Canonical Form and Finite Blocklength Bounds for Stabilizer Codes

Dimiter Ostrev*

Abstract

First, a canonical form for stabilizer parity check matrices of arbitrary size and rank is derived. Next, it is shown that the closely related canonical form of the Clifford group can be computed in time $O(n^3)$ for n qubits, which improves upon the previously known time $O(n^6)$. Finally, the related problem of finite blocklength bounds for stabilizer codes and Pauli noise is studied. A finite blocklength refinement of the hashing bound is derived, and it is shown that no argument that uses guessing the error as a substitute for guessing the coset can lead to a significantly better achievability bound.

1 Introduction

Stabilizer codes [12, 6], quantum analogues of classical linear codes, are a widely studied method for protecting quantum states from noise. This article considers canonical forms and finite blocklength bounds for stabilizer codes.

Matrix decompositions and canonical forms have long been important in mathematics. In quantum error correction, various decompositions of Clifford group elements have been proposed, with applications ranging from compilation of quantum circuits to generation of uniformly random Clifford group elements; see for example [8, 5, 13, 2] and references therein for more details.

In another recent work [17], a joint matrix decomposition of a pair of orthogonal matrices plays an important role in the design and analysis of the proposed family of quantum Calderbank-Shor-Steane codes. The codes in [17] have a set of linearly dependent stabilizer measurements, and therefore can be used to correct errors on the qubits and errors in the syndrome simultaneously. The matrix decomposition in [17] gives information about both the quantum CSS code and the classical linear code that protects against errors in the syndrome.

It would be desirable to generalize the matrix decomposition of [17] from CSS to arbitrary stabilizer codes. This is achieved in the present article. Theorem 5 below gives a canonical form for stabilizer parity check matrices of arbitrary size and rank.

*German Aerospace Center (DLR), Institute of Communications and Navigation, Münchener Str. 20, 82234 Weßling, Germany

The main idea in the proof of existence of the matrix decomposition, both in [17] and in Theorem 5 in the present article, is to use a variant of Gaussian elimination with both row and column operations, with the additional restriction that the column operations are chosen so as to preserve the symplectic products of the rows at each step. Thus, the algorithm for computing the matrix decomposition is a close relative of the "disentangling" algorithms for the Clifford group, such as the one described in [16, Theorem 10.6 and Exercise 10.40] and more recently in [5, Section II].

To prove uniqueness of the matrix decomposition for stabilizer codes, the present article extends and simplifies the methods used previously in [5] to establish a canonical form for the Clifford group. [5] describes a collection of subgroups of the Borel group, parametrized by elements of the Weyl group. This collection of subgroups is then used to describe the canonical form and prove its uniqueness. See also [15, Section V] for a more detailed description of the Borel subgroup, the Weyl subgroup, and the Bruhat decomposition of the symplectic group.

The approach in [5] is not directly applicable to stabilizer parity check matrices; the obstacle is roughly speaking that the latter are not invertible and need not even have linearly indepedent rows. To overcome the difficulty, the present article defines a more general family of groups of lower triangular matrices. These groups are shown to arise naturally during Gaussian elimination, and are used to describe and prove uniqueness of the canonical form for stabilizer parity check matrices.

An approach centered on Gaussian elimination has various theoretical and practical advantages. Besides yielding a canonical form for stabilizer parity check matrices, it also gives new insights into the canonical form for the Clifford group.

For example, [5, Section II] gave a two-stage algorithm to compute the canonical form of the Clifford group. The first stage, a disentangling algorithm closely related to Gaussian elimination, takes time $O(n^3)$ for n qubits. However, the second stage takes time $O(n^6)$, so the overall runtime to compute the canonical form is $O(n^6)$.

The present article gives a more precise analysis of the Gaussian elimination stage. In Theorem 6 below, it is shown that Gaussian elimination produces the canonical form directly, without the need for post-processing. Thus, the overall runtime is reduced from $O(n^6)$ to $O(n^3)$, which roughly squares the number of qubits for which it is practically feasible to compute the canonical form.

Moreover, the present article gives a simplified description of the relevant subgroups and self-contained proofs. The exposition here relies only on the structure of Gaussian elimination, rather than on more advanced topics such as Bruhat decomposition and (B, N)-pairs.

Insights in the structure of stabilizer codes may be applied to further problems. For example, the canonical form in Theorem 5 gives an encoding circuit for a given stablizer parity check matrix. The canonical form can also be used to generate a uniformly random stabilizer parity check matrix of a given size and rank by an algorithm that uses the optimal number of random bits.

These applications are not considered in detail here; the argument for the encoding circuit is similar to [17, Section III], where a decomposition of CSS parity check matrices is used to derive the encoding circuit and the algorithm for generation of uniformly random stabilizer parity check matrices of given size and rank is similar to the algorithm in [5, Section III] for the generation of uniformly random Clifford group elements.

Instead, this article considers the following question: given Pauli noise affecting a finite number n of qubits, what is the highest rate of a stabilizer code that keeps the error probability below a target ϵ, and what is the lowest error probability of a stabilizer code of a given rate r?

Theorem 7 below gives a finite blocklength refinement of the hashing bound, applicable to arbitrary n-qubit Pauli noise. Then, Theorems 8 and 9 show how the general bound can be computed efficiently in the important special cases of n independent qubit erasure channels and n independent qubit depolarizing channels.

Given an achievability bound, it is natural to ask whether it is in some sense optimal. In classical error correction, there is a well developed theory of finite blocklength bounds [18]. In particular, in the special cases of the classical erasure and binary symmetric channels, the finite blocklength achievability bounds come with nearly matching finite blocklength converse bounds.

Unfortunately, the corresponding problem in quantum error correction is less well understood. [1] focuses on achievability bounds for stabilizer and CSS codes on the depolarizing channel, but does not give any converse bounds. [21] derives both achievability and converse bounds applicable to general quantum error correcting codes, not just stabilizer codes. However, the achievability and converse bounds of [21] nearly match only in the case of the qubit dephasing channel, which is closely related to the classical binary symmetric channel, and in the case of the quantum erasure channel with classical post processing, which is closely related to the classical erasure channel. On the other hand, for the depolarizing channel, there is a constant size gap between the achievability and converse bounds of [21], which corresponds to the gap between the hashing lower and dephasing upper bounds on the capacity of the depolarizing channel.

