w_k by $1/2$ if the candidate is associated to P_1. *(For example, for $k = 3$, the weight of the number of active candidates spanning 3 consecutive replicas is $(L - 2)$.)*
12) Store the weighted count associated with each 1 in $\mathbf{\Pi}_1^{3,\mathrm{p}}$, which is indexed by (i', j'), in $\psi'_{i',j'}$.
13) Calculate the counting variables $\psi_{i,j}$, $\forall i, j$, associated with the 1's in $\mathbf{H}^{\mathrm{p}}$ from the counting variables $\psi'_{i',j'}$ associated with the 1's in $\mathbf{\Pi}_1^{3,\mathrm{p}}$ (computed in Steps 11 and 12) using the following formula:

$$\psi_{i,j} = \sum_{i':\overline{i'}=i} \sum_{\substack{j':\widetilde{j'}=j \\ \mathbf{\Pi}_1^{3,\mathrm{p}}[i'][j'] \neq 0}} \psi'_{i',j'}, \tag{59}$$

14) Sort these $\gamma\kappa$ 1's of $\mathbf{H}^{\mathrm{p}}$ in a list descendingly according to the counts in $\psi_{i,j}$, $\forall i, j$.
15) Pick a subset of 1's from the top of this list, and change the circulant powers associated with them.
16) Using these interim powers, do Steps 7, 8, 9, and 10.
17) If F_{SC} is reduced while maintaining no cycles of length 4 and no $(4, 0)$ objects (in the case of $\gamma = 3$) in $\mathbf{H}_{\mathrm{SC}}$, update F_{SC} and the circulant powers, then go to Step 11.
18) Otherwise, return to Step 15 to pick a different set of circulant powers or/and a different subset of 1's (from the 1's in $\mathbf{H}^{\mathrm{p}}$).
19) Iterate until the target F_{SC} (set by the code designer) is achieved, or the reduction in F_{SC} approaches zero.

Step 15 in the CPO algorithm is performed heuristically.

Example 1. *Suppose we are designing an SC code with $\gamma = 3$, $\kappa = 7$, $z = 13$, $m = 1$, and $L = 10$ using the OO-CPO approach for PR systems. Solving the optimization problem in (52) gives an optimal vector $\mathbf{t}^* = [t_0\ t_1\ t_2\ t_{\{0,1\}}\ t_{\{0,2\}}\ t_{\{1,2\}}\ t_{\{0,1,2\}}]^{\mathrm{T}} = [3\ 3\ 4\ 0\ 1\ 2\ 0]^{\mathrm{T}}$, with $F^*_{\mathrm{tot}} = 5170$ patterns (rounded weighted sum) in the graph of $\mathbf{H}^{\mathrm{p}}_{\mathrm{SC}}$. Fig. 3(a) shows how the partitioning is applied on $\mathbf{H}^{\mathrm{p}}$ (or $\mathbf{H}$). Next, applying the CPO results in 2613 $(4, 4)$ UASs in the graph of $\mathbf{H}_{\mathrm{SC}}$. Fig. 3(b) shows the final circulant power arrangement for all circulants in $\mathbf{H}$.*



(a)

0	8	12	2	2	12	0
12	6	5	4	7	5	1
5	2	4	4	8	10	0

(b)

Fig. 3. (a) The OO partitioning of $\mathbf{H}^{\mathrm{p}}$ (or $\mathbf{H}$) of the SC code in Example 1. Entries with circles (resp., squares) are assigned to $\mathbf{H}^{\mathrm{p}}_0$ (resp., $\mathbf{H}^{\mathrm{p}}_1$). (b) The circulant power arrangement for the circulants in $\mathbf{H}$.

Remark 3. *After introducing the concept of patterns in this work, the OO-CPO approach can be easily extended to target other common substructures if needed.*

VI. Experimental Results

In this section, we propose experimental results demonstrating the effectiveness of the OO-CPO approach compared to other code design techniques in PR (1-D MR) systems.

Remark 4. *In this section, all the codes used have no cycles of length 4. Moreover, we opted to work with circulant sizes $z > \kappa$ in order to give more freedom to the CPO, which results in less detrimental objects.*

First, we compare the total number of instances of the common substructure of interest in the unlabeled graphs of SC codes designed using various techniques. All the codes in this comparison have $\gamma = 3$ (i.e., the common substructure of interest is the $(4, 4)$ UAS in Fig. 1) and $m = 1$. In addition to the uncoupled setting ($\mathbf{H}_0 = \mathbf{H}$ and $\mathbf{H}_1 = \mathbf{0}$), we show results for the following three SC code design techniques:

1) The CV technique (see [10]).
2) The OO technique with no CPO applied.
3) The OO technique with circulant powers optimized via the CPO (the OO-CPO approach).

In the uncoupled setting and the first two techniques, circulant powers as in SCB codes, $f_{i,j} = f(i)f(j) = (i^2)(2j)$, are used.

