Quantum Data-Syndrome Codes

Alexei Ashikhmin, *Fellow, IEEE*, Ching-Yi Lai, *Member, IEEE*, and Todd A. Brun, *Senior Member, IEEE*

***Abstract*— Performing active quantum error correction to protect fragile quantum states highly depends on the correctness of measured error syndromes. To obtain reliable error syndromes using imperfect physical circuits, we propose syndrome measurement (SM) and quantum data-syndrome (DS) codes. SM codes protect syndrome with linearly dependent redundant stabilizer measurements. DS codes generalize this idea for simultaneous correction of both data qubits and syndrome bits errors. We study fundamental properties of quantum DS codes, including split weight enumerators, generalized MacWilliams identities, and linear programming bounds. In particular, we derive Singleton and Hamming-type upper bounds on the minimum distance of degenerate quantum DS codes. Then we study random DS codes and show that random DS codes with a relatively small additional syndrome measurements achieve the Gilbert-Varshamov bound of stabilizer codes. Finally, we propose a family of CSS-type quantum DS codes based on classical cyclic codes, which include the Steane code and the quantum Golay code.**

***Index Terms*— Data-syndrome (DS) codes, quantum codes, stabilizer codes, macWilliams identities, linear programming bounds, quantum syndrome errors.**

I. Introduction

QUANTUM error-correcting codes provide a method of actively protecting quantum information [1]. In particular, in a quantum stabilizer code, quantum information is stored in the joint $+1$ eigenspace of a set of Pauli operators, called *stabilizers*. To perform quantum error correction, we have to measure a generating set of the stabilizers and get the *error syndrome* (in classical bits). Realistically, the quantum gates used to perform stabilizer measurements are themselves faulty and thus the measurement outcomes can be wrong which lead to a wrong error syndrome and further to a wrong error-correction result.

In this work, we are interested in eliminating the effect of faulty syndrome measurement. Typically, this can be done with the syndromes measured repeatedly in the case of Shor's syndrome extraction [2]. This, however, is equivalent to protecting each syndrome bit individually with a repetition code, which is very inefficient from the information theoretical point of view.

Manuscript received July 3, 2019; revised December 16, 2019; accepted January 6, 2020. Date of publication January 30, 2020; date of current version April 3, 2020. The work of Ching-Yi Lai was supported by the Young Scholar Fellowship Program by Ministry of Science and Technology (MOST) in Taiwan under Grant MOST108-2636-E-009-004. The work of Todd A. Brun was supported in part by NSF under Grant QIS-1719778. This article was presented in part at ISIT 2014 and in part at ISIT 2016. *(Corresponding author: Alexei Ashikhmin.)*

Alexei Ashikhmin is with the Nokia Bell Labs, Murray Hill, NJ 07974 USA (e-mail: alexei.ashikhmin@nokia-bell-labs.com).

Ching-Yi Lai is with the Institute of Communications, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: cylai@nctu.edu.tw).

Todd A. Brun is with the Electrical Engineering Department, University of Southern California, Los Angeles, CA 90089 USA (e-mail: tbrun@usc.edu).

Digital Object Identifier 10.1109/JSAC.2020.2968997

Another approach it to measure an overcomplete set of code generators, which allows obtaining an extended syndrome whose bits are linearly dependent, and therefore a classical error correction can be used for identifying and removing syndrome errors. Even better results can be obtained if both data (qubits) and syndrome errors are decoded together. This approach, under the names of Syndrome-Measurement (SM) and Data-Syndrome (DS) codes, was proposed and studied in [3]–[5]. Similar ides were proposed in [6]. Independently, in the context of higher-dimensional toric and/or color quantum codes similar approach was proposed and studied in [7]–[9]. Very recently Campbel has proposed a theory for this one-shot error correction [10].

In [3] we showed that in the case of Steane SM codes drastically reduce the syndrome measurement error, and considered construction of SM codes as low density generator matrix codes. In [4] we outlined a general framework for deriving upper and lower bounds on the minimum distance of DS codes, and presented some bounds without proofs. In [5] constructions and decoding algorithms of convolutional DS codes have been studied.

In this paper, we give a comprehensive study of quantum SM and DS codes with complete proofs and details omitted in [3], [4]. We show that SM codes drastically reduce the probability of syndrome errors for ceratin families of high rate quantum codes, present proofs of the Hamming and Singleton bounds for degenerate DS codes, our new results on asymptotic bounds for degenerate DS codes, full derivations of the weight enumerators of random DS codes and comparisons of Gilbert-Varshamov bounds for random DS and stabilizer codes.

The paper is organized as follows. In Section II we briefly introduce the main notions of quantum stabilizer codes and error syndromes. In Section III we define SM and DS quantum codes, show how they make syndrome measurement more error resilient and further define the main parameters of SM and DS codes. Further in Section IV we discuss SM codes and demonstrate that for some families of high rate quantum stabilizer codes, SM codes provide a very large reduction of the wrong syndrome measurement compared with repetitive syndrome measurements followed by the majority voting. In Section V we introduce the notion of split weight enumerators of DS codes and show how these enumerators are connected with the minimum distance of DS codes. We use the split weight enumerators in Section VI for developing a linear programming method for deriving upper bounds on the minimum distance of unrestricted (degenerate or non-degenerate) quantum DS codes. We use this method for obtaining the Hamming and Singleton bounds on degenerate DS codes. We further use this method for deriving asymptotic bounds for

unrestricted DS codes. In Section VII we study the split weight enumerators for random DS codes and use them for obtaining the Gilbert-Varshamov (GV) lower bound on the code minimum distance as a function of the normalized number, ρ, of redundant syndrome measurements. We show that small ρ is enough that the GV bound for DS codes would achieve the GV bound of usual stabilizer codes. In VIII we present constructions of CSS-type ([11], [12]) DS codes. In Conclusion we summarize the presented results and observations.

II. Preliminaries

An n-qubit state space is a 2^n-dimensional complex Hilbert space $\mathbb{C}^{2^n}$ and a pure quantum state is a unit vector in $\mathbb{C}^{2^n}$. A basis of the linear operators on the n-qubit state space is the n-fold Pauli operators $\{M_1 \otimes \cdots \otimes M_n : M_j \in \{I, X, Y, Z\}\}$, where

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$
$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad Y = iXZ,$$

are the Pauli matrices. The n-fold Pauli group $\mathcal{G}_n$ is the set of n-fold Pauli operators with possible phases $\pm 1, \pm i$. Note that Pauli operators either commute or anticommute with each other. We can define an inner product in $\mathcal{G}_n$: for $g, h \in \mathcal{G}_n$,

$$\langle g, h \rangle_{\mathcal{G}_n} = \begin{cases} 0, & \text{if } gh = hg; \\ 1, & \text{otherwise.} \end{cases} \tag{1}$$

It is convenient to connect quantum codes with codes over the Galios filed $\mathbb{F}_4 = \{0, 1, \omega, \omega^2\}$ [13], via a homomorphism τ on $\mathcal{G}_1$ that maps I, X, Z, Y to $0, 1, \omega, \omega^2$, respectively, regardless of a possible phase $\pm 1, \pm i$ in front of a Pauli matrix. This homomorphism extends to an n-fold Pauli operator naturally. For example, $\tau(\pm iX \otimes Y \otimes Z \otimes I \otimes I) = (1\omega^2\omega 00)$. For an n-fold Pauli operator g, we denote by $\mathbf{g} \in \mathbb{F}_4^n$ the corresponding vector and vice versa. We define a trace inner product for $\mathbf{x} = (x_1, x_2, \ldots, x_n), \mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{F}_4^n$ by

$$\mathbf{x} * \mathbf{y} = \mathrm{Tr}_{\mathbb{F}_2}^{\mathbb{F}_4}\left(\sum_{i=0}^{n} x_i \bar{y}_i\right), \tag{2}$$

where $\bar{y}_i$ denotes the conjugation of y_i in $\mathbb{F}_4 = \{0, 1, \omega, \omega^2\}$ with $\bar{0} = 0, \bar{1} = 1, \bar{\omega} = \omega^2$, and $\bar{\omega^2} = \omega$. It can be checked that $\langle g, h \rangle_{\mathcal{G}_n} = \mathbf{g} * \mathbf{h}$ for $g, h \in \mathcal{G}_n$ and $\mathbf{g} = \tau(g), \mathbf{h} = \tau(h)$.

Suppose $\mathcal{S} = \langle g_1, \ldots, g_{n-k} \rangle$ is an Abelian subgroup of $\mathcal{G}_n$, where g_j are independent generators of $\mathcal{S}$, such that the minus identity $-I^{\otimes n} \notin \mathcal{S}$. Then $\mathcal{S}$ defines a quantum *stabilizer* code $Q = \{|\psi\rangle \in \mathbb{C}^{2^n} : g|\psi\rangle = |\psi\rangle, \forall g \in \mathcal{S}\}$ of dimension 2^k. The vectors $|\psi\rangle \in \mathcal{C}(\mathcal{S})$ are called the *codewords* of $\mathcal{C}(\mathcal{S})$ and the operators $g \in \mathcal{S}$ are called the *stabilizers* of $\mathcal{C}(\mathcal{S})$. By the quantum error correction conditions [14], [15], it suffices to consider error correction on a discrete set of error operators. Thus we only treat errors that are Pauli operators in this paper.

Let $\mathbf{g}_1, \ldots, \mathbf{g}_{n-k} \in \mathbb{F}_4^n$ be vectors over $\mathbb{F}_4$ corresponding to $g_1, g_2, \ldots, g_{n-k} \in \mathcal{G}_n$, and let consider check matrix

$$H = \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{n-k} \end{bmatrix}. \tag{3}$$

Let C be the classical $[n, n-k]$ additive code generated by H, so $|C| = 2^{n-k}$, and $C^\perp$ be its dual with respect to the trace inner product (2). We have $C \subseteq C^\perp$ since $\mathbf{g}_i * \mathbf{g}_j = 0$ for all i, j. Suppose that a quantum codeword $|\psi\rangle \in Q$ is corrupted by a Pauli error $e \in \mathcal{G}_n$ and let $\mathbf{e} = \tau(e) \in \mathbb{F}_4^n$ be the corresponding vector. Then the syndrome of e is $\mathbf{s} = (s_1, \ldots, s_{n-k}) \in \mathbb{F}_2^{n-k}$, where $s_j = \mathbf{g}_j * \mathbf{e}$. The syndrome $\mathbf{s}$ shows the commutation relations between the Pauli error e and the stabilizers $g_1, g_2, \ldots, g_{n-k}$ and it can be obtained by measuring the observables g_j's on $e|\psi\rangle$.

There are several possible quantum circuits for measuring g_j's. We assume that Shor's syndrome extraction [2] is used. Let p_m be the error rate of the measurement blocks in the Z basis. Let all the other quantum gates (like $CNOT$ and H) have error rate p_g. For some technologies, e.g., quantum trapped ions, we may have $p_m \gg p_g$. Using this observation, we adopt the simple *error model* in which $p_g = 0$. This means that the quantum state does not change during the syndrome measurement. Error models with $p_g > 0$, and perhaps other sources of errors, would be more realistic. However, developing theories starting with simple models and assumptions proved to be productive. For this reason and for avoiding overwhelming details, we use the above simple error model.