In this context, the present article makes the following contribution: it shows that the finite blocklength refinement of the hashing bound in Theorems 8 and 9 is nearly optimal among a certain class of arguments: those that use guessing the most likely error from the syndrome as a substitute for guessing the most likely stabilizer coset from the syndrome.

Such a result is important because the relaxation from guessing the coset to guessing the error is widely used in the literature on stabilizer codes. In particular, the previous finite blocklength achievability bound of [1] for stabilizer codes on the depolarizing channel also uses this relaxation.

The rest of the article is structured as follows: Section 2 covers preliminaries and some background material. Section 3 introduces a family of groups of lower triangular matrices, and Section 4 uses these groups to establish canonical forms for unrestricted and stabilizer parity check matrices and for the symplectic and Clifford group. Section 5 gives the finite blocklength refinement of the hashing bound and the proof that it is nearly optimal among the class of arguments that substitute guessing the error for guessing the coset. Section 6 concludes the article and gives some directions for future work.

2 Preliminaries and notation

This section introduces notation that will be used later on, and covers some background material on quantum stabilizer codes and Pauli channels. Some prior acquaintance with quantum computation and quantum information is assumed; see, for example, [16].

After some basic notation (subsection 2.1), the Pauli group (subsection 2.2) and Clifford group (subsection 2.3) are introduced, along with the associated homomorphisms to vectors and matrices over the field with two elements. Next, the class of Pauli channels is described in subsection 2.4; these channels apply a random Pauli error and possibly give some classical side information correlated with the error. This class of channels has a number of useful properties: it is closed under conjugation by the Clifford group (subsection 2.5), closed under preparation and measurement of a subset of the qubits in the computational basis (subsection 2.6), and closed under Pauli corrections conditional on the side information (subsection 2.7). Moreover, the diamond distance of a Pauli channel from the identity can be easily computed (subsection 2.8). The preceding observations show that the search for the optimal stabilizer code to correct Pauli noise admits a convenient combinatorial description (subsection 2.9). A relaxation of this combinatorial problem is given in the final subsection 2.10.

2.1 Some notation

Let $[n]$ denote the finite set $\{1,\ldots,n\}$. Let $\mathbb{F}_2$ denote the field with two elements and $\mathbb{F}_2^n$ denote the space of column vectors with n components from $\mathbb{F}_2$. For $i \in [n]$, let $e_{n,i}$ denote the i-th standard basis vector of $\mathbb{F}_2^n$. Let $\mathbb{F}_2^{m\times n}$ denote the space of $m \times n$ matrices over $\mathbb{F}_2$. Let I_n denote the $n \times n$ identity, and let Ω_n denote the $n \times n$ reverse diagonal matrix

$$\Omega_n = \sum_{i=1}^{n} e_{n,i} e_{n,n+1-i}^{\top} \tag{1}$$

2.2 The Pauli group

Consider a quantum system with n qubits. Let $X_{n,i}, Z_{n,i}$ denote the Pauli X and Z operations acting on the i-th qubit, and let

$$Q_{2n} = (X_{n,1},\ldots,X_{n,n},Z_{n,n},\ldots,Z_{n,1}) \tag{2}$$

be the $2n$-tuple of single qubit Pauli X and Z operations ordered in a particular way. The reason for this particular ordering will become clear later. Individual components of Q_{2n} will be denoted by $Q_{2n,i}$; for example, $Q_{2n,1} = X_{n,1}$ and $Q_{2n,2n} = Z_{n,1}$ are the Pauli X and Z operations on the first out of n qubits.

The Pauli group on n qubits $\mathcal{P}_n$ is generated by the components of Q_{2n} and by iI_{2^n}. Two elements P, P' of the Pauli group either commute or anti-commute; this will be denoted as follows:

$$\langle P, P' \rangle = \begin{cases} 1 & \text{if } PP' + P'P = 0 \\ 0 & \text{if } PP' - P'P = 0 \end{cases} \tag{3}$$

The operation $\langle , \rangle$ has the following properties

$$\langle P, P' \rangle = \langle P', P \rangle \tag{4}$$

$$\langle PP', P'' \rangle = \langle P, P'' \rangle \oplus \langle P', P'' \rangle \tag{5}$$

$$\langle P, P' \rangle = \left\langle UPU^{-1}, UP'U^{-1} \right\rangle \tag{6}$$

for any $P, P', P'' \in \mathcal{P}_n$ and any unitary U.

Finally, the map $\Theta_n : \mathcal{P}_n \to \mathbb{F}_2^{2n}$ given by

$$\Theta_n(P) = \begin{pmatrix} \langle Q_{2n,2n}, P \rangle \\ \vdots \\ \langle Q_{2n,1}, P \rangle \end{pmatrix} \tag{7}$$

is a surjective group homomorphism with kernel $\{\pm I, \pm iI\}$. Θ_n sends $Q_{2n,i}$ to $e_{2n,i}$. Additinally, the identity

$$\forall P, P' \in \mathcal{P}_n, \langle P, P' \rangle = \Theta_n(P)^\top \Omega_{2n} \Theta_n(P') \tag{8}$$

holds.

2.3 The Clifford group

The Clifford group on n qubits consists of unitary matrices that map Pauli group elements to Pauli group elements under conjugation:

$$\mathcal{C}_n = \{W \in \mathbb{C}^{2^n \times 2^n} : WW^\dagger = I \text{ and } \forall P \in \mathcal{P}_n, WPW^{-1} \in \mathcal{P}_n\} \tag{9}$$

The symplectic group consists of matrices in $\mathbb{F}_2^{2n \times 2n}$ that preserve the symplectic form $x^\top \Omega_{2n} y$:

$$\mathcal{SP}_{2n} = \left\{ C \in \mathbb{F}_2^{2n \times 2n} : C^\top \Omega_{2n} C = \Omega_{2n} \right\} \tag{10}$$

The map $\Psi_n : \mathcal{C}_n \to \mathcal{SP}_{2n}$ defined by

$$\forall i, j \in [2n], e_{2n,i}^\top \Psi_n(W) e_{2n,j} = \left\langle Q_{2n,2n+1-i}, WQ_{2n,j}W^{-1} \right\rangle \tag{11}$$

is a surjective group homomorphism with kernel $\{cP : c \in \mathbb{C}, P \in \mathcal{P}_n\}$. Additional useful properties of Ψ_n are

$$\Psi_n(W^{-1}) = \Omega_{2n} \Psi_n(W)^\top \Omega_{2n} \tag{12}$$

$$\forall P \in \mathcal{P}_n, \Theta_n(WPW^{-1}) = \Psi_n(W) \Theta_n(P) \tag{13}$$