TABLE I
NUMBER OF $(4,4)$ UASS IN SC CODES WITH $\gamma = 3$ AND $m = 1$ DESIGNED USING DIFFERENT TECHNIQUES.

Design technique	Number of $(4,4)$ UASs			
	$\kappa = 7$, $z = 13$, $L = 10$	$\kappa = 11$, $z = 23$, $L = 10$	$\kappa = 17$, $z = 37$, $L = 10$	$\kappa = 19$, $z = 46$, $L = 5$
Uncoupled with SCB	32370	254610	1700890	2425120
SC CV with SCB	9464	91333	652347	845434
SC OO with SCB	6500	53130	440818	579968
SC OO-CPO	2613	32361	254005	184667

The results for different choices of κ, z, and L are listed in Table I. For a particular choice of κ, z, and L codes designed using these different techniques all have block length $= \kappa z L$ and rate $\approx \left[1 - \frac{3(L+1)}{\kappa L}\right]$. Table I demonstrates the significant gains achieved by the OO-CPO approach compared to other techniques. In particular, the proposed OO-CPO approach achieves a reduction in the number of $(4,4)$ UASs that ranges between 85% and 92% compared to the uncoupled setting, and between 61% and 78% compared to the CV technique. Moreover, the importance of the two stages (the OO and the CPO) is highlighted by the numbers in the table.

Second, we present simulation results of SC codes designed using various techniques over the PR channel. We use the PR channel described in [5]. This channel incorporates inter-symbol interference (intrinsic memory), jitter, and electronic noise. The normalized channel density [17], [18] we use is 1.4, and the PR equalization target is $[8\ 14\ 2]$. The receiver consists of filtering units followed by a Bahl Cocke Jelinek Raviv (BCJR) detector, which is based on pattern-dependent noise prediction (PDNP), in addition to a fast Fourier transform based q-ary sum-product algorithm (FFT-QSPA) LDPC decoder [19]. The number of global (detector) iterations is 10, and the number of local (decoder) iterations is 20. Unless a codeword is reached, the decoder performs its prescribed number of local iterations for each global iteration. More details about this PR system can be found in [5].

In the simulations, we use five different codes. All the codes are defined over GF(4). Codes 1, 2, 3, and 4 have $\gamma = 3$, $\kappa = 19$, $z = 46$, $m = 1$, and $L = 5$. Thus, these codes have block length $= 8740$ bits and rate ≈ 0.81. Code 1 is uncoupled. Code 2 is an SC code designed using the CV technique for PR channels [10]. Codes 1 and 2 have SCB circulant powers of the form $f_{i,j} = (i^2)(2j)$. Code 3 is an SC code designed using the OO-CPO approach. Codes 1, 2, and 3 do not have optimized edge weights. Code 4 is the result of applying the WCM framework to Code 3 in order to optimize its edge weights. The number of $(4,4)$ UASs in the unlabeled graphs of Codes 1, 2, and 3 are given in the last column of Table I. Code 5 is a block (BL) code, which is also protograph-based (PB), designed as in [6] and [7]. Code 5 has column weight $= 3$, circulant size $= 46$, block length $= 8832$ bits, rate ≈ 0.81 (similar to all the other codes), and unoptimized weights (similar to all codes except Code 4). Note that because our main focus in this work is the performance, a relatively small value of L (which is 5) along with block decoding are used for SC Codes 1, 2, 3, and 4.

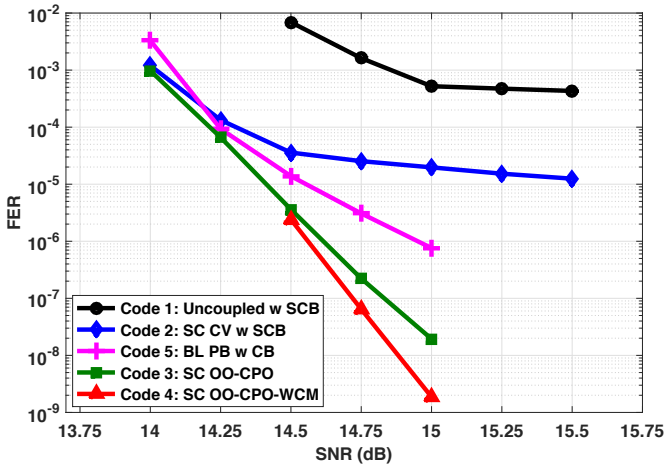


Fig. 4. Simulation results over the PR Flash channel for SC codes designed using different techniques and a BL code.