Let $\mathbf{s} = (s_1, \ldots, s_{n-k})$ be the correct syndrome and $\widehat{\mathbf{s}} = (\widehat{s}_1, \ldots, \widehat{s}_{n-k})$ be the corrupted syndrome. According to Shor's syndrome extraction circuit and the above error model we have

$$\begin{aligned} p_{\text{err}}(\mathbf{g}_j) &= \Pr(\widehat{s}_j \neq s_j) \\ &= \sum_{i \text{ is odd}} \binom{\text{wt}(\mathbf{g}_j)}{i} p_m^i (1 - p_m)^{\text{wt}(\mathbf{g}_j) - i}. \end{aligned} \tag{4}$$

III. Quantum Data-Syndrome Codes

Quantum error correction can be done by finding the most likely error operator e with a measured syndrome $\mathbf{s} \in \mathbb{F}_2^{n-k}$. However, $\mathbf{s}$ itself could be measured with an error, and we may get $\widehat{\mathbf{s}} = \mathbf{s} + \mathbf{z}$, where $\mathbf{z} \in \mathbb{F}_2^{n-k}$ is syndrome error. Herein we discuss stabilizer codes capable of correcting both data errors and syndrome errors. To shorten notation, we will denote $m \triangleq n - k$, and for vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}_4^n \times \mathbb{F}_2^l$ define the following inner product:

$$\mathbf{x} \star \mathbf{y} = \mathrm{Tr}_{\mathbb{F}_2}^{\mathbb{F}_4}\left(\sum_{i=1}^{n} x_i \bar{y}_i\right) + \sum_{j=1}^{l} x_{n+j} y_{n+j}. \tag{5}$$

To deal with syndrome errors we define a new parity-check matrix

$$\hat{H} = [H \quad I_m], \tag{6}$$

where I_m is considered as a matrix over $\mathbb{F}_2$. We define codes

$$D = \{\mathbf{w} = \mathbf{u}\hat{H} : \mathbf{u} \in \mathbb{F}_2^m\} \subset \mathbb{F}_4^n \times \mathbb{F}_2^m,$$
$$D^\perp = \{\mathbf{v} : \mathbf{w} \star \mathbf{v} = 0, \forall \mathbf{w} \in D\} \subset \mathbb{F}_4^n \times \mathbb{F}_2^m,$$

where $D^\perp$ is the dual code of D with respect to the inner product (5). By decoding $D^\perp$, one can correct both data and syndrome errors. Fujiwara, [6], noticed that choosing generators properly, we can get a code capable of correcting simultaneously multiple data and syndrome errors. In addition, the error-correcting capabilities of $D^\perp$ can be further enhanced. The standard approach to reduce the probability of syndrome measurement error is to repeat measurements several times and take a majority vote. We propose a generalization of this idea by measuring additional stabilizers according to more powerful linear classical codes.

Let C be an $[m+r, m]$ linear binary code with a generator matrix in the systematic form

$$G_C = [I_m \quad A], \tag{7}$$

where $A = [a_{i,j}]$ is an $m \times r$ binary matrix. We define a new set of r stabilizers $\mathbf{f}_j$ by

$$\mathbf{f}_j = a_{1,j}\mathbf{g}_1 + \cdots + a_{m,j}\mathbf{g}_m, \quad \text{for } j = 1, \ldots, r. \tag{8}$$

These $\mathbf{f}_j$ belong to the stabilizer group $\mathcal{S}$, and can be measured without disturbing the underlying quantum codewords. We call C *the syndrome measurement (SM) code*. Let

$$H'^T = \begin{bmatrix} \mathbf{f}_1^T & \cdots & \mathbf{f}_r^T \end{bmatrix}.$$

Measuring additional r stabilizers is equivalent to considering the code defined by the parity-check matrix

$$\begin{bmatrix} H & I_m & 0 \\ H' & 0 & I_r \end{bmatrix},$$

which can be transformed into the form

$$H_{\mathrm{DS}} = \begin{bmatrix} H & I_m & 0 \\ 0 & A^T & I_r \end{bmatrix}. \tag{9}$$

We will say that (9) defines a quantum *data-syndrome code* Q_{DS}. It is convenient to define codes

$$C_{\mathrm{DS}} = \{\mathbf{w} = \mathbf{u}H_{\mathrm{DS}} : \mathbf{u} \in \mathbb{F}_2^{m+r}\} \subset \mathbb{F}_4^n \times \mathbb{F}_2^{m+r},$$
$$C_{\mathrm{DS}}^\perp = \{\mathbf{v} : \mathbf{w} \star \mathbf{v} = 0, \forall \mathbf{w} \in C_{\mathrm{DS}}\} \subset \mathbb{F}_4^n \times \mathbb{F}_2^{m+r},$$

where $C_{\mathrm{DS}}^\perp$ is the dual code of C_{DS} with respect to (5). Slightly abusing terminology, we will call $C_{\mathrm{DS}}^\perp$ also a *data-syndrome* code. We will say that Q_{DS} (or $C_{\mathrm{DS}}^\perp$) has *length* n, *dimension* k, and *size* 2^k. Such a code encodes k (information) qubits into n (code) qubits.

It is easy to see that

$$|C_{\mathrm{DS}}| = 2^{m+r} \text{ and } |C_{\mathrm{DS}}^\perp| = 2^{2n+m+r}/|C_{\mathrm{DS}}| = 2^{2n},$$

so the size of $C_{\mathrm{DS}}^\perp$ does not depend on k and r. For a given matrix (9), we always can find vectors $\mathbf{g}_{m+1}, \ldots, \mathbf{g}_n$ and $\mathbf{h}_1, \ldots, \mathbf{h}_n$ over $\mathbb{F}_4$ such that

$$\mathbf{g}_i * \mathbf{g}_j = 0, \ \mathbf{g}_i * \mathbf{h}_j = 0 \quad \text{for } i \neq j \text{ and } \mathbf{g}_i * \mathbf{h}_i = 1.$$

These vectors allow us to write down a generator matrix of $C_{\mathrm{DS}}^\perp$ in the following explicit form

$$G_{C_{\mathrm{DS}}^\perp} = \begin{bmatrix} G & \mathbf{0} & \mathbf{0} \\ H_1 & \mathbf{0} & \mathbf{0} \\ H_2 & I_m & A \end{bmatrix}, \tag{10}$$

where

$$G = \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_n \end{bmatrix}, \quad H_1 = \begin{bmatrix} \mathbf{h}_{m+1} \\ \vdots \\ \mathbf{h}_n \end{bmatrix}, \ H_2 = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_m \end{bmatrix},$$

and $\mathbf{0}$s are all zero matrices of appropriate sizes.

We will say that a code defined by (9) is an $[[n, k : r]]$ DS code. If the minimum distance d (defined in Section V) is known, it is called an $[[n, k, d : r]]$ code.

IV. Syndrome Measurement Codes

In this section, we consider some examples of stabilizer codes for which it is easy to find an efficient SM code that beat the repetitive syndrome measurement approach.

Suppose that an $[m+r, m]$ SM code C is used. Denote by s_j and z_j the results of correct measurement of $\mathbf{g}_j$ and $\mathbf{f}_j$, respectively. Then

$$\mathbf{x} = (s_1, \ldots, s_m, z_1, \ldots, z_r)$$

is a codeword of C. Because of possible measurement errors, we obtain a vector

$$\widehat{\mathbf{x}} = (\widehat{s}_1, \ldots, \widehat{s}_m, \widehat{z}_1, \ldots, \widehat{z}_r).$$

The probability $\Pr(s_j \neq \widehat{s}_j)$ (similarly $\Pr(z_i \neq \widehat{z}_i)$) is defined by (4). We can correct quantum and syndrome errors simultaneously by decoding vector

$$(\underbrace{0, \ldots, 0}_{n \text{ times}}, \widehat{\mathbf{x}})$$

using a decoder of $C_{\mathrm{DS}}^\perp$. Alternatively we can first correct syndrome errors by decoding $\widehat{\mathbf{x}}$ using a decoder of the SM code C, and next correct quantum errors. The latter approach is typically simpler, though it is suboptimal. In this section we consider this type of decoding.

Applying a decoding algorithm of C to $\widehat{\mathbf{x}}$, we obtain bits $\tilde{s}_1, \ldots, \tilde{s}_m$. We define the *syndrome decoding error* and *average syndrome decoding error*, respectively, as

$$P_{se} = \Pr((s_1, \ldots, s_m) \neq (\tilde{s}_1, \ldots, \tilde{s}_m)), \tag{11}$$

$$P_{SBER} = \frac{1}{m}\sum_{j=1}^{m} \Pr(\tilde{s}_j \neq s_j). \tag{12}$$

The l-fold repeated syndrome measurement can be considered as the SM code with generator matrix

$$G = [\underbrace{I_m \cdots I_m}_{l \text{ times}}].$$

Note that choosing a good SM code is not equivalent to finding a good $[m+r, m]$ linear code in the usual sense, since such a code will have a large minimum distance. Hence the matrix A in (7) will have "heavy" columns, and this will lead to

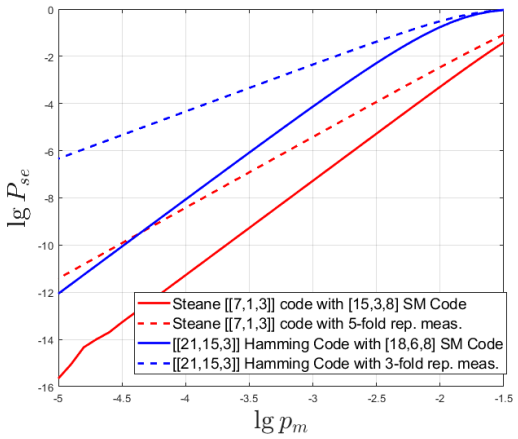

Fig. 1. P_{se} for the $[[7,1,3]]$ Steane and $[[21,15,3]]$ Hamming codes with SM codes and 5-fold and 3-fold repeated measurements.

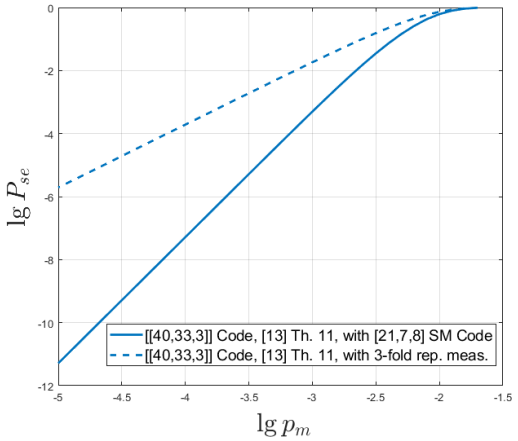

Fig. 2. $[[40,33,3]]$ code from [13], Th.11 with SM code and 3-fold repeated measurements.

$\mathrm{wt}(\mathbf{f}_j) \gg \mathrm{wt}(\mathbf{g}_l)$ and further to $p_{\mathrm{err}}(\mathbf{f}_j) \gg p_{\mathrm{err}}(\mathbf{g}_i)$, which, in turn, will lead to large P_{se} and P_{SBER}.

Below we present several families of high rate quantum codes with the property that all their stabilizers $\mathbf{g} \in \mathcal{S}$ have the same or almost the same weights and therefore any good linear codes can be used for robust syndrome measurement.

Let S_a be a generator matrix of the $[2^a-1, a, 2^{a-1}]$ simplex code. The generators

$$\begin{bmatrix} S_a & 0 \\ 0 & S_a \end{bmatrix}$$

define an $[[2^a-1, 2^a-1-2a, 3]]$ CSS code. Any liner combination of the first (second) a generators is a vector of weight 2^{a-1}.

Another important family is the $[[n, n-2a, 3]]$ quantum Hamming codes $\mathcal{H}_a$ for $n = (4^a-1)/3$ [13, V]. It is not difficult to prove that all generators of $\mathcal{H}_a$ have weight 4^{a-1}.

In [13, Thm 11] a family of $[[n, n-a-2, 3]]$ codes with

$$n = \sum_{i=1}^{(a-1)/2} 2^{2i+1}$$

is defined for odd a. The generators of these codes can have only weights 2^a-2 and 2^a.

For all these families a good $[m+r, m]$ linear code can be used as SM code. On Fig. 1 and Fig. 2 we present P_{se} for quantum codes from above families in combination with SM codes and syndrome repeated measurements. The parameters of the SM codes and repeated measurements are chosen so that the total number of measurements be the same in both cases. We take SM codes from the table of the best linear codes available online http://www.codetables.de/. One can see that SM codes provide drastically smaller P_{se} compared with syndrome repeated measurements.

V. Minimum Distance and Split Weight Enumerators

Let $\mathbf{e}_{\mathrm{DS}} = (\mathbf{g}, \mathbf{0}) \in \mathbb{F}_4^n \times \mathbb{F}_2^{m+r}$, with $\mathbf{g} \in C$. Since $\mathbf{g} \in C$, we have $g \in \mathcal{S}$ and thus $\mathbf{e}_{\mathrm{DS}}$ is harmless. If $\mathbf{e}_{\mathrm{DS}} = (\mathbf{e}, \mathbf{z}) \in C_{\mathrm{DS}}^{\perp} \setminus \{(\mathbf{g}, \mathbf{0}) : \mathbf{g} \in C\}$, then $\mathbf{e} \notin C$. Therefore the operator e does not belong to $\mathcal{S}$ and acts on Q nontrivially. Since $H_{\mathrm{DS}} \star \mathbf{e}_{\mathrm{DS}} = \mathbf{0}^T$ by definition, we conclude that such $\mathbf{e}_{\mathrm{DS}}$ is an undetectable and harmful error. Naturally, the weight $\mathrm{wt}(\mathbf{e}, \mathbf{z})$ is defined as the number of its nonzero entries. We define the *minimum distance* d of Q_{DS} (equivalently $C_{\mathrm{DS}}^{\perp}$) as the minimum weight of any element in $C_{\mathrm{DS}}^{\perp} \setminus \{(\mathbf{g}, \mathbf{0}) : \mathbf{g} \in C\}$.