2.4 Pauli channels

Definition 1. *An n qubit Pauli channel with side information (or just Pauli channel) is specified by the joint distribution p_{UV} of a random variable U taking values in $\mathbb{F}_2^{2n}$ and another discrete random variable V. The channel specified by the joint distribution p will be denoted by $\mathcal{N}_p$; it transforms n qubit states as follows:*

$$\rho \mapsto \mathcal{N}_p(\rho) = \sum_{u,v} p_{UV}(u,v) Q_{2n}^u \rho \left(Q_{2n}^u\right)^\dagger \otimes |v\rangle \langle v| \tag{14}$$

where $Q_{2n}^u = Q_{2n,1}^{u_1} Q_{2n,2}^{u_2} \dots Q_{2n,2n}^{u_{2n}}$.

Examples of Pauli channels are the qubit erasure channel and the qubit depolarizing channel. The qubit erasure channel with parameter δ transforms the state ρ of a single qubit to

$$\mathcal{E}_\delta(\rho) = (1-\delta)\rho \otimes |0\rangle \langle 0| + \frac{\delta}{4}(\rho + X\rho X + Z\rho Z + XZ\rho ZX) \otimes |1\rangle \langle 1| \tag{15}$$

The side information indicates whether an erasure has occurred or not. The qubit depolarizing channel with parameter δ transforms the state ρ of a single qubit to

$$\mathcal{D}_\delta(\rho) = (1-\delta)\rho + \frac{\delta}{3}(X\rho X + Z\rho Z + XZ\rho ZX) \tag{16}$$

In this case, there is no side information.

2.5 Conjugation of a Pauli channel by a Clifford group element

Conjugating a Pauli channel by a Clifford unitary gives another Pauli channel. Specifically, if (14) is conjugated by $W \in \mathcal{C}_n$ the resulting channel transforms n-qubit states as follows:

$$\begin{aligned}\rho \mapsto &\sum_{u,v} p_{UV}(u,v) W Q_{2n}^u W^{-1} \rho W \left(Q_{2n}^u\right)^\dagger W^{-1} \otimes |v\rangle \langle v| \\ &= \sum_{u,v} p_{UV}(u,v) Q_{2n}^{\Psi_n(W)u} \rho \left(Q_{2n}^{\Psi_n(W)u}\right)^\dagger \otimes |v\rangle \langle v| \end{aligned} \tag{17}$$

2.6 Preparation and measurement

Consider now the n-qubit Pauli channel specified by (17). Suppose the n qubits are divided in two registers: the first m qubits, and the remaining $k = n - m$ qubits. Suppose further that the first m qubits are prepared in the zero state, then the channel is applied, then the first m qubits are measured in the computational basis. These operations transform the n-qubit Pauli channel into a k-qubit Pauli channel acting on the last k qubits. The channel on the last k qubits transforms k-qubit states as follows:

$$\rho \mapsto \sum_{u,v} p_{UV}(u,v) Q_{2k}^{\sum_{i=1}^{2k} e_{2k,i} e_{2n,m+i}^{\top} \Psi_n(W) u} \rho \left(Q_{2k}^{\sum_{i=1}^{2k} e_{2k,i} e_{2n,m+i}^{\top} \Psi_n(W) u} \right)^{\dagger} \otimes \left| v, \sum_{i=1}^{m} e_{m,i} e_{2n,i}^{\top} \Psi_n(W) u \right\rangle \left\langle v, \sum_{i=1}^{m} e_{m,i} e_{2n,i}^{\top} \Psi_n(W) u \right| \tag{18}$$

2.7 Conditional correction

Finally, a Pauli correction is applied on the remaining qubits to transform (18) into a channel that is close to the identity. To simplify notation, rewrite (18) as

$$\rho \mapsto \sum_{u',v'} p'_{U'V'}(u',v') Q_{2k}^{u'} \rho \left(Q_{2k}^{u'} \right)^{\dagger} \otimes |v'\rangle \langle v'| \tag{19}$$

A conditional correction specified by some conditional probabilities $\hat{p}_{\hat{U}'|V'}$, followed by tracing out the side information result in the channel

$$\rho \mapsto \sum_{u',v',\hat{u}'} p'_{U'V'}(u',v') \hat{p}_{\hat{U}'|V'}(\hat{u}'|v') Q_{2k}^{u'+\hat{u}'} \rho \left(Q_{2k}^{u'+\hat{u}'} \right)^{\dagger} \tag{20}$$

2.8 Diamond distance from the identity

The diamond distance [22, Section 9.1.6] between the channel in equation (20) and the identity channel is $2\mathbb{P}\left(\hat{U}' \neq U'\right)$. Indeed, the triangle inequality shows that the diamond distance is at most $2\mathbb{P}\left(\hat{U}' \neq U'\right)$, and using a maximally entangled state in the maximization problem for the diamond distance [22, Theorem 9.1.1] shows that it is at least $2\mathbb{P}\left(\hat{U}' \neq U'\right)$.

2.9 Optimal stabilizer codes and coset decoding

Given a finite blocklength n and an n-qubit Pauli channel $\mathcal{N}_p$, it is desirable to understand the highest rate of a stabilizer code that can correct the noise $\mathcal{N}_p$ while keeping the probability of error below a given ϵ, and the lowest probability of error of a stabilizer code of a given rate r. The preceeding discussion shows that these can be defined combinatorially as follows:

Definition 2. *Let p_{UV} be a joint probability distribution of a random vector U in $\mathbb{F}_2^{2n}$ and another discrete random variable V.*

For $C \in \mathcal{SP}_{2n}$, partition $n = m + k$ of the qubits and decoding algorithm D, the probability of error is

$$\mathbb{P}\left(D\left(V, \sum_{i=1}^{m} e_{m,i} e_{2n,i}^{\top} CU \right) \neq \sum_{i=1}^{2k} e_{2k,i} e_{2n,m+i}^{\top} CU \right) \tag{21}$$

For $\epsilon > 0$, let $R_{coset}(p_{UV}, \epsilon)$ be the maximum value of k/n such that there exist C, D with probability of error at most ϵ.