Fig. 4 demonstrates the effectiveness of the proposed OO-CPO approach in designing high performance SC codes for PR channels. In particular, Code 3 (designed using the OO-CPO approach) outperforms Code 2 (designed using the CV technique) by about 3 orders of magnitude at SNR $= 15$ dB, and by about 1.1 dB at FER $\approx 10^{-5}$. More intriguingly, Code 3 outperforms Code 5 (the block code) by about 1.6 orders of magnitude at SNR $= 15$ dB, and by almost 0.4 dB at FER $\approx 10^{-6}$. The performance of Code 3 is better than the performance Code 5 not only in the error floor region, but also in the waterfall region. An interesting observation is that, in the error profile of Code 3, we found no codewords of weights $\in \{6, 8\}$ (which are $(6,0,0,9,0)$ and $(8,0,0,12,0)$ BASTs) despite the dominant presence of such low weight codewords in the error profiles of Codes 1, 2, and 5 (see also [5] and [7]). From Fig. 4, the WCM framework achieves 1 order of magnitude additional gain.

VII. CONCLUSION

We proposed the OO-CPO approach to optimally design binary and non-binary SC codes for PR channels, via minimizing the number of detrimental objects in the graph of the code. SC codes designed using the OO-CPO approach were shown to significantly outperform SC codes designed using techniques from the literature. More importantly, SC codes designed using our approach were demonstrated to outperform structured block codes with the same parameters.

ACKNOWLEDGEMENT

The work is supported in part by an ASTC-IDEMA grant.

REFERENCES

[1] B. Vasic and E. Kurtas, *Coding and Signal Processing for Magnetic Recording Systems.* CRC Press, 2005.

[2] G. Colavolpe and G. Germi, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 818–825, May 2005.

[3] W. Phakphisut, P. Supnithi, and N. Puttarak, "EXIT chart analysis of nonbinary protograph LDPC codes for partial response channels," *IEEE Trans. Magn.*, vol. 50, no. 11, Nov. 2014, Art. no. 3101904.

[4] P. Chen, C. Kui, L. Kong, Z. Chen, M. Zhang, "Non-binary protograph-based LDPC codes for 2-D-ISI magnetic recording channels," *IEEE Trans. Magn.*, vol. 53, no. 11, Nov. 2017, Art. no. 8108905.

[5] A. Hareedy, B. Amiri, R. Galbraith, and L. Dolecek, "Non-binary LDPC codes for magnetic recording channels: error floor analysis and optimized code design," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3194–3207, Aug. 2016.

[6] A. Hareedy, C. Lanka, and L. Dolecek, "A general non-binary LDPC code optimization framework suitable for dense Flash memory and magnetic storage," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 9, pp. 2402–2415, Sep. 2016.

[7] A. Hareedy, C. Lanka, N. Guo, and L. Dolecek, "A combinatorial methodology for optimizing non-binary graph-based codes: theoretical analysis and applications in data storage," Jun. 2017. [Online]. Available: http://arxiv.org/abs/1706.07529

[8] A. J. Felstrom and K. S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2181–2191, Sep. 1999.

[9] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Dec. 2013.

[10] H. Esfahanizadeh, A. Hareedy, and L. Dolecek, "Spatially-coupled codes optimized for magnetic recording applications," *IEEE Trans. Magn.*, vol. 53, no. 2, pp. 1–11, Feb. 2016.

[11] A. R. Iyengar, M. Papaleo, P. H. Siegel, J. K. Wolf, A. Vanelli-Coralli, and G. E. Corazza, "Windowed decoding of protograph-based LDPC convolutional codes over erasure channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2303–2320, Apr. 2012.

[12] H. Esfahanizadeh, A. Hareedy, and L. Dolecek, "A novel combinatorial framework to construct spatially-coupled codes: minimum overlap partitioning," in *Proc. IEEE ISIT*, Aachen, Germany, Jun. 2017, pp. 1693–1697.

[13] D. G. M. Mitchell and E. Rosnes, "Edge spreading design of high rate array-based SC-LDPC codes," in *Proc. IEEE ISIT*, Aachen, Germany, Jun. 2017, pp. 2940–2944.

[14] H. Esfahanizadeh, A. Hareedy, and L. Dolecek, "Finite-length construction of high performance spatially-coupled codes via optimized partitioning and lifting," Feb. 2018. [Online]. Available: http://arxiv.org/abs/1802.06481

[15] A. Hareedy, H. Esfahanizadeh, and L. Dolecek, "High performance non-binary spatially-coupled codes for flash memories," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 229–233.

[16] M. P. C. Fossorier, "Quasi-cyclic low-density parity-check codes from circulant permutation matrices," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1788–1793, Aug. 2004.

[17] S. Srinivasa, Y. Chen, and S. Dahandeh, "A communication-theoretic framework for 2-DMR channel modeling: performance evaluation of coding and signal processing methods," *IEEE Trans. Magn.*, vol. 50, no. 3, pp. 6–12, Mar. 2014.

[18] T. Souvignier, M. Öberg, P. Siegel, R. Swanson, and J. Wolf, "Turbo decoding for partial response channels," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1297 - 1308, Aug. 2000.

[19] D. Declercq and M. Fossorier, "Decoding algorithms for nonbinary LDPC codes over GF(q)," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 633–643, Apr. 2007.