It is not difficult to see that Q_{DS} (or equivalently $C_{\mathrm{DS}}^{\perp}$) can correct any error $\mathbf{e}_{\mathrm{DS}} = (\mathbf{e}, \mathbf{z})$ (here we do not assume $\mathbf{e}_{\mathrm{DS}} \in C_{\mathrm{DS}}^{\perp}$) with $\mathrm{wt}(\mathbf{e}) = t_{\mathrm{D}}, \mathrm{wt}(\mathbf{z}) = t_{\mathrm{S}}$ if $t_{\mathrm{D}} + t_{\mathrm{S}} < \frac{d}{2}$. Apparently the minimum distance of a DS code cannot be greater than that of the underlying stabilizer code.

Define the split weight enumerators of C_{DS} and $C_{\mathrm{DS}}^{\perp}$, respectively, by

$$\begin{aligned} B_{i,j} &= B_{i,j}(C_{\mathrm{DS}}) \\ &= |\{\mathbf{w} \in C_{\mathrm{DS}} : \mathrm{wt}(w_1, \ldots, w_n) = i, \\ &\quad \mathrm{wt}(w_{n+1}, \ldots, w_{n+m+r}) = j\}|, \end{aligned} \tag{13}$$

and

$$\begin{aligned} B_{i,j}^{\perp} &= B_{i,j}(C_{\mathrm{DS}}^{\perp}) \\ &= |\{\mathbf{w} \in C_{\mathrm{DS}}^{\perp} : \mathrm{wt}(w_1, \ldots, w_n) = i, \\ &\quad \mathrm{wt}(w_{n+1}, \ldots, w_{n+m+r}) = j\}|. \end{aligned} \tag{14}$$

The minimum distance d of Q_{DS} implies that

$$B_{i,0}^{\perp} = \sum_{j=0}^{m+r} B_{i,j}, \quad \text{for } i = 1, \ldots, d-1. \tag{15}$$

We will say that Q_{DS} is a *degenerate* quantum DS code if there exists $B_{i,j} > 0$ for $i < d$. Otherwise, it is a *nondegenerate* quantum DS code. Clearly, we also have

$$B_{i,0}^{\perp} \geqslant \sum_{j=1}^{m+r} B_{i,j}, \quad i = d, \ldots, n, \tag{16}$$

$$B_{0,0} = B_{0,0}^{\perp} = 1 \text{ and } B_{i,0} = 0 \quad \text{for } i \geqslant 1. \tag{17}$$

For $0 \leqslant i \leqslant m+r$, let us define $d(i)$ as the smallest integer such that

$$\begin{cases} B_{d(0),0}^{\perp} > \sum_{j=1}^{m+r} B_{d(0),j}, \\ B_{d(i),i}^{\perp} > 0, \quad \text{for } i = 1, \ldots, m+r. \end{cases} \tag{18}$$

Then the minimum distance of $Q_{\rm DS}$ is

$$d = \min_{0 \leqslant i \leqslant m+r} d(i) + i. \quad (19)$$

Denote the q-ary Krawtchouk polynomial of degree i by

$$K_i(x; n, q) = \sum_{j=0}^{i} (-1)^j (q-1)^{i-j} \binom{x}{j} \binom{n-x}{i-j}. \quad (20)$$

We list the properties of Krawtchouk polynomials needed in this work in Appendix A. Their proof and other information on these polynomials can be found in [16]–[18]. Let $f(x, y)$ be a two-variable polynomial and its maximal degrees of x and y be $d_x \leqslant n$ and $d_y \leqslant m + r$, respectively. Then the following Krawtchouk expansion of this polynomial holds:

$$f(x, y) = \sum_{i=0}^{d_x} \sum_{j=0}^{d_y} f_{i,j} K_i(x; n, q_1) K_j(y; m + r, q_2), \quad (21)$$

where

$$f_{i,j} = \frac{1}{q_1^n q_2^{m+r}} \sum_{x=0}^{n} \sum_{y=0}^{m+r} f(x, y) K_x(i; n, q_1) \times K_y(j; m + r, q_2). \quad (22)$$

Proofs of these equalities are straightforward generalizations of the proofs (see [17, Chapter 5]) for equivalent expressions for single variable polynomials.

In what follows we will need the following generalization of MacWilliams identities [17].

Theorem 1:

$$B_{x,y} = \frac{1}{4^n} \sum_{i=0}^{n} \sum_{j=0}^{m+r} B^{\perp}_{i,j} K_x(i; n, 4) K_y(j; m + r, 2). \quad (23)$$

A proof of this theorem can be found in Appendix B.

Like [13], [19], [20], for small n one could apply linear programming techniques to obtain upper bounds on the minimum distance of $[n, k : r]]$ DS codes. More explicitly, we have the following linear program: given n, k, d, and r,

$$\text{Find nonnegative intergers } B_{i,j}, B^{\perp}_{i,j}$$
$$\text{subject to } (15), (16), (17), \text{ and } (23).$$

If there is no solution to this feasibility problem, it means that no $[[n, k, d : r]]$ DS code exists.

Example 2: Let $n = 7, m = 6, r = 0$, and $d = 3$. MAPLE tells us that the liner program is feasible, and therefore a $[[7, 1, 3]]$ code may be capable of fighting a syndrome bit error by measuring six stabilizer generators. Indeed, a $[[7, 1, 3 : 6]]$ DS code exists [6].

VI. Upper Bounds on Unrestricted (Degenerate and Non-Degenerate) DS Codes

In this Section we propose a general method defined in Theorem 3 for deriving upper bounds on the minimum distance of both non-degenerate and degenerate DS codes. Next we use this method for obtaining several explicit bounds for DS codes with $r = 0$. Theorem 3 can be used for deriving bounds in the case of $r > 0$, but this will be done in future work.

Let $1 \leqslant d_{\rm D} \leqslant n$ be an integer,

$$\mathcal{N} = \{(i, j) : 0 \leqslant i \leqslant n, 1 \leqslant j \leqslant m + r\},$$

and $\mathcal{A} \subset \mathcal{N}$ and $\bar{\mathcal{A}} = \mathcal{N} \backslash \mathcal{A}$. We want to upper bound *quantum code rate* $R = k/n$ under the conditions:

$$B^{\perp}_{i,0} = \sum_{j=0}^{m+r} B_{i,j}, \quad i = 0, \ldots, d_{\rm D} - 1, \quad (24)$$

$$B^{\perp}_{i,j} = 0, \quad (i, j) \in \mathcal{A}. \quad (25)$$

Theorem 3: Let $f(x, y)$ be a polynomial with $f_{i,j} \geqslant 0$ satisfying the conditions:

$$f(x, 0) \leqslant 0, \quad \text{if } x \geqslant d_{\rm D}, \text{ and} \quad (26)$$

$$f(x, y) \leqslant 0, \quad \text{if } (x, y) \in \bar{\mathcal{A}}. \quad (27)$$

Then the following claims hold.

1) For non-degenerate $C^{\perp}_{\rm DS}$, it must hold that

$$f(0, 0)/f_{0,0} \geqslant 2^{2n}. \quad (28)$$

2) For unrestricted $C^{\perp}_{\rm DS}$, it must hold that

$$\max \left\{ \frac{f(0, 0)}{f_{0,0}}, \max_{1 \leqslant x \leqslant d_{\rm D}-1} \frac{f(x, 0)}{\min_{1 \leqslant j \leqslant m+r} f_{x,j}} \right\} \geqslant 2^{2n}. \quad (29)$$

Proof: We prove the second claim. Let $M = |C^{\perp}_{\rm DS}| = 2^{2n}$. Using (23), (21), and (16), we get

$$\begin{aligned} M & \sum_{i=0}^{d_{\rm D}-1} \sum_{j=0}^{m+r} f_{i,j} B_{i,j} \\ & \leqslant M \sum_{i=0}^{n} \sum_{j=0}^{m+r} f_{i,j} B_{i,j} \quad (30) \\ & = M \sum_{i=0}^{n} \sum_{j=0}^{m+r} f_{i,j} \\ & \quad \times \frac{1}{M} \sum_{x=0}^{n} \sum_{y=0}^{m+r} B^{\perp}_{x,y} K_i(x; n, 4) K_j(y; m + r, 2) \\ & = \sum_{x=0}^{n} B^{\perp}_{x,0} f(x, 0) + \sum_{(i,j) \in \mathcal{A}} B^{\perp}_{i,j} f(i, j) \\ & \quad + \sum_{(i,j) \in \overline{\mathcal{A}}} B^{\perp}_{i,j} f(i, j) \\ & \leqslant \sum_{x=0}^{d_{\rm D}-1} B^{\perp}_{x,0} f(x, 0) = \sum_{x=0}^{d_{\rm D}-1} \sum_{j=0}^{m+r} B_{x,j} f(x, 0). \quad (31) \end{aligned}$$

From this and (17), we get

$$\begin{aligned} 2^{2n} & \leqslant \frac{\sum_{x=0}^{d_{\rm D}-1} \sum_{j=0}^{m+r} B_{x,j} f(x, 0)}{\sum_{i=0}^{d_{\rm D}-1} \sum_{j=0}^{m+r} f_{i,j} B_{i,j}} \\ & \leqslant \max \left\{ \frac{f(0, 0)}{f_{0,0}}, \max_{1 \leqslant x \leqslant d_{\rm D}-1} \frac{f(x, 0)}{\min_{1 \leqslant j \leqslant m+r} f_{x,j}} \right\}. \quad (32) \end{aligned}$$

■

For getting a bound on the size of DS codes with minimum distance d, it suffices to choose

$$d_{\rm D} = d, \quad \mathcal{A} = \{(i, j) : j \geq 1, 0 \leq i + j \leqslant d - 1\}, \quad (33)$$

and a polynomial $f(x,y)$ that satisfies constraints (26) and (27). In the following subsections, we discuss two polynomials and their corresponding bounds on DS codes with $r=0$.

We also have upper bounds for a DS code inherited from its underlying stabilizer code. Let $Q_{\rm DS}$ be a DS code defined by (9) ($r=0$) with minimum distance $d(Q_{\rm DS})$. Let C be the $[n,n-k]$ code with generator matrix H used in (6) and $C^{\perp}$ be its dual code. Let Q be the $[[n,k]]$ stabilizer code defined by C with minimum distance $d(Q)$. From (10) it follows that vectors of the form

$$(\mathbf{v},\mathbf{0}),\quad \mathbf{v}\in C^{\perp},\ \mathbf{0}=(\underbrace{0,\ldots,0}_{n-k}),$$

form a subcode of $C_{\rm DS}^{\perp}$ and therefore $d(Q_{\rm DS})\leqslant d(Q)$. Thus any upper bound on degenerate $[[n,k]]$ stabilizer code Q is also an upper bound on the minimum distance of degenerate $[[n,k:0]]$ DS code. The same is true for non-degenerate codes.

A. Singleton Bound

As we mentioned in Section III code $C_{\rm DS}^{\perp}$ has size 2^{2n}, and if $\mathbf{v}=(v_1,\ldots,v_n,w_1,\ldots,w_{n-k})\in C_{\rm DS}^{\perp}$ then $v_i\in\mathbb{F}_4$ and $w_i\in\mathbb{F}_2$. This leads to the Singleton bound for nondegenerate DS codes.

Theorem 4: For any nondegenerate $[[n,k,d:0]]$ DS code, we have

$$k\leq n-2(d-1). \tag{34}$$

The proof of this theorem is a simple generalization of the well known case of codes over $\mathbb{F}_q$, see [21].