For $r = k/n$, let $\varepsilon_{coset}(p_{UV}, r)$ be the lowest error probability of a pair C, D that uses $m = n - k$ bits of syndrome.

Note that $\sum_{i=1}^{m} e_{m,i} e_{2n,2n+1-i}^{\top} CU$ is irrelevant to the distance from the identity channel in section 2.7, and therefore it is also not mentioned in the combinatorial formulation in Definition 2. Thus, the Pauli error U, about which m bits are observed as the syndrome and $2k$ bits are guessed by the decoder, is determined only up to a coset of a vector space of dimension m.

2.10 Guessing the error from the syndrome

Analyzing the optimal rate and error probability for coset decoding presents considerable difficulties. Therefore, it is worthwhile to introduce a relaxation of this problem, and to consider the optimal rate and error probability for the related task of guessing the error from the syndrome.

Definition 3. *Consider a probability distribution p_{UV} as before.*

For $C \in \mathcal{SP}_{2n}$, partition $n = m + k$ of the qubits and decoding algorithm D, the probability of incorrect guess of U from the syndrome is

$$\mathbb{P}\left(D\left(V, \sum_{i=1}^{m} e_{m,i} e_{2n,i}^{\top} CU\right) \neq U\right) \tag{22}$$

For $\epsilon > 0$, let $R_{errorguess}(p_{UV}, \epsilon)$ be the maximum value of k/n such that there is a pair C, D with the probability of guessing U incorrectly at most ϵ.

For $r = k/n$, let $\varepsilon_{errorguess}(p_{UV}, r)$ be the lowest probability of guessing U incorrectly of a pair C, D that uses $m = n - k$ bits of syndrome.

Since a pair C, D that can guess U can be adapted to coset decoding with equal or lower error probability, the optimal rates and error probabilities satisfy

$$R_{errorguess}(p_{UV}, \epsilon) \leq R_{coset}(p_{UV}, \epsilon) \tag{23}$$

$$\varepsilon_{coset}(p_{UV}, r) \leq \varepsilon_{errorguess}(p_{UV}, r) \tag{24}$$

3 Groups of lower triangular matrices

Subsection 3.1 introduces a family of groups of lower triangular matrices. Subsection 3.2 shows that each of these groups is generated by a suitable subset of the Gaussian move matrices. Subsection 3.3 considers the intersection of these groups with the symplectic group and Subsection 3.4 shows that these intersections are generated by suitable symplectic analogues of the Gaussian moves.

3.1 A family of groups from transitive sets of pairs

Definition 4. *For $n \in \mathbb{N}$, let*

$$Pairs\,(n) = \{(i,j) : i,j \in [n], i > j\} \tag{25}$$

Definition 5. *A subset T of $Pairs\,(n)$ is called transitive if*

$$(i,j) \in T, (j,k) \in T \implies (i,k) \in T \tag{26}$$

Definition 6. *To a transitive subset T, associate the set of lower triangular matrices*

$$\mathcal{L}_n\,(T) = I_n + span\left\{e_{n,i}e_{n,j}^\top : (i,j) \in T\right\} \tag{27}$$

Lemma 1. *$\mathcal{L}_n\,(T)$ is closed under matrix multiplication and matrix inverse.*

Proof. Take any

$$A = \sum_{(i,j)\in T} a_{i,j}e_{n,i}e_{n,j}^\top, \quad B = \sum_{(k,l)\in T} b_{k,l}e_{n,k}e_{n,l}^\top \tag{28}$$

From the transitive property, deduce $AB \in span\left\{e_{n,i}e_{n,j}^\top : (i,j) \in T\right\}$. Then,

$$(I_n + A)(I_n + B) = I_n + A + B + AB \in \mathcal{L}_n\,(T) \tag{29}$$

Moreover, A is nilpotent. Let d be the largest integer such that $A^{2^d} \neq 0$. Then,

$$(I_n - A)^{-1} = (I + A)(I + A^2)(I + A^4)\dots(I + A^{2^d}) \tag{30}$$

is also an element of $\mathcal{L}_n\,(T)$. □

Some examples

The lower triangular group is $\mathcal{L}_n\,(Pairs\,(n))$.

Matrices with off-diagonal entries only in certain rows and columns: take subsets R, C of $[n]$. Then, $T = Pairs\,(n) \cap (R \times C)$ is transitive. $\mathcal{L}_n\,(T)$ is a group of lower triangular matrices whose non-zero off-diagonal entries appear only in rows in R and columns in C.

The inversions and non-inversions of a permutation: let π be a permutation of $[n]$. Let

$$inv(\pi) = \{(i,j) : i > j, \pi(i) < \pi(j)\} \tag{31}$$
$$ninv(\pi) = \{(i,j) : i > j, \pi(i) > \pi(j)\} \tag{32}$$

Both $inv(\pi)$ and $ninv(\pi)$ are transitive. The associated groups $\mathcal{L}_n\,(inv(\pi))$ and $\mathcal{L}_n\,(ninv(\pi))$ played a role in the canonical form for invertible $n \times n$ matrices [5].

3.2 Generators and canonical form of the groups $\mathcal{L}_n(T)$

For $i \neq j \in [n]$, let

$$G_{n,i,j} = I_n + e_{n,i}e_{n,j}^\top \in \mathbb{F}_2^{n\times n} \tag{33}$$

denote the corresponding Gaussian move matrix over $\mathbb{F}_2$. Left multiplication by $G_{n,i,j}$ adds row j to row i. Right multiplication by $G_{n,i,j}$ adds column i to column j. $G_{n,i,j}G_{n,j,i}G_{n,i,j}$ acts on the left as a row swap and on the right as a column swap.

Consider the identity:

Lemma 2. *Let $\{a_{i,j} : (i,j) \in Pairs(n)\}$ be any collection of elements of $\mathbb{F}_2$ indexed by $Pairs(n)$. Then,*

$$I_n + \sum_{(i,j)\in Pairs(n)} a_{i,j}e_{n,i}e_{n,j}^\top = \prod_{(i,j)\in Pairs(n)} G_{n,i,j}^{a_{i,j}} \tag{34}$$

where in the product, the order of terms from left to right is

$$(2,1),\dots,(n,1),(3,2),\dots,(n,2),\dots,(n,n-1) \tag{35}$$

Proof. Think of the product as a polynomial in the standard basis vectors and their transposes. The ordering is such that all terms of degree 4 and higher vanish, and only the terms of degree 2 and the identity remain. □

This identity gives a set of generators and a canonical form for all the groups $\mathcal{L}_n(T)$.

Corollary 1. *Take $n \in \mathbb{N}$ and transitive $T \subset Pairs(n)$. The group $\mathcal{L}_n(T)$ is generated by*

$$\{G_{n,i,j} : (i,j) \in T\} \tag{36}$$

Each element of $\mathcal{L}_n(T)$ can be uniquely written as a product

$$\prod_{(i,j)\in T} G_{n,i,j}^{a_{i,j}} \tag{37}$$

where the product respects the order (35)

3.3 Intersection of the groups $\mathcal{L}_{2n}(T)$ with the symplectic group

Definition 7. *A subset T of $Pairs(2n)$ will be called closed under reversal if*

$$(i,j) \in T \iff (2n+1-j, 2n+1-i) \in T \tag{38}$$

Definition 8. *To a subset T that is both transitive and closed under reversal, associate the group*

$$\mathcal{B}_{2n}(T) = \mathcal{L}_{2n}(T) \cap \mathcal{SP}_{2n} \tag{39}$$

3.4 Generators and canonical form of the groups $\mathcal{B}_{2n}(T)$

3.4.1 The Clifford Gaussian moves

For $i \neq j \in \{1, \ldots, 2n\}$ let

$$W_{2n,i,j} = \begin{cases} \sqrt{Q_{2n,i}} & \text{if } i+j = 2n+1 \\ \frac{1}{2}\left(I + Q_{2n,2n+1-j} + Q_{2n,i} - Q_{2n,2n+1-j}Q_{2n,i}\right) & \text{otherwise} \end{cases} \tag{40}$$

These unitaries will be called the Clifford Gaussian moves, because they play a role analogous to the classical Gaussian moves $G_{n,i,j}$.

The Clifford gaussian moves are also related to the standard Clifford one and two qubit gates:

1. For $i \in [n]$, $W_{2n,i,2n+1-i}$ is a phase gate conjugated by a Hadamard gate on qubit i.
2. For $i \in [n]$, $W_{2n,2n+1-i,i}$ is a phase gate on qubit i.
3. For $i \in [n]$, $W_{2n,2n+1-i,i}W_{2n,i,2n+1-i}W_{2n,2n+1-i,i} = \sqrt{Z_{n,i}}\sqrt{X_{n,i}}\sqrt{Z_{n,i}} = \frac{1+i}{\sqrt{2}}H_{n,i}$ is a Hadamard gate on qubit i.
4. For $i \neq j \in [n]$, $W_{2n,i,j} = W_{2n,(2n+1-j),(2n+1-i)}$ is a CNOT gate with control qubit j and target qubit i.
5. For $i \neq j \in [n]$, $W_{2n,i,j}W_{2n,j,i}W_{2n,i,j}$ is a SWAP of qubits i, j.
6. For $i > n \geq j, i+j \neq 2n+1$, $W_{2n,i,j} = W_{2n,2n+1-j,2n+1-i}$ is a CZ gate on qubits $2n+1-i, j$.
7. For $i \leq n < j$, $W_{2n,i,j} = W_{2n,2n+1-j,2n+1-i}$ is a CZ gate conjugated by Hadamard gates on qubits $i, 2n+1-j$.

3.4.2 The symplectic Gaussian moves

The symplectic Gaussian moves are the images of the Clifford Gaussian moves under the homomorphism Ψ_n: for $i \neq j \in \{1, \ldots, 2n\}$ let

$$\begin{aligned} S_{2n,i,j} &= \Psi_n(W_{2n,i,j}) \\ &= \begin{cases} I_{2n} + e_{2n,i}e_{2n,j}^\top & \text{if } i+j = 2n+1 \\ I_{2n} + e_{2n,i}e_{2n,j}^\top + e_{2n,2n+1-j}e_{2n,2n+1-i}^\top & \text{otherwise} \end{cases} \end{aligned} \tag{41}$$

3.4.3 Clearing entire rows and columns

It is well-known that there are simple combinations of classical Gaussian moves that can be used to clear entire rows and columns. It is less obvious that there are combinations of the more complicated Clifford and symplectic Gaussian moves that have similar properties. To illustrate the similarities and differences of the classical and symplectic case, both are now given explicitly.

With classical Gaussian moves: Take $i \in [n]$. Take a vector

$$v = \sum_{j=1}^{n} v_j e_{n,j} \in \mathbb{F}_2^n \tag{42}$$

that does not have an i-th component (i.e. $v_i = 0$). Let

$$G_{n,v,i} = I_n + v e_{n,i}^\top = \prod_{j \neq i} G_{n,j,i}^{v_j} \tag{43}$$

and let $G_{n,i,v^\top} = G_{n,v,i}^\top$.

With symplectic Gaussian moves: Take $i \in [2n]$. Take a vector

$$v = \sum_{j=1}^{2n} v_j e_{2n,j} \in \mathbb{F}_2^{2n} \tag{44}$$

that does not have an i-th component (i.e. $v_i = 0$). Let

$$\begin{aligned} S_{2n,v,i} = I_{2n} + v e_{2n,i}^\top + \Omega_{2n} e_{2n,i} v^\top \Omega_{2n} + v_{2n+1-i} e_{2n,2n+1-i} e_{2n,i}^\top \\ = S_{2n,2n+1-i,i}^{\sum_{j=1}^{n} v_j v_{2n+1-j}} \prod_{j \neq i} S_{2n,j,i}^{v_j} \end{aligned} \tag{45}$$

and let $S_{2n,i,v^\top} = S_{2n,v,i}^\top$. Similar combinations of Clifford gates appeared previously in [5, Lemma 4].