In [16] several upper bounds for degenerate stabilizer codes have been derived. In particular, the Singleton bound $k\leqslant n-2(d(Q)-1)$ has been proven. Thus we conclude that bound (34) also holds for degenerate DS codes.

It is instructive to prove this result using (29). To do this, we first note that if $f(x,y)=0$ for $y\geqslant 1$, then the coefficients $f_{i,j}$ do not depend on j. Indeed, let $f(x,y)=g(x)\delta_{y,0}$ and $f(x,0)=g(x)=\sum_{i=0}^{n}g_iK_i(x;n,4)$. Then, according to (21) and (61), we have

$$\begin{aligned} f_{i,j}&=\frac{1}{4^n2^m}\sum_{x=0}^{n}g(x)K_x(i;n,4)\sum_{y=0}^{m}\delta_{y,0}K_y(j;m,2)\\ &=\frac{1}{2^m}g_iK_0(j;m,2)=\frac{1}{2^m}g_i. \end{aligned}\tag{35}$$

Theorem 5 (Singleton Bound): For an unrestricted (non-degenerate or degenerate) DS code, we have

$$k\leq n-2(d-1). \tag{36}$$

Proof: We will use the polynomial

$$f(x,y)=\frac{4^{n-d+1}2^m}{\binom{n}{d-1}}\binom{n-x}{n-d+1}\delta_{y,0}. \tag{37}$$

Using (22) and (65), we obtain

$$\begin{aligned} f_{i,j}&=\frac{1}{4^n2^m}\frac{4^{n-d+1}2^m}{\binom{n}{d-1}}\sum_{x=0}^{n}\binom{n-x}{n-d+1}K_x(i;n,4)\\ &\quad\times\sum_{y=0}^{m}\delta_{y,0}K_y(j;m,2)=\frac{\binom{n-i}{d-1}}{\binom{n}{d-1}}\geqslant 0,\quad\forall i,j. \end{aligned}\tag{38}$$

It is easy to see that $f_{i,j}\geqslant 0$ and $f(i,0)=0$ for $i\geqslant d$. Simple computations show that

$$\frac{f(0,0)}{f_{0,0}}>\frac{f(l,0)}{f_{l,j}}=\frac{f(l,0)}{f_{l,0}}\quad\text{for } 1\leqslant l\leqslant d-1.$$

Finally, $f(0,0)/f_{0,0}=4^{n-d+1}2^m$. ∎

This approach gives us additional information on DS codes achieving the Singleton bound. For such a code, say $Q_{\rm DS,MDS}$, we must have equality in (30). Noticing that $f_{i,j}=0$ for $i>n-d+1$, we conclude that $Q_{\rm DS,MDS}$ must have $B_{i,j}=0$ for $d\leqslant n-d+1$ and $j\geqslant 0$. In (31) we always have equality since $f(x,0)=0$ if $x\geqslant d$. Finally, in order to have

$$\begin{aligned} \frac{\sum_{x=0}^{d_{\rm D}-1}\sum_{j=0}^{m}B_{x,j}f(x,0)}{\sum_{i=0}^{d_{\rm D}-1}\sum_{j=0}^{m}f_{i,j}B_{i,j}}&=\frac{\sum_{x=0}^{d_{\rm D}-1}f(x,0)\sum_{j=0}^{m}B_{x,j}}{\sum_{i=0}^{d_{\rm D}-1}f_{i,0}\sum_{j=0}^{m}B_{i,j}}\\ &=f(0,0)/f_{0,0} \end{aligned}$$

in (32), code $Q_{\rm DS,MDS}$ must have $\sum_{j=0}^{m}B_{x,j}=0$ for $1\leqslant x\leqslant d_{\rm D}-1$. Thus $B_{x,j}=0$ for $1\leqslant x\leqslant n-d_{\rm D}-1$ and $j\geqslant 0$. This means that any generator $\mathbf{g}$ of $Q_{\rm DS,MDS}$ should have large weight, $\mathrm{wt}\,(\mathbf{g})\geqslant n-d$. Hence such $Q_{\rm DS,MDS}$ will have large syndrome measurement error.

Extensive research have been conducted on construction of quantum codes meeting the Singleton bound (see for example [22], [23], [24], [25], references within, and numerous other papers on this subject). The above result however shows that such codes most likely will not be useful for practical applications due to their large syndrome measurement error probability.

B. Hamming Bound

Let $C_{\rm DS}^{\perp}$ be a non-degenerate DS code with minimum distance $d=2t+1$. Standard combinatorial arguments (see [6]) lead to that $k\leqslant\tilde{k}$, where $\tilde{k}$ is the largest integer such that

$$2^{2n}\leq\frac{4^n2^{n-\tilde{k}}}{\sum_{i=0}^{t}\binom{n}{i}3^i\sum_{j=0}^{t-i}\binom{n-\tilde{k}}{j}}. \tag{39}$$

This is the *Hamming Bound for non-degenerate* DS codes. Below we show that this bounds also holds for degenerate DS codes if n is sufficiently large. Let $d_{\rm D}$ and $\mathcal{A}$ be defined as in (33).

Lemma 6: For a positive integer λ, let $f^{(k)}(x,y)$ be the polynomial defined by the coefficients

$$f_{i,j}^{(k)}=\left(\sum_{a=0}^{t}\sum_{g=0}^{t-\lambda a}K_a(j;m,2)K_g(i;n,4)\right)^2.$$

Then

$$\begin{aligned} f^{(k)}(x,y)=4^n2^m\sum_{a=0}^{t}\sum_{b=0}^{t}\beta(y,a,b)\\ \times\sum_{g=0}^{t-\lambda a}\sum_{h=0}^{t-\lambda b}\sum_{w=0}^{n-x}\alpha(x,g,h,w), \end{aligned}$$

where

$$\alpha(x,g,h,w) = \binom{x}{2x+2w-g-h}\binom{n-x}{w} \times \binom{2x+2w-g-h}{x+w-h} 2^{g+h-2w-q} 3^w, \quad (40)$$

and

$$\beta(y,a,b) = \binom{m-y}{(a+b-y)/2}\binom{y}{(a-b+y)/2}. \quad (41)$$

A proof can be found in Appendix C.

Note that in the above lemma, λ is a parameter over which we will optimize our bound. Next we give an important property of the polynomial $f^{(k)}(x,y)$.

Lemma 7: For $x+y \geqslant d$, we have $f^{(k)}(x,y)=0$.

The proof is given in Appendix D.

Thus $f^{(k)}(x,y)$ satisfies constraints (26) and (27) and hence we can use it for obtaining a bound on the minimum distance of DS codes. Choosing $\lambda = 1$, we get a polynomial with $f^{(\tilde{k})}(0,0)/f^{(\tilde{k})}_{0,0}$ equal to the right hand side of (39). Numerical computations show that for large n, the first entry in the set defined in (29) dominates. Thus, for large n, this polynomial gives the Hamming bound (39) for unrestricted (non-degenerate and degenerate) DS codes. The "disadvantage" of this polynomial is that its coefficients $f^{(k)}_{i,j}$ may aggressively decrease with j, which for certain parameters makes $\min_{1\leqslant j\leqslant m} f(x,j)$ in (29) being very small, that results in a loose bound.

If we choose $\lambda = t+1$, we get $f^{(k)}(x,y) = f(x)\delta_{0,y}$, where $f(x)$ is the polynomial with $f_i = (\sum_{g=0}^{t} K_g(i;n,4))^2$, that is the polynomial that leads to the Hamming bound for classical codes over $\mathbb{F}_4$, see [17, Chapter 17]. For this polynomial, the value $f^{(k)}(0,0)/f^{(k)}_{0,0}$ is larger than in the case of $\lambda = 1$. However, its advantage is that its coefficients $f^{(k)}_{i,j}$ do not decrease with j (in fact they do not depend on j), which often leads to better bound than with $\lambda = 1$.

We conclude that for obtaining the best bound we have to find an optimal value $\lambda \in [1, t+1]$.

Theorem 8 (Hamming Bound for Unrestricted DS Codes): For an unrestricted DS code, we have $k \leqslant \bar{k}$, where $\bar{k}$ is the largest integer such that

$$\min_{1\leqslant\lambda\leqslant t+1} \max\left\{ \frac{f^{(\bar{k})}(0,0)}{f^{(\bar{k})}_{0,0}}, \max_{1\leqslant x\leqslant d_{\mathrm{D}}-1} \frac{f^{(\bar{k})}(x,0)}{\min_{1\leq j\leq m} f^{(\bar{k})}_{x,j}} \right\} \geqslant 2^{2n}. \quad (42)$$

For $d=7$, the Hamming bounds (39) and (42) are shown in Fig. 3. For small values of n, bound (42) is only marginally weaker than (39), and for $n \geqslant 36$, these bounds coincide. We observed the same behavior for other values of d. So we make the following conjecture.

Conjecture 9: For any d, there exists $n(d)$ such that for $n \geqslant n(d)$, the Hamming bound (39) holds for unrestricted DS codes.

In [6] Fujiwara obtained a *hybrid Hamming bound* for nondegenerate DS codes that can correct any t_{D} data and t_{S} syndrome errors: $k \leqslant \hat{k}$, where $\hat{k}$ is the largest integer such that

$$2^{2n} \leqslant \frac{2^{2n}2^{n-\hat{k}}}{\sum_{i=0}^{t_{\mathrm{D}}}\sum_{j=0}^{t_{\mathrm{S}}}\binom{n}{i}3^i\binom{n-\hat{k}}{j}}. \quad (43)$$

We can also derive this hybrid bound using Theorem 3 with

$$\mathcal{A} = \{(i,j) : 0 \leqslant i \leqslant 2t_{\mathrm{D}} \text{ and } 1 \leqslant j \leqslant 2t_{\mathrm{S}}\},$$

and polynomial

$$f^{(k)}(x,y) = 4^n 2^m \sum_{i=0}^{t_{\mathrm{D}}}\sum_{j=0}^{t_{\mathrm{D}}}\sum_{h=0}^{n-x} \alpha(x,i,j,h) \times \sum_{u=0}^{t_{\mathrm{S}}}\sum_{v=0}^{t_{\mathrm{S}}} \beta(y,u,v).$$

Tedious but straightforward computations show that $f^{(k)}_{i,j} \geqslant 0$, $f^{(k)}(x,y)=0$ if $(x,y)\in\bar{\mathcal{A}}$, and that $f^{(\hat{k})}(0,0)/f^{(\hat{k})}_{0,0}$ is equal to the right hand side of (43).

Thus we obtained a different proof of (43). We cannot use this polynomial for degenerate DS codes, since for some $i \leqslant d_{\mathrm{D}}-1$, we have $f^{(k)}_{i,j}=0$. Finding good polynomials for deriving hybrid bounds on degenerate DS codes is an open problem.

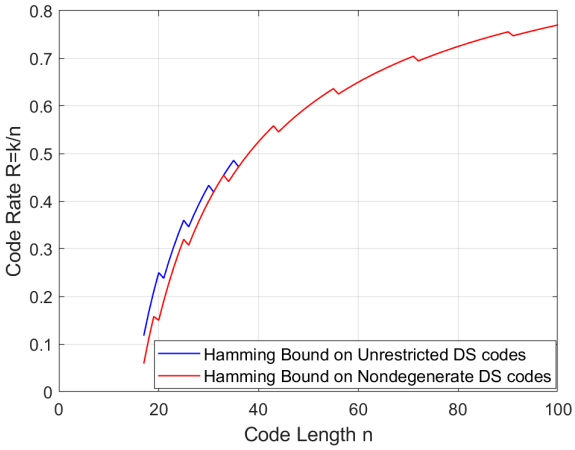

Fig. 3. Hamming bounds for nondegenerate and unrestricted DS codes, $d=7$.

C. Asymptotic Bounds

In this subsection, we consider the asymptotic regime in which both the code length n and the number of information qubits k tend to infinity but the code rate $R = k/n$ remains constant.

It is instructive to consider the Hamming bound (39) for nondegenerate DS codes in this regime. In order of doing this we have to find the leading term of the denominator of (39).