Some relevant properties of $S_{2n,v,i}$ are:

Lemma 3. *1. For fixed i the symplectic Gaussian moves $\{S_{2n,j,i} : j \in [2n] \backslash \{i\}\}$ commute.*

2. *The two different expressions for $S_{2n,v,i}$ in equation (45), one with the identity and outer products, the other with the symplectic Gaussian moves, are equal.*

3. *$S_{2n,v,i}^2 = I_{2n}$.*

4. *$S_{2n,v,i}$ is supported only on the main diagonal, column i and row $2n+1-i$.*

5. *$S_{2n,v,i} e_{2n,i} = e_{2n,i} + v$*

6. *If $e_{2n,i}^\top u = v^\top \Omega_{2n} u = 0$, then $S_{2n,v,i} u = u$.*

Proof. From the definitions using simple calculations. The most work is required to show that the two expressions in equation (45) are equal. To do this, consider the action of the two expressions on the standard basis vectors, exploiting the commutativity from part 1 and the fact that $S_{2n,i,j}$ acts as the identity on all standard basis vectors except $e_{2n,j}, e_{2n,2n+1-i}$. $\square$

3.4.4 An identity reveals the generators and canonical form of the groups $\mathcal{B}_{2n}(T)$

In the classical case, the generators and canonical form of the groups $\mathcal{L}_n(T)$ were obtained from the identity in Lemma 2. A symplectic analogue of that identity will now be given.

Because of the symplectic constraint, the off-diagonal entries of matrices in $\mathcal{L}_{2n}(Pairs(2n)) \cap \mathcal{SP}_{2n}$ cannot all be chosen freely. The following linear function focuses on a set of entries that can be chosen freely, and that together determine the rest.

Definition 9. *For $n \in \mathbb{N}$, let $\mathcal{V}_n : \mathbb{F}_2^{2n\times 2n} \to \mathbb{F}_2^{2n\times n}$ be the linear function*

$$\mathcal{V}_n(B) = \sum_{j=1}^{n} \sum_{i=j+1}^{2n+1-j} e_{2n,i} e_{2n,i}^\top B e_{2n,j} e_{n,j}^\top \tag{46}$$

Now, consider the identity:

Theorem 1. *For every $B \in \mathcal{B}_{2n}(Pairs(2n))$,*

$$B = S_{2n,\mathcal{V}_n(B)e_{n,1},1} S_{2n,\mathcal{V}_n(B)e_{n,2},2} \cdots S_{2n,\mathcal{V}_n(B)e_{n,n},n} \tag{47}$$

Proof. Define a sequence of matrices by

$$B_0 = B, B_k = S_{2n,\mathcal{V}_n(B)e_{n,k},k} B_{k-1}, k = 1, \dots, n \tag{48}$$

It will be shown by induction on k that the matrices B_k satisfy:

$$e_{2n,i}^\top B_k e_{2n,j} = \begin{cases} e_{2n,i}^\top B e_{2n,j} & \text{if } i,j \in \{k+1, \dots, 2n-k\} \\ e_{2n,i}^\top e_{2n,j} & \text{otherwise} \end{cases} \tag{49}$$

For $B_0 = B$, the claim holds. Suppose the claim holds for B_{k-1}. Note that

$$\begin{aligned} B_k &= S_{2n,\mathcal{V}_n(B)e_{n,k},k} B_{k-1} \\ &= B_{k-1} + \mathcal{V}_n(B) e_{n,k} e_{2n,k}^\top B_{k-1} + e_{2n,2n+1-k} e_{n,k}^\top \mathcal{V}_n(B)^\top \Omega_{2n} B_{k-1} \\ &\qquad + e_{2n,2n+1-k} e_{2n,2n+1-k}^\top B e_{2n,k} e_{2n,k}^\top B_{k-1} \end{aligned} \tag{50}$$

Since $e_{2n,k}^\top B_{k-1} = e_{2n,k}^\top$, the second term is supported only on the k-th column. The third and fourth terms are supported only on row $2n+1-k$. Thus, $B_k - B_{k-1}$ is supported only on column k and row $2n+1-k$. Moreover, column k of B_{k-1} is $e_{2n,k} + \mathcal{V}_n(B)e_{n,k}$, so column k of B_k is just $e_{2n,k}$. Finally, row $2n+1-k$ of B_k is determined by column k via the symplectic constraint:

$$e_{2n,2n+1-k}^\top B_k = e_{2n,k}^\top \Omega_{2n} B_k = e_{2n,k}^\top B_k^\top \Omega_{2n} B_k = e_{2n,k}^\top \Omega_{2n} = e_{2n,2n+1-k}^\top \tag{51}$$

This completes the induction. Then, $B_n = I_{2n}$, which proves the theorem. $\square$

The following Theorem gives properties of the groups $\mathcal{B}_{2n}(T)$ related to generators and a canonical form:

Theorem 2. *Let $T \subset Pairs\,(2n)$ be both transitive and closed under reversal. Then,*

1. $(i, j) \in T$ *implies* $S_{2n,i,j} \in \mathcal{B}_{2n}\,(T)$.
2. *If* $i \in [2n]$ *and* $v \in \mathbb{F}_2^{2n}$ *are such that* v *is supported on positions* $\{j : (j, i) \in T\}$, *then* $S_{2n,v,i} \in \mathcal{B}_{2n}\,(T)$.
3. $\{S_{2n,i,j} : j \in [n], j+1 \leq i \leq 2n+1-j, (i, j) \in T\}$ *generates* $\mathcal{B}_{2n}\,(T)$.
4. *The identity in Theorem 1 provides a canonical expression of each* $B \in \mathcal{B}_{2n}\,(T)$ *in terms of the generators.*

Proof. Part 1 follows directly from the definitions.

Part 2 follows from part 1, equation (45) and the following claim, which is used to deal with the correction term $S_{2n,2n+1-i,i}^{\sum_{j=1}^{n} v_j v_{2n+1-j}}$ in equation (45):

$$\text{Claim: } (j, i) \in T \wedge (2n+1-j, i) \in T \implies (2n+1-i, i) \in T \tag{52}$$

Indeed, closure under reversal implies $(2n+1-i, 2n+1-j) \in T$ and then transitivity implies $(2n+1-i, i) \in T$.

Parts 3 and 4 follow from part 2 and Theorem 1. □

4 Canonical forms for unrestricted and stabilizer parity check matrices, and for the symplectic and Clifford groups

The preparatory subsection 4.1 introduces the main ideas in the simpler case of unrestricted matrices. It also establishes notation and some lemmas. Next, subsection 4.2 gives the canonical form for stabilizer parity check matrices. Finally, subsection 4.3 considers the canonical form for the symplectic and Clifford groups.

4.1 Gaussian elimination produces a canonical form for matrices of a given size

An application of the groups $\mathcal{L}_n\,(T)$ will now be described. Groups of this type will be seen to arise naturally during Gaussian elimination. They will be used to show that the associated matrix decomposition is in fact a canonical form.

One of the many variants of Gaussian elimination will now be recalled in some detail, with special attention on which row and column operations can potentially be used, as a function of the pivot positions.

The search for pivots must proceed in some definite order; different orders result in different but analogous matrix decompositions [20]. The present article uses the order "left and down": the search for pivots starts in the first row from the last element to the first, then in the second row from the last element to the first, etc. This choice produces decompositions with lower triangular matrices both on the left and on the right.

The positions of the pivots can be described using a pair of functions. Let $\alpha : [r] \to [m]$ be increasing, and let $\beta : [r] \to [n]$ be injective; to these functions correspond the r pivot positions $(\alpha(i), \beta(i)), i = 1, \dots, r$. These r pivot positions can also be summarized in a matrix: let

$$\Pi(\alpha, \beta) = \sum_{i=1}^{r} e_{m,\alpha(i)} e_{n,\beta(i)}^{\top} \tag{53}$$

$\Pi(\alpha, \beta)$ may be called an incomplete permutation matrix, because it has at most one 1 in each row and column, but may be non-square and may have zero rows and columns.

During Gaussian elimination, the pivot positions are not revealed all at once, but step-by-step. Take $s > r$, and suppose that increasing $\alpha' : [s] \to [m]$ and injective $\beta' : [s] \to [n]$ agree with α, β on $[r]$. Then, call (α', β') an s-extension of (α, β), and call (α, β) the r-predecessor of (α', β'). A given collection of pivot positions can have many possible extensions to a given higher rank, but it has a unique predecessor at a given lower rank.

The algorithm progresses by taking a partially reduced matrix and simplifying it further. For α, β as above, call $A \in \mathbb{F}_2^{m \times n}$ (α, β)-partially reduced if rows $1, \dots, \alpha(r)$ and columns $\beta(1), \dots, \beta(r)$ of A coincide with the corresponding rows and columns of $\Pi(\alpha, \beta)$. If in addition $A = \Pi(\alpha, \beta)$, then A is fully reduced, and the algorithm ends.

If A is partially but not fully reduced, then extend (α, β) to (α', β') so that $\alpha'(r+1), \beta'(r+1)$ is the position of the next pivot in the left and down order. Then, transform A:

$$A' = \left(\prod_{i=\alpha'(r+1)+1}^{m} G_{m,i,\alpha'(r+1)}^{e_{m,i}^{\top} A e_{n,\beta'(r+1)}} \right) A \left(\prod_{j=1}^{\beta'(r+1)-1} G_{n,\beta'(r+1),j}^{e_{m,\alpha'(r+1)}^{\top} A e_{n,j}} \right) \tag{54}$$

to a matrix A' that is (α', β')-partially simplified.

Now, observe that on the left, only $G_{m,i,\alpha'(r+1)}$ with $i > \alpha'(r+1)$ are used. Similarly, on the right, only $G_{n,\beta'(r+1),j}$ with $j < \beta'(r+1)$ are used. However, on the right, there is an additional restriction: since columns $\beta(1), \dots, \beta(r)$ of A are already simplified, $G_{n,\beta'(r+1),j}$ with $j \in Im(\beta)$ are not used.[1]

These observations prompt the following definitions.

Definition 10. *For increasing $\alpha : [r] \to [m]$, let*

$$T^{Left}(\alpha) = \{(i, j) : j \in Im(\alpha), i > j\} \subset Pairs\,(m) \tag{55}$$

[1]If the search for pivots proceeded along columns rather than along rows, i.e. in the order last column top to bottom, then second last column top to bottom, etc., the additional restriction would be on the left rather than on the right.

Definition 11. *For injective $\beta : [r] \to [n]$, let*

$$T^{Right}(\beta) = \left\{(i,j) : i \in Im(\beta), j < i, j \notin \left\{\beta(1), \dots, \beta\left(\beta^{-1}(i) - 1\right)\right\}\right\} \subset Pairs\,(n) \quad (56)$$

Note that these sets are transitive and behave naturally under extensions: for α' an extension of α, $T^{Left}(\alpha) \subset T^{Left}(\alpha')$, and for β' and extension of β, $T^{Right}(\beta) \subset T^{Right}(\beta')$.

Putting everything so far together gives the following matrix decomposition:

Theorem 3. *For all $A \in \mathbb{F}_2^{m\times n}$ there exist $r \leq \min(m,n)$, increasing $\alpha : [r] \to [m]$, injective $\beta : [r] \to [n]$, and matrices $L \in \mathcal{L}_m\left(T^{Left}(\alpha)\right)$, $R \in \mathcal{L}_n\left(T^{Right}(\beta)\right)$ such that*

$$A = L\Pi(\alpha,\beta)R \quad (57)$$

(r,α,β,L,R) can be computed from A in time $O(mnr)$ by Gaussian elimination.[2]

The given matrix decomposition is in fact unique; thus, it is a canonical form for $m \times n$ matrices. This will now be proved in detail. For convenience, say that a quintuple (r,α,β,L,R) is (m,n)-allowed if $r \leq \min(m,n)$, $\alpha : [r] \to [m]$ is increasing, $\beta : [r] \to [n]$ is injective, $L \in \mathcal{L}_m\left(T^{Left}(\alpha)\right)$ and $R \in \mathcal{L}_n\left(T^{Right}(\beta)\right)$.

Theorem 4. *Let (r,α,β,L,R), $(r',\alpha',\beta',L',R')$ be (m,n)-allowed quintuples such that*

$$L\Pi(\alpha,\beta)R = L'\Pi(\alpha',\beta')R' \quad (58)$$

Then,

$$(r,\alpha,\beta,L,R) = (r',\alpha',\beta',L',R') \quad (59)$$

Proof. The first step is to establish uniqueness of the incomplete permutation matrix. In [5, Lemma 14], an argument for the case of full permutation matrices is given; unfortunately, it is not clear how to extend this approach to the case of incomplete permutation matrices. On the other hand, a rank-of-minors argument is given in [20], also for the case of full permutation matrices, and this proof extends without problem to the case of incomplete permutation matrices.