Recall that if v grows linearly with n and $a_{j^*} > a_j, j \neq j^*$, then

$$\frac{1}{n}\log_2 \sum_{j=1}^{v} 2^{na_j} = a_{j^*} + o(1). \quad (44)$$

Let $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ be the binary entropy function. Denoting $\xi = j/n, \tau = t/n$, and $\iota = i/n$

and using (44), we get for the second sum of the denominator of (39)

$$\begin{aligned}\frac{1}{n}\log_2 \sum_{j=0}^{t-i}\binom{n-\tilde{k}}{j} &= \frac{1}{n}\max_{0\leqslant\xi\leqslant\tau-\iota}\log_2 2^{(n-\tilde{k})H\left(\frac{\xi n}{n-\tilde{k}}\right)}+o(1)\\ &= \max_{0\leqslant\xi\leqslant\tau-\iota}(1-R)H\left(\frac{\xi}{1-R}\right)+o(1),\end{aligned}$$

where $o(1)$ is a function that tends to 0 as n increases and we have used Stirling's approximation, see [17], that

$$\frac{1}{n}\log_2\binom{n}{i} = H(i/n)+o(1).$$

This function achieves its maximum at $\xi=\frac{1}{2}(1-R)$. However, according to the Singleton bound, the *relative distance* $\delta \triangleq \frac{d}{n}\leqslant\frac{1}{2}(1-R)$ and therefore

$$\tau = t/n = \delta/2 \leqslant \frac{1}{4}(1-R)\leqslant\frac{1}{2}(1-R).$$

Thus the maximum is achieved at $\xi=\tau-\iota$. Hence for the denominator of (39) we have

$$\begin{aligned}&\frac{1}{n}\log_2\sum_{i=0}^{t}\binom{n}{i}3^i\sum_{j=0}^{t-i}\binom{n-\tilde{k}}{j}\\ &\quad= \max_{0\leqslant\iota\leqslant\tau} H(\iota)+\iota\log_2(3)+(1-R)H\left(\frac{\tau-\iota}{1-R}\right)+o(1).\end{aligned} \tag{45}$$

Taking the derivative and finding its roots, we conclude that the maximum is achieved at

$$\iota^* = 1-\frac{1}{4}R+\frac{1}{2}\tau-\frac{1}{4}\sqrt{16-8R-8\tau+R^2-4R\tau+4\tau^2}. \tag{46}$$

One can show that ι^* is always smaller than τ. Thus the exponent of the denominator of (39) is

$$H(\iota^*)+\iota^*\log_2(3)+(1-R)H((\tau-\iota^*)/(1-R))+o(1).$$

The exponents of the left part and the numerator of (39) are 2 and $\frac{1}{n}\log_2 4^n 2^{n-\tilde{k}} = 3-R$, respectively. Combining the above results, we obtain the following theorem.

Theorem 10: For a given δ, the code rate R cannot exceed the root, say $R_{Ham,nondeg}(\delta)$, of

$$H(\iota^*)+\iota^*\log_2(3)+(1-R)H\left(\frac{\delta/2-\iota^*}{1-R}\right)+R-1=0. \tag{47}$$

In [16] the Hamming and so-called first linear programming (LP1) bounds have been derived in asymptotic form for unrestricted (degenerate and non-degenerate) quantum codes:

$$R\leqslant 1-\delta/2\log_2(3)-H(\delta/2)+o(1),\ \text{for } 0\leqslant\delta\leqslant 1/3,\quad \text{(Hamming)} \tag{48}$$

$$R\leqslant H(w)+w\log_2(3)-1+o(1),\ \text{(LP1)} \tag{49}$$

where $w=\frac{3}{4}-\frac{1}{2}\delta-\frac{1}{2}\sqrt{3\delta(1-\delta)}$, for $0\leqslant\delta\leqslant 0.3152$. The Hamming bound was obtained by applying the polynomial $f_{Ham}(x)$ defined by its coefficients $f_i=K_{\delta/2}(i)^2$. LP1 bound was obtained with the help of the polynomial

$$f_{LP1}(x)=\frac{1}{a-x}(K_{t+1}(x)K_t(a)-K_t(x)K_{t+1}(a))^2,$$

where $t=\frac{\delta}{2}n$ and a is a real number located between the first roots r_{t+1} and r_t of $K_{t+1}(x)$ and $K_t(x)$, and chosen so that $K_t(a)/K_{t+1}(a)=-1$.

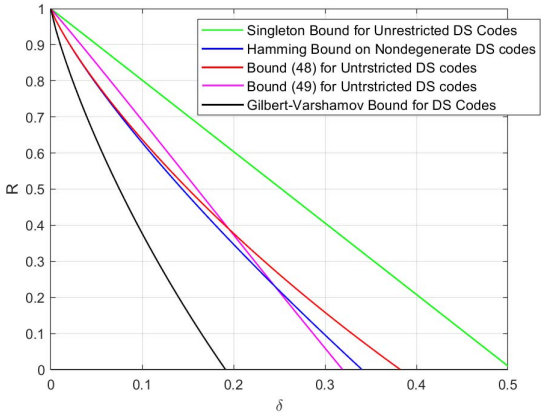

Fig. 4. Upper bounds on unrestricted DS codes and an achievability bound on DS codes.

As we mentioned prior to Section VI-A, any bound on degenerate quantum codes is also a bound on degenerate DS codes with the corresponding n and k. Thus bounds (48) and (49) also hold for unrestricted (degenerate or non-degenerate) DS codes. Note that these bounds can be also obtained using Theorem 3 and polynomials $f(x,y)=f_{Ham}(x)\delta_{y,0}$ and $f(x,y)=f_{LP1}(x)\delta_{y,0}$. As we showed in (35) the coefficients $f_{i,j}$ of these polynomials do not depend on j.

The bounds (36), (47), (48), and (49) for unrestricted DS codes are shown in Fig. 4. One can see that at certain interval the Hamming bound for non-degenerate quantum codes beats all the bounds for degenerate DS codes.

It looks natural to try to improve bounds (48), and (49) by using polynomials $f(x,y)$ whose coefficients $f_{i,j}$ depend on both indices i and j. At this moment we did not find such polynomials and leave this as an interesting open problem.

VII. Random DS Codes

The enumerators $B^{\perp}_{i,j}$ define the decoding error probability of a DS code in a number of communication/computational scenarios, similar to [26]. Below we study the behavior of $B_{i,j}$ and $B^{\perp}_{i,j}$ of random DS codes. In particular, we are interested in how the normalized minimum distance $d(r)/n$ depends on the ratio r/n when $n\to\infty$.

We will consider the ensemble $\mathcal{E}_{n,k,r}$ of $C_{\rm DS}$ codes defined by matrices of the form (9) with $r\leqslant n-k$ and full rank matrices A, i.e., $\mathrm{rank}(A)=r$. We will use this ensemble to show that the minimum distance of random DS codes with a relatively small $r\leqslant n-k$ achieves the Gilbert-Varshamov bound of stabilizer codes [27].

Let $\mathcal{E}^{\perp}_{n,k,r}$ be the ensemble of $C^{\perp}_{\rm DS}$ codes that are dual to codes from $\mathcal{E}_{n,k,r}$. Note that $|\mathcal{E}^{\perp}_{n,k,r}|=|\mathcal{E}_{n,k,r}|$. Define the

average enumerators (weight distribution) of codes from $\mathcal{E}_{n,k,r}$ and $\mathcal{E}^{\perp}_{n,k,r}$, respectively, by

$$\overline{B}_{i,j} = \frac{1}{|\mathcal{E}_{n,k,r}|} \sum_{C \in \mathcal{E}_{n,k,r}} B_{i,j}(C), \text{ and}$$
$$\overline{B}^{\perp}_{i,j} = \frac{1}{|\mathcal{E}^{\perp}_{n,k,r}|} \sum_{C^{\perp} \in \mathcal{E}^{\perp}_{n,k,r}} B_{i,j}(C^{\perp}),$$

where $B_{i,j}(C)$ and $B_{i,j}(C^{\perp})$ are defined in (13) and (14). The following theorem finds these weight distributions explicitly.

Theorem 11: For $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m+r$, we have

$$\overline{B}_{0,0} = 1, \quad \overline{B}_{i,0} = 0,$$
$$\overline{B}_{0,j} = \frac{1}{2^m - 1}\left(\binom{m+r}{j} - \binom{m}{j} - \binom{r}{j}\right), \tag{50}$$
$$\overline{B}_{i,j} = \frac{1}{(4^n-1)(2^m-1)}\binom{n}{i}3^i\left((2^m-2)\binom{r+m}{j} + \binom{m}{j} + \binom{r}{j}\right), \tag{51}$$
$$\overline{B}^{\perp}_{i,0} = \frac{1}{(4^n-1)(2^m-1)}\binom{n}{i}3^i\left(4^n - 2^m + 1 - \frac{4^n}{2^m}\right), \tag{52}$$
$$\overline{B}^{\perp}_{i,j} = \frac{4^n}{(4^n-1)2^{r+m}(2^m-1)}\binom{n}{i}3^i\left(\binom{m+r}{j}2^m - \binom{r}{j}2^m - \binom{m}{j}2^r\right), \tag{53}$$
$$\overline{B}^{\perp}_{0,0} = 1, \quad \overline{B}^{\perp}_{0,j} = 0. \tag{54}$$

A combinatorial proof of this result can be found in Appendix E.

Let us now consider the asymptotic case when the code length $n \to \infty$. Again, denote

$$\iota = i/n, \quad \xi = j/n, \text{ and } \rho = r/n.$$

For a DS code with $B_{i,j}$ and $B^{\perp}_{i,j}$, we define

$$b_{\iota,\xi} = \frac{1}{n}\log_2 B_{\lfloor \iota n \rfloor, \lfloor \xi n \rfloor}, \quad b^{\perp}_{\iota,\xi} = \frac{1}{n}\log_2 B^{\perp}_{\lfloor \iota n \rfloor, \lfloor \xi n \rfloor},$$

and $\delta = d/n$, where the minimum distance d is defined by (18) and (19). Denote by d_Q the minimum distance of a generic quantum code. It was shown in [27] that there are quantum codes, and quantum stabilizer codes in particular, whose normalized minimum distance $\delta_Q = d_Q/n$ is at least as large as the quantum Gilbert-Varshamov (GV) bound $\delta_{GV}(R)$. This bound is defined by the equation

$$H(\delta_{GV}(R)) + \delta_{GV}(R)\log_2(3) = 1 - R.$$

In the next theorem, we prove that there exist DS codes whose weight distributions $B_{i,j}$ and $B^{\perp}_{i,j}$ are upper bounded by the analytical expressions presented in the theorem for all i and j, and present a GV bound $\delta_{DS,GV}$ for such codes.

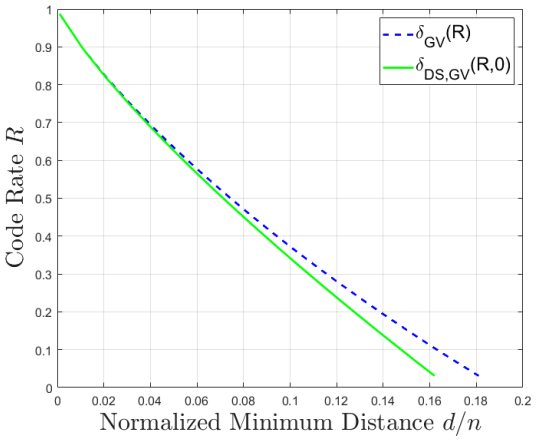

Fig. 5. Gilbert-Varshamov bounds for stabilizer and DS codes with $\rho = 0$.

Theorem 12: For $r \leqslant m$, there exist DS codes with rate R and $b_{\iota,\xi} \leqslant \overline{b}_{\iota,\xi}$ and $b^{\perp}_{\iota,\xi} \leqslant \overline{b}^{\perp}_{\iota,\xi}$, where

$$\overline{b}_{0,\xi} = (1-R+\rho)H\left(\frac{\xi}{1-R+\rho}\right) - 1 + R + o(1), \tag{55}$$
$$\overline{b}_{\iota,\xi} = H(\iota) + \iota\log_2(3) + (1-R+\rho)H\left(\frac{\xi}{1-R+\rho}\right) - 2 + o(1), \tag{56}$$
$$\overline{b}^{\perp}_{\iota,\xi} = H(\iota) + \iota\log_2(3) + (1-R+\rho)H\left(\frac{\xi}{1-R+\rho}\right) - (1-R+\rho) + o(1), \tag{57}$$
$$\overline{b}^{\perp}_{\iota,0} = H(\iota) + \iota\log_2(3) - 1 + R + o(1), \tag{58}$$

and the normalized minimum distance

$$\delta_{\mathrm{DS}} \geqslant \delta_{\mathrm{DS,GV}}(R,\rho),$$

where

$$\delta_{DS,GV}(R,\rho) = \min\left\{d_{GV}(R), \min_{\iota} \iota + H^{-1}\left(1 - H\left(\frac{\iota + \iota\log_2(3)}{1-R+\rho}\right)\right)\right\}. \tag{59}$$

A proof can be found in Appendix F.

It is instructive to compare the bounds $\delta_{DS,GV}(R,\rho)$ and $\delta_{GV}(R)$. In Fig. 5, we plot these bounds for the case $\rho = 0$. One can see that $\delta_{DS,GV}(R,0) < \delta_{GV}(R)$, especially for low rate quantum codes. This means that DS codes with $\rho = 0$ have inferior performance compared to stabilizer codes (in which only qubits are vulnerable to errors). However, we can improve DS codes by taking nonzero ρ. It is not difficult to see that $\delta_{DS,GV}(R,\rho)$ grows with ρ. So, for each R we can choose $\rho^*(R)$ so that $\delta_{DS,GV}(R,\rho^*(R)) = \delta_{GV}(R)$. It happened that $\rho^*(R) < 1 - R$ for any R (that is the corresponding $r^*(R) = \rho^*(R)n < n - k$, what we assumed for ensemble $\mathcal{E}_{n,k,r}$). In Fig. 6 we plot the normalized length of syndrome $\mu = m/n = 1 - R$ for stabilizer codes and $\mu + \rho^*(R) = 1 - R + \rho^*(R)$ for DS codes. One can observe that $\rho^*(R)$ is not very large even for low rate quantum codes. This means that relatively small number of additional generator measurements are needed for achieving the quantum GV bound by DS codes.

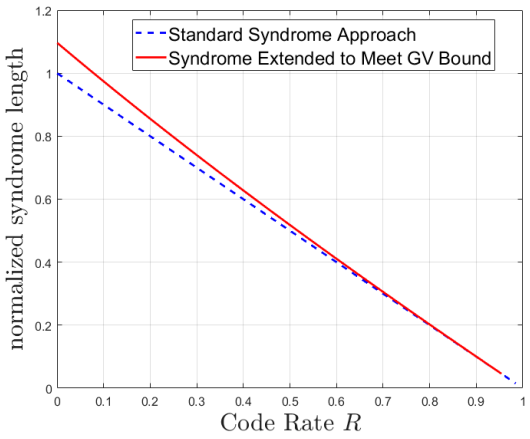

Fig. 6. Normalized number of syndrome bits $(n-k)/n$ for standard stabilizer codes and extended syndrome bits $(n-k+r)/n$ for DS codes, where r is chosen such that $\delta_{GV,DS}(R, r/n) = \delta_{GV}(R)$.

VIII. CSS-Type Quantum DS Codes

In this section we discuss CSS-type DS codes with $r \geq 0$. Suppose that

$$H_{\text{CSS}} = \left(\begin{array}{c|c} H & 0 \\ 0 & H \end{array}\right)$$

defines an $[[n, k, d]]$ CSS code, where H is a binary $(\frac{n-k}{2}) \times n$ matrix and $HH^T = \mathbf{0}$. Let

$$H_{\text{DS}} = \left(\begin{array}{cc|cc} H' & I_{(m+r)/2} & 0 & 0 \\ 0 & 0 & H' & I_{(m+r)/2} \end{array}\right), \qquad (60)$$

where

$$H'^T = \begin{pmatrix} H^T & \mathbf{f}_1^T & \cdots & \mathbf{f}_{r/2}^T \end{pmatrix},$$

and vectors $\mathbf{f}_j$ are obtained as linear combinations of rows of H. The matrix $\tilde{H} = [H' \; I_{(m+r)/2}]$ defines a classical $[n', k', d']$ code. The minimum distance of the corresponding DS code is $d' \leq d$ and therefore we obtain an $[[n, k, d' : r]]$ quantum DS code. Below we discuss how to extend H to H' so that the minimum distance of the DS code would not decrease and remain equal to d.

For a vector $\mathbf{y} \in \mathbb{F}_2^{n+(m+r)/2}$ we define the extended syndrome as $\mathbf{s} = (s_1, \ldots, s_{(m+r)/2}) = H'\mathbf{y}$. One can see that these syndromes belong to the column space of H'. Hence if any nonzero vector $\mathbf{w}$ from the column space of H' has weight $\text{wt}(\mathbf{w}) \geqslant d$, then for any two extended syndromes, say $\mathbf{s}$ and $\mathbf{s}'$, we have $\text{dist}(\mathbf{s}, \mathbf{s}') \geqslant d$ and hence the DS code can correct any $\lfloor \frac{d-1}{2} \rfloor$ syndrome bit errors. If the CSS code defined by H_{CSS} also has minimum distance d or larger then the DS code can correct any combination of qubit and syndrome errors whose total number does not exceed $\lfloor \frac{d-1}{2} \rfloor$. This leads us to the following result.

Theorem 13: If there exists an $[n, k, d]$ classical dual-containing cyclic code C with $2k > n$, then there exists an $[[n, 2k-n, d : r]]$ quantum DS code with $r \leq 2k$.

Proof: Suppose that H is an $n \times (n-k)$ parity-check matrix of C. Since C contains its dual code $C^\perp$, we have that $HH^T = 0$. Hence H can be used for construction of a CSS code according to H_{CSS}. Let $\mathbf{c} = (c_0 \; c_1 \; \cdots \; c_{n-1}) \in C^\perp$.

TABLE I

The Distance of $\tilde{H}$ Corresponding to Different Number of Rows

r	0	1	2	3	4	5	6	7	8	9
d	3	4	4	4	5	5	5	6	6	7

Since $C^\perp$ is also cyclic, any cyclic shift of $\mathbf{c}$ is also a codeword of $C^\perp$. Hence $n-k$ cyclic shifts of $\mathbf{c}$ can be used to construct H and additional k cyclic shifts can be used to construct H'. Thus, we can construct H' as follows

$$H' = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_1 & c_2 & \cdots & c_0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_0 & \cdots & c_{n-2} \end{pmatrix}.$$

Clearly, the column space and the row spaces of H' are the same and they generate code $C^\perp$. Since $C^\perp \subseteq C$ we have $d(C^\perp) \geqslant d$. Now from the arguments preceding this theorem, it follows that H' defines an $[[n, 2k-n, d : 2k]]$ DS code. ■

To demonstrate an application of the above theorem, we consider quadratic-residue (QR) codes. QR codes are cyclic codes and they are dual-containing for certain parameters [17], [28], and therefore can be used for construction of CSS quantum codes. In particular, they lead to $[[p, 1, d]]$ CSS codes with $d^2 - d + 1 \geq p$ for $p = 8j - 1$.

Theorem 14: For $p = 8j - 1$, there exist $[[p, 1, d]]$ CSS codes with $d^2 - d + 1 \geq p$. Moreover, there are $[[p, 1, d : r]]$ quantum DS codes with $r \leq p + 1$.

Example 15: Consider the QR code with $p = 23$. Suppose H' is cyclicly generated by $11 + r$ cyclic shifts of a code vector of the dual code. Table I provides the distances of the corresponding DS codes with different values of r. Table I shows that there exists $[[23, 1, 7 : 18]]$ quantum DS code and therefore we need only 18 additional redundant stabilizers, instead of 24 by Theorem 13.

The family of quantum QR codes in Theorem 14, which includes the Steane code and the quantum Golay code, are important in the theory of fault-tolerant quantum computation. In particular they are used for finding error thresholds [29]. Here we have shown that these codes also induce nontrivial quantum DS codes.

IX. Conclusion

In this work we proposed to use Syndrome Measurement (SM) and Data Syndrome (DS) codes for making the syndrome of quantum stabilizer codes robust against measurement errors. We demonstrated that if stabilizers of a code have a small spread of weights SM codes give large performance gain over repeated syndrome measurement approach. We next show that DS codes that can simultaneously correct both qubit data and syndrome errors. We derived upper bounds on their minimum distance and note that the bounds hold for both non-degenerate and degenerate DS codes. We next studied the weight distribution of random DS codes and used it for deriving lower (Gilbert-Varshamov) bound on the code minimum distance. Finally we presented a construction of CSS-type DS codes.

APPENDIX

A. Properties of Krawtchouk Polynomial

The following equalities holds (see [17, Chapter 5])

$$K_0(x;n,q) = 1, \tag{61}$$

$$K_j(0;n,q) = (q-1)^j \binom{n}{j}, \tag{62}$$

$$\sum_{i=0}^{n} K_r(i;n,c)K_i(s;n,c) = c^n \delta_{rs}, c = 2 \text{ or } 4, \tag{63}$$

$$\sum_{j=0}^{n} \binom{n}{j}(q-1)^j K_i(j;n,q) = q^n \delta_{i,0}, \tag{64}$$

$$\sum_{i=0}^{n} \binom{n-i}{n-j} K_i(x;n,q) = q^j \binom{n-x}{j}. \tag{65}$$

In [30, eq. A.19] and [16, Lemma 2], it is shown that

$$K_a(j;m,2)K_b(j;m,2) = \sum_{u=0}^{m} \beta(u,a,b)K_u(j;m,2), \tag{66}$$

$$K_g(i;n,4)K_h(i;n,4) = \sum_{q=0}^{n}\sum_{w=0}^{n-q} \alpha(q,g,h,w)K_q(i;n,4), \tag{67}$$

where $\beta(u,a,b)$ and $\alpha(q,g,h,w)$ are defined in (41) and (40) respectively.

Lemma 16:

$$\sum_{j=0}^{m} \binom{m}{j} K_u(j;n,2) = 2^m \binom{n-m}{u}. \tag{68}$$

Proof: The generating function of the binary Krawtchouk polynomials (see [17, Sec. 5.7]) is

$$(1+x)^{n-j}(1-x)^j = \sum_{u=0}^{n} K_u(j;n,2)x^u.$$

Using this equation, we obtain

$$\sum_{j=0}^{m} \binom{m}{j}(1+x)^{n-j}(1-x)^j = \sum_{u=0}^{n} x^u \sum_{j=0}^{m} \binom{m}{j} K_u(j;n,2).$$

At the same time

$$\begin{aligned}
&\sum_{j=0}^{m} \binom{m}{j}(1+x)^{n-j}(1-x)^j \\
&\quad = \sum_{j=0}^{m} \binom{m}{j}(1+x)^{n-m}(1+x)^{m-j}(1-x)^j \\
&\quad = (1+x)^{n-m}\sum_{j=0}^{m} \binom{m}{j}(1+x)^{m-j}(1-x)^j \\
&\quad = (1+x)^{n-m}2^m = 2^m \sum_{u=0}^{n-m} \binom{n-m}{u} x^u.
\end{aligned}$$

Comparing these two expressions, we finish the proof. ∎

B. Proof of Theorem 1

We can use the techniques in [31] as follows. We define a Fourier transform operator with respect to the inner product (5) and find a MacWilliams identity that relates the two split weight enumerators. Then Theorem 1 follows directly.

C. Proof of Lemma 6

Using (66) and (67), we obtain

$$\begin{aligned}
f_{i,j}^{(k)} &= \sum_{a=0}^{t}\sum_{b=0}^{t} K_a(j;m,2)K_b(j;m,2) \\
&\quad \times \sum_{g=0}^{t-\lambda a}\sum_{h=0}^{t-\lambda b} K_g(i;n,4)K_h(i;n,4) \\
&= \sum_{a=0}^{t}\sum_{b=0}^{t}\sum_{u=0}^{m} \beta(u,a,b)K_u(j;m,2) \\
&\quad + \sum_{g=0}^{t-\lambda a}\sum_{h=0}^{t-\lambda b}\sum_{q=0}^{n}\sum_{w=0}^{n-q} \alpha(q,g,h,w)K_q(i;n,4).
\end{aligned}$$

Now, using (21), we obtain

$$\begin{aligned}
&f^{(k)}(l,r) \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m} f_{i,j}^{(k)} K_i(l;n,4)K_j(r;m,2) \\
&= \sum_{i=0}^{n}\sum_{j=0}^{m}\left[\sum_{a=0}^{t}\sum_{b=0}^{t}\sum_{u=0}^{m} \beta(u,a,b)K_u(j;m,2)K_j(r;m,2)\right. \\
&\quad \left. + \sum_{g=0}^{n}\sum_{h=0}^{t-\lambda a}\sum_{q=0}^{t-\lambda b}\sum_{w=0}^{n-q} \alpha(q,g,h,w)K_q(i;n,4)K_i(l;n,4)\right] \\
&= \sum_{a=0}^{t}\sum_{b=0}^{t}\sum_{u=0}^{m} \beta(u,a,b)\sum_{j=0}^{m} K_u(j;m,2)K_j(r;m,2) \\
&\quad + \sum_{g=0}^{t-\lambda a}\sum_{h=0}^{t-\lambda b}\sum_{q=0}^{n}\sum_{w=0}^{n-q} \alpha(q,g,h,w) \\
&\quad \times \sum_{i=0}^{n} K_q(i;n,4)K_i(l;n,4) \\
&= 4^n 2^m \sum_{a=0}^{t}\sum_{b=0}^{t} \beta(r,a,b) \sum_{g=0}^{t-\lambda a}\sum_{h=0}^{t-\lambda b}\sum_{w=0}^{l-q} \alpha(l,g,h,w),
\end{aligned}$$

where in the last step we used the orthogonality property of Krawtchouk polynomials (63). ∎

D. Proof of Lemma 7

A binomial coefficient $\binom{i}{j}$ is assumed to be zero if: 1) $i < j$, 2) $j < 0$, or 3) j is not an integer.

The polynomial $f^{(k)}(x,y)$ is a sum of non negative terms:

$$\begin{aligned}
&\binom{m-y}{(a+b-y)/2}\binom{y}{(a-b+y)/2}\binom{x}{2x+2w-g-h} \\
&\qquad \times \binom{n-x}{w}\binom{2x+2w-g-h}{x+w-h} 2^{g+h-2w-q} 3^w.
\end{aligned}$$

A particular term is not zero if all the five binomial coefficients are not zeros. In the following discussion, we drop condition 3 since it is not needed for our purpose. It is not difficult to see that conditions 1 and 2 imply that $f^{(k)}(x,y) > 0$ for $x + y \geqslant 2t + 1$ only if there is a solution to the system of linear inequalities

$$A \cdot (a, b, w, g, h, x, y)^T \leqslant \mathbf{b},$$

with $A, \mathbf{b}$ given in (69) and (70), as shown at the bottom of this page. Conducting the Fourier–Motzkin elimination [32] in the order of h, y, g, a, b, w, x (other orders also work), we come to the incompatible condition $0 \leqslant -1/2$. This completes the proof. ∎

E. Proof of Theorem 11

We would like to analyze the weight distribution of a random DS code from $\mathcal{E}_{n,k,r}$ with a generator matrix of the form (9) where $m = n - k$. Let $\mathcal{E}_{n,m}$ be the set of DS codes with a generator matrix of the form $[H\ I_m]$, and $\mathcal{F}_{m,r}$ be the set of binary codes with a generator matrix of the form $[A\ I_r]$, where A has rank r. A code from $\mathcal{E}_{n,k,r}$ can be considered as a combination of codes from $\mathcal{E}_{n,m}$ and $\mathcal{F}_{m,r}$.

Lemma 17: The size of the ensemble $\mathcal{E}_{n,m}$ is

$$|\mathcal{E}_{n,m}| = |\mathcal{E}_{n,m}^{\perp}| = \prod_{u=0}^{m-1} \frac{(2^{2(n-u)} - 1)(2^m - 2^u)}{2^{u+1} - 1}, \quad (71)$$

and any vector $\mathbf{w} = (\mathbf{a}, \mathbf{b})$ with $\mathbf{a} \in \mathbb{F}_4^n \setminus \mathbf{0}$ and $\mathbf{b} \in \mathbb{F}_2^m \setminus \mathbf{0}$ is contained in

$$L = \prod_{u=1}^{m-1} \frac{(2^{2(n-u)} - 1)(2^m - 2^u)}{2^u - 1} \quad (72)$$

codes from $\mathcal{E}_{n,m}$.

Proof: It is proved in [26] that the number of $[n, m]$ additive self-orthogonal codes over $\mathbb{F}_4$ is

$$S \triangleq \prod_{u=0}^{m-1} \frac{(2^{2(n-u)} - 1)}{2^{u+1} - 1}.$$

For any $[n, m]$ additive self-orthogonal code, we can choose m generators (rows of matrix H) in

$$T \triangleq \prod_{u=0}^{m-1} (2^m - 2^u)$$

ways. Hence, using any $[n, m]$ self-orthogonal code, we can form T different matrices $[H\ I_m]$. Thus, $|\mathcal{E}_{n,m}| = ST$.

It is shown in [26] that any nonzero vector $\mathbf{a} \in \mathbb{F}_4^n \setminus \mathbf{0}$ is contained in

$$P = \prod_{u=1}^{m-1} \frac{2^{2(n-u)} - 1}{2^u - 1}$$

$[n, m]$ self-orthogonal codes. We can use any of those codes for building a code from $\mathcal{E}_{n,m}$ with vector $(\mathbf{a}, \mathbf{b})$ as its first basis vector. The other $(m-1)$ basis vectors can be chosen in

$$R = \prod_{u=1}^{m-1} (2^m - 2^u)$$

ways. Hence any $(\mathbf{a}, \mathbf{b})$ is contained in PR codes from $\mathcal{E}_{n,m}$. ∎

Lemma 18: The size of the ensemble $\mathcal{F}_{m,r}$ is

$$|\mathcal{F}_{m,r}| = \begin{bmatrix} m \\ r \end{bmatrix} \prod_{u=0}^{r-1} (2^r - 2^u) = \prod_{u=0}^{r-1} (2^m - 2^u), \quad (73)$$

where

$$\begin{bmatrix} m \\ r \end{bmatrix} = \frac{(2^m - 1)(2^{m-1} - 1) \cdots (2^{m-r+1} - 1)}{(2^r - 1)(2^{r-1} - 1) \cdots (2 - 1)}$$

is the *Gaussian binomial coefficient*, and any vector $(\mathbf{b}, \mathbf{c})$ for $\mathbf{b} \in \mathbb{F}_2^m \setminus \mathbf{0}$ and $\mathbf{c} \in \mathbb{F}_2^r \setminus \mathbf{0}$ is contained in

$$\begin{bmatrix} m-1 \\ r-1 \end{bmatrix} \prod_{u=1}^{r-1} (2^r - 2^u) = \prod_{u=1}^{r-1} (2^m - 2^u) \quad (74)$$

codes from $\mathcal{F}_{m,r}$.

The proof of this lemma is similar to the previous one and is omitted.

Lemma 19: The following four claims hold.

1) Any vector $(\mathbf{a}, \mathbf{b}, \mathbf{c})$, where $\mathbf{a} \in \mathbb{F}_4^n, \setminus \mathbf{0}$, $\mathbf{b} \in \mathbb{F}_2^m, \setminus \mathbf{0}$, $\mathbf{c} \in \mathbb{F}_2^r \setminus \mathbf{0}$, is contained in
$$F(\mathbf{a}, \mathbf{b}, \mathbf{c}) = L(2^m - 2) \prod_{u=1}^{r-1} (2^m - 2^u) \quad (75)$$
codes from $\mathcal{E}_{n,k,r}$.
2) Any $(\mathbf{a}, \mathbf{b}, \mathbf{0})$, where $\mathbf{a} \in \mathbb{F}_4^n \setminus \mathbf{0}$, $\mathbf{b} \in \mathbb{F}_2^m \setminus \mathbf{0}$, $\mathbf{0} \in \mathbb{F}_2^r$, is contained in
$$F(\mathbf{a}, \mathbf{b}, \mathbf{0}) = L \prod_{u=0}^{r-1} (2^m - 2^u) \quad (76)$$
codes from $\mathcal{E}_{n,k,r}$.

$$A^T = \begin{pmatrix}
-1 & 1 & -1 & 1 & & & & & -1 & 1 & & & 1 & & & & & \\
-1 & 1 & 1 & -1 & & & & & & & -1 & 1 & & & 1 & & & \\
 & & & & -2 & 2 & -1 & -1 & & & & & & & & & -1 & 1 & \\
 & & & & 1 & -1 & & 1 & & & & & -1 & 1 & & & & & \\
 & & & & 1 & -1 & 1 & & & & & & & & -1 & 1 & & & \\
 & & & & -2 & 1 & -1 & -1 & & & & & & & & & 1 & -1 \\
1 & 1 & -1 & -1 & & & & & & & & & & & & & & -1
\end{pmatrix} \quad (69)$$

$$\mathbf{b} = \begin{pmatrix} & 2m & & & & & & & t & & t & & t & & t & & n & -2t-1 \end{pmatrix} \quad (70)$$

3) Any $(\mathbf{0},\mathbf{b},\mathbf{c})$, where $\mathbf{0}\in\mathbb{F}_4^n$, $\mathbf{b}\in\mathbb{F}_2^m\setminus\mathbf{0}$, $\mathbf{c}\in\mathbb{F}_2^r\setminus\mathbf{0}$, is contained in
$$F(\mathbf{0},\mathbf{b},\mathbf{c})=ST\prod_{u=1}^{r-1}(2^m-2^u) \tag{77}$$
codes from $\mathcal{E}_{n,k,r}$.
4) Any $(\mathbf{a},\mathbf{0},\mathbf{c})$, where $\mathbf{a}\in\mathbb{F}_4^n,\backslash\mathbf{0}$, $\mathbf{0}\in\mathbb{F}_2^m$, $\mathbf{c}\in\mathbb{F}_2^r\setminus\mathbf{0}$, is contained in
$$F(\mathbf{a},\mathbf{0},\mathbf{c})=L\prod_{u=0}^{r-1}(2^m-2^u) \tag{78}$$
codes from $\mathcal{E}_{n,k,r}$.

Proof: We prove the first claim and the other three follow similarly.

A vector $(\mathbf{a},\mathbf{b},\mathbf{c})$ can be obtained only as the sum of a vector $(\mathbf{a},\mathbf{x},\mathbf{0})$ (where $\mathbf{x}\in\mathbb{F}_2^m\setminus\mathbf{0},\mathbf{x}\neq\mathbf{b}$, $\mathbf{0}\in\mathbb{F}_2^r$, and $(\mathbf{a},\mathbf{x})$ is a code vector of a code from $\mathcal{E}_{n,m}$) and a code vector $(\mathbf{0},\mathbf{b}+\mathbf{x},\mathbf{c})$ (where $(\mathbf{b}+\mathbf{x},\mathbf{c})$ is a code vector of a code from $\mathcal{F}_{m,r}$). Any given $(\mathbf{b}+\mathbf{x},\mathbf{c})$ is contained in $\prod_{u=1}^{r-1}(2^m-2^u)$ codes from $\mathcal{F}_{m,r}$ by Lemma 18. Since $(\mathbf{a},\mathbf{x})$ is contained in L codes from $\mathcal{E}_{n,m}$ and vector $\mathbf{x}$ can be chosen in 2^m-2 ways, we have $(\mathbf{a},\mathbf{b},\mathbf{c})$ contained in $L(2^m-2)\prod_{u=1}^{r-1}(2^m-2^u)$ codes from $\mathcal{E}_{n,k,r}$, which gives (75). ∎

In addition, we note that the total number of codes in $\mathcal{E}_{n,k,r}$ is
$$\begin{aligned}N=|S_{n,k,r}|&=ST\begin{bmatrix}m\\r\end{bmatrix}\prod_{u=0}^{r-1}(2^r-2^u)\\&=ST\prod_{u=0}^{r-1}(2^m-2^u).\end{aligned} \tag{79}$$

Now we have all the ingredients needed for finding $\overline{B}_{i,j}$. We again assume that $\binom{a}{b}=0$ if $a<b$ or $b<0$. Let consider the case of $i>0$ and $j>0$. Then
$$\begin{aligned}\overline{B}_{i,j}&=\frac{1}{N}\sum_{C\in\mathcal{E}_{n,k,r}}B_{i,j}(C)\\&=\frac{1}{N}\Bigg(\sum_{\substack{(\mathbf{a},\mathbf{b},\mathbf{0}):\\ \mathrm{wt}(\mathbf{a})=i,\ \mathrm{wt}(\mathbf{b})=j}}F(\mathbf{a},\mathbf{b},\mathbf{0})\\&\quad+\sum_{\substack{(\mathbf{a},\mathbf{0},\mathbf{c}):\\ \mathrm{wt}(\mathbf{a})=i,\ \mathrm{wt}(\mathbf{c})=j}}F(\mathbf{a},\mathbf{0},\mathbf{c})\\&\quad+\sum_{\substack{(\mathbf{a},\mathbf{b},\mathbf{c}):\\ \mathrm{wt}(\mathbf{a})=i,\ \mathrm{wt}(\mathbf{b})+\mathrm{wt}(\mathbf{c})=j}}F(\mathbf{a},\mathbf{b},\mathbf{c})\Bigg)\\&=\frac{1}{N}\binom{n}{i}3^i\left(\binom{m}{j}F(\mathbf{a},\mathbf{b},\mathbf{0})+\binom{r}{j}F(\mathbf{a},\mathbf{0},\mathbf{c})\right.\\&\quad\left.+F(\mathbf{a},\mathbf{b},\mathbf{c})\sum_{u=1}^{j-1}\binom{m}{u}\binom{r}{j-u}\right).\end{aligned}$$
Taking into account that
$$\sum_{u=0}^{j}\binom{m}{u}\binom{r}{j-u}=\binom{m+r}{j},$$
after some computations, we obtain (51). Equation (50) is obtained in a similar way.

To derive $\overline{B}_{i,j}^{\perp}$ we use MacWilliams identities (23), which also hold for average weight enumerators $\overline{B}_{i,j}$ and $\overline{B}_{i,j}^{\perp}$. Changing the role of codes C and $C^{\perp}$, we get, similar to (23):
$$\begin{aligned}\overline{B}_{i,j}^{\perp}&=\frac{1}{2^{m+r}(4^n-1)(2^m-1)}\sum_{l=1}^{n}\binom{n}{l}l^3\ K_i(l;n,4)\\&\quad\times\sum_{t=1}^{m+r}\left((2^m-2)\binom{r+m}{t}+\binom{m}{t}+\binom{r}{t}\right)\\&\quad\times K_j(t;m+r,2)\\&\quad+\frac{1}{2^{m+r}}K_i(0;n,4)\sum_{t=1}^{m+r}\left(\binom{m+r}{t}-\binom{m}{t}-\binom{r}{t}\right)\\&\quad\times K_j(t;m+r,2)+\frac{1}{2^{m+r}}K_i(0;n,4)K_j(0;m+r,2).\end{aligned}$$

Using (62), (64), and (68), after long manipulations, we obtain (52), (53), and (54). ∎

F. Proof of Theorem 12

According to Markov's inequality for a given pair i and j, we have
$$\begin{aligned}&\Pr\left(B_{i,j}(C^{\perp})\geqslant((n+1)(m+r+1))^{1+\epsilon}\overline{B}_{i,j}^{\perp}\right)\\&\qquad\leqslant\frac{1}{((n+1)(m+r+1))^{1+\epsilon}},\end{aligned}$$
for any $\epsilon>0$. Applying the union bound, we obtain
$$\begin{aligned}&\Pr(B_{i,j}(C^{\perp})\geqslant((n+1)(m+r+1))^{1+\epsilon}\overline{B}_{i,j}^{\perp}\\&\qquad\text{for at least one pair } i,j)\leqslant\frac{1}{((n+1)(m+r+1))^{\epsilon}},\end{aligned}$$
and further
$$\begin{aligned}&\Pr\left(B_{i,j}(C^{\perp})<((n+1)(m+r+1))^{1+\epsilon}\overline{B}_{i,j}^{\perp}\quad\text{for all } i,j\right)\\&\qquad\geqslant1-\frac{1}{((n+1)(m+r+1))^{\epsilon}}.\end{aligned}$$
Hence there exists a code $C^{\perp}\in\mathcal{E}_{n,m}^{\perp}$ such that
$$B_{i,j}(C^{\perp})\leqslant((n+1)(m+r+1))^{1+\epsilon}\overline{B}_{i,j}^{\perp},\quad\forall i,j. \tag{80}$$
Now we consider codes of growing lengths, i.e., $n\to\infty$. Note that $m/n=(n-k)/n=1-R$. Recall, see [17], that
$$\frac{1}{n}\log_2\binom{n}{i}=H(i/n)+o(1).$$
So the three terms of the last factor of (53) are that
$$\begin{aligned}\frac{1}{n}\log_2\binom{m+r}{j}2^m&=(1-R+\rho)H\left(\frac{\xi}{1-R+\rho}\right)\\&\quad+1-R,\end{aligned} \tag{81}$$
$$\frac{1}{n}\log_2\binom{r}{j}2^m=\rho H\left(\frac{\xi}{\rho}\right)+1-R, \tag{82}$$
$$\frac{1}{n}\log_2\binom{m}{j}2^r=(1-R)H\left(\frac{\xi}{1-R}\right)+\rho. \tag{83}$$

Simple analysis shows that for $\rho \leqslant 1-R$ (which is the same as $r \leqslant n-k$), we have that (81) is always larger than (82) and (83). Hence

$$\begin{aligned}\overline{b}^{\perp}_{\iota,\xi} &\triangleq \frac{1}{n}\log_2((n+1)(m+r+1))^{1+\epsilon}\overline{B}^{\perp}_{i,j}\\ &= \frac{1}{n}\log_2 \overline{B}^{\perp}_{i,j} + o(1)\\ &= H(\iota) + \iota\log_2(3) + (1-R+\rho)H\left(\frac{\xi}{1-R+\rho}\right)\\ &\quad -(1-R+\rho) + o(1).\end{aligned}$$

The equation (56) is obtained in a similar way.

Let us have $C^{\perp}$ that satisfies (80). Since $C^{\perp}$ is linear, all $B_{i,j}(C^{\perp})$ are integers. Hence if ι^* and ξ^* are such that $\overline{b}_{\iota,\xi} \leqslant 0$ for $\iota \leqslant \iota^*$ and $\xi \leqslant \xi^*$, then $B_{i,j}(C^{\perp}) = 0$ for $1 \leqslant i \leqslant (\iota^*-\epsilon)n$, $1 \leqslant j \leqslant (\xi^*-\omega)n$ for any $\epsilon, \omega > 0$ and sufficiently large n. It is not difficult to see that if $\iota \leqslant \delta_{GV}(R)$, then $\overline{b}^{\perp}_{\iota,0} \leqslant 0$. Similarly, if

$$\xi(\iota) = H^{-1}\left(1 - H\left(\frac{\iota+\iota\log_2(3)}{1-R+\rho}\right)\right),$$

then $\overline{b}^{\perp}_{\iota,\xi(\iota)} = 0$. Thus $B_{i,0}(C^{\perp}) = 0$ for all $i \leqslant (\delta_{GV}(R)-\epsilon)n$ and $B_{i,j}(C^{\perp}) = 0$ if $i + j \leqslant (\iota + \xi(\iota) - \omega)n$. Hence (59) follows. ∎

REFERENCES

[1] D. A. Lidar and T. A. Brun, Eds., *Quantum Error Correction*. Cambridge, U.K.: Cambridge Univ. Press, Oct. 2013.

[2] P. W. Shor, "Fault-tolerant quantum computation," in *Proc. 37th Annu. Symp. Theory Comput. Sci.* Los Alamitos, CA, USA: IEEE Press, 1996, pp. 56–65.

[3] A. Ashikhmin, C.-Y. Lai, and T. A. Brun, "Robust quantum error syndrome extraction by classical coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 546–550.

[4] A. Ashikhmin, C.-Y. Lai, and T. A. Brun, "Correction of data and syndrome errors by stabilizer codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2274–2278.

[5] W. Zeng, A. Ashikhmin, M. Woolls, and L. P. Pryadko, "Quantum convolutional data-syndrome codes," *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Cannes, France, Jul. 2019, pp. 1–5.

[6] Y. Fujiwara, "Ability of stabilizer quantum error correction to protect itself from its own imperfection," *Phys. Rev. A, Gen. Phys.*, vol. 90, Dec. 2014, Art. no. 062304.

[7] H. Bombín, "Single-shot fault-tolerant quantum error correction," *Phys. Rev. X*, vol. 5, Sep. 2015, Art. no. 031043.

[8] B. Brown, N. Nickerson, and D. Browne, "Fault-tolerant error correction with the gauge color code," *Nature Commun.*, vol. 7, p. 12302, Sep. 2016.

[9] N. P. Breuckmann, K. Duivenvoorden, D. Michels, and B. M. Terhal, "Local decoders for the 2D and 4D toric code," *Quant. Inf. Comput.*, vol. 17, nos. 3–4, p. 181, 2017.

[10] E. T. Campbell, "A theory of single-shot error correction for adversarial noise," *Quantum Sci. Technol.*, vol. 4, no. 2, Jan. 2019, Art. no. 025006.

[11] A. R. Calderbank and P. W. Shor, "Good quantum error-correcting codes exist," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 2, pp. 1098–1105, Jul. 2002.

[12] A. M. Steane, "Error correcting codes in quantum theory," *Phys. Rev. Lett.*, vol. 77, no. 5, pp. 793–797, Jul. 1996.

[13] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction via codes over $GF(4)$," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1369–1387, Jul. 1998.

[14] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed state entanglement and quantum error correction," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 5, pp. 3824–3851, 1996.

[15] E. Knill and R. Laflamme, "A theory of quantum error-correcting codes," *Phys. Rev. A, Gen. Phys.*, vol. 55, no. 2, pp. 900–911, 1997.

[16] A. Ashikhmin and S. Litsyu, "Upper bounds on the size of quantum codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1206–1215, May 1999.

[17] F. J. MacWilliams and N. J. A. Sloane, *The Theory Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[18] V. Levenshtein, "Krawtchouk polynomials and universal bounds for codes and designs in Hamming spaces," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1303–1321, Sep. 1995.

[19] C.-Y. Lai, T. A. Brun, and M. M. Wilde, "Duality in entanglement-assisted quantum error correction," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 4020–4024, Jun. 2013.

[20] C.-Y. Lai and A. Ashikhmin, "Linear programming bounds for entanglement-assisted quantum error-correcting codes by split weight enumerators," *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 622–639, Jan. 2018.

[21] R. C. Singleton, "Maximum distance q-nary codes," *IEEE Trans. Inf. Theory*, vol. IT-10, no. 2, pp. 116–118, Apr. 1964.

[22] B. Chen, S. Ling, and G. Zhang, "Application of constacyclic codes to quantum MDS codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1474–1484, Mar. 2015.

[23] X. He, L. Xu, and H. Chen, "New q-ary quantum MDS codes with distance bigger than $q/2$," *Quantum Inf. Process*, vol. 15, pp. 2745–2758, Jul. 2016.

[24] W. Fang and F. Fu, "Some new constructions of quantum MDS codes," *IEEE Trans. Inf. Theory*, vol. 65, no. 12, pp. 7840–7847, 2019.

[25] X. Shi, Q. Yue, and Y. Wu, "New quantum MDS codes with large minimum distance and short length from generalized Reed–Solomon codes," *Discrete Math.*, vol. 342, no. 7, pp. 1989–2001, 2019.

[26] A. Ashikhmin, "Fidelity lower bounds for stabilizer and CSS quantum codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3104–3116, Jun. 2014.

[27] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, "Quantum error detection. II. Bounds," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 789–800, May 2000.

[28] C.-Y. Lai and C.-C. Lu, "A construction of quantum stabilizer codes based on syndrome assignment by classical parity-check matrices," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 7163–7179, Oct. 2011.

[29] A. Paetznick and B. W. Reichardt, "Fault-tolerant ancilla preparation and noise threshold lower bounds for the 23-qubit Golay code," *Quant. Inf. Comput.*, vol. 12, pp. 1034–1080, Jun. 2012.

[30] R. Mceliece, E. Rodemich, H. Rumsey, and L. Welch, "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 2, pp. 157–166, Mar. 1977.

[31] C.-Y. Lai, M.-H. Hsieh, and H.-F. Lu, "On the MacWilliams identity for classical and quantum convolutional codes," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3148–3159, Aug. 2016.

[32] A. Schrijver, *Theory of Linear and Integer Programming*. Hoboken, NJ, USA: Wiley, 1998.