Lemma 4. *Let $\Pi(\alpha,\beta)$, $\Pi(\alpha',\beta')$ be two $m\times n$ incomplete permutation matrices. If there exist lower triangular $L \in \mathcal{L}_m(Pairs\,(m))$, $R \in \mathcal{L}_n(Pairs\,(n))$ such that*

$$L\Pi(\alpha,\beta) = \Pi(\alpha',\beta')R \quad (60)$$

then $\Pi(\alpha,\beta) = \Pi(\alpha',\beta')$.

[2] Assuming L, R are output as lists of off-diagonal non-zero positions; otherwise, outputting the full L, R would take time $\max(m,n)^2$.

Proof. Take any $k \in [m], l \in [n]$, and divide the matrices into blocks with $(k, m-k)$ and $(l, n-l)$ rows/columns:

$$\begin{pmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} \Pi(\alpha,\beta)_{11} & \Pi(\alpha,\beta)_{12} \\ \Pi(\alpha,\beta)_{21} & \Pi(\alpha,\beta)_{22} \end{pmatrix} = \begin{pmatrix} \Pi(\alpha',\beta')_{11} & \Pi(\alpha',\beta')_{12} \\ \Pi(\alpha',\beta')_{21} & \Pi(\alpha',\beta')_{22} \end{pmatrix} \begin{pmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{pmatrix} \tag{61}$$

Then, the upper right blocks $\Pi(\alpha,\beta)_{12}, \Pi(\alpha',\beta')_{12}$ have the same rank. This holds for any k, l, so $\Pi(\alpha,\beta) = \Pi(\alpha',\beta')$. $\square$

The next step is to establish uniqueness of the right factor. For the case of full permutation matrices, both [20, 5] argue by considering whether conjugating a lower triangular matrix by a permutation leads to another lower triangular matrix. Unfortunately, this approach breaks down for incomplete permutation matrices, and a different approach is needed.

Lemma 5. *Let $\beta : [r] \to [n]$ be injective, and let R be in $\mathcal{L}_n\left(T^{Right}(\beta)\right)$. If there exist m, lower triangular $L \in \mathcal{L}_m(Pairs(m))$, and increasing $\alpha : [r] \to [m]$ such that $L\Pi(\alpha,\beta) = \Pi(\alpha,\beta)R$, then $R = I$.*

Proof. Take any $i, j \in [n]$. The goal is to show $e_{n,i}^\top R e_{n,j} = e_{n,i}^\top e_{n,j}$. Consider cases.

Case 1: $i \le j$**.** Then, $e_{n,i}^\top R e_{n,j} = e_{n,i}^\top e_{n,j}$ because R is lower triangular.

Case 2: $i > j$**,** $i \notin Im(\beta)$**.** Then, $e_{n,i}^\top R e_{n,j} = 0$ because $R \in \mathcal{L}_n\left(T^{Right}(\beta)\right)$.

Case 3: $i > j$**,** $\exists k : i = \beta(k)$**,** $j \notin Im(\beta)$**.** Then,

$$e_{n,i}^\top R e_{n,j} = e_{m,\alpha(k)}^\top \Pi(\alpha,\beta) R e_{n,j} = e_{m,\alpha(k)}^\top L\Pi(\alpha,\beta) e_{n,j} = 0 \tag{62}$$

Case 4: $i > j$**,** $\exists k, l : i = \beta(k), j = \beta(l), l < k$**.** Then, $e_{n,i}^\top R e_{n,j} = 0$ because $R \in \mathcal{L}_n\left(T^{Right}(\beta)\right)$.

Case 5: $i > j$**,** $\exists k, l : i = \beta(k), j = \beta(l), k < l$**.** Then,

$$e_{n,i}^\top R e_{n,j} = e_{m,\alpha(k)}^\top \Pi(\alpha,\beta) R e_{n,j} = e_{m,\alpha(k)}^\top L\Pi(\alpha,\beta) e_{n,j} = e_{m,\alpha(k)}^\top L e_{m,\alpha(l)} = 0 \tag{63}$$

because α is increasing and L is lower triangular. $\square$

Finally, note that since incomplete permutation matrices are not necessarily invertible, uniqueness of the left factor does not follow automatically from uniqueness of the incomplete permutation matrix and of the right factor. Fortunately, the following lemma holds:

Lemma 6. *Let $\alpha : [r] \to [m]$ be increasing, and let L be in $\mathcal{L}_m\left(T^{Right}(\alpha)\right)$. If there exist n and injective $\beta : [r] \to [n]$ such that $L\Pi(\alpha,\beta) = \Pi(\alpha,\beta)$, then $L = I$.*

Proof. Take any $i \in [m]$. If $i \notin Im(\alpha)$, then $Le_{m,i} = e_{m,i}$ by the definition of $\mathcal{L}_m\left(T^{Left}(\alpha)\right)$. If $i \in Im(\alpha)$, then

$$Le_{m,i} = L\Pi(\alpha,\beta)e_{n,\beta(\alpha^{-1}(i))} = \Pi(\alpha,\beta)e_{n,\beta(\alpha^{-1}(i))} = e_{m,i} \tag{64}$$

Thus, $Le_{m,i} = e_{m,i}$ holds for all $i \in [m]$. □

Lemmas 4, 5, 6 imply Theorem 4. □

4.2 Canonical form for $m \times 2n$ stabilizer parity check matrices

Now, an application of the groups $\mathcal{B}_{2n}(T)$ will be described. Recall from section 4.1 that Gaussian elimination produces a canonical form for unrestricted matrices in $\mathbb{F}_2^{m\times n}$. Now, consider $m \times 2n$ stabilizer parity check matrices. This section shows that Gaussian elimination modified to use symplectic Gaussian moves on the right produces a canonical form for such matrices, and the right factor belongs to a group $\mathcal{B}_{2n}(T)$ for T a suitable function of the pivot positions.

First, some definitions:

Definition 12. *$A \in \mathbb{F}_2^{m\times 2n}$ is a stabilizer parity check matrix if $A\Omega_{2n}A^\top = 0$.*

When Gaussian elimination is performed on a stabilizer parity check matrix with symplectic moves on the right, the column indices of the pivots have special structure:

Definition 13. *Let $q_n : [2n] \to [n]$ be the function*

$$q_n(i) = \min(i, 2n+1-i) \tag{65}$$

that maps each $i \in [2n]$ to the qubit on which Pauli $Q_{2n,i}$ acts.

Definition 14. *A function $\beta : [r] \to [2n]$ will be called qubit-injective if $q_n \circ \beta : [r] \to [n]$ is injective.*

As in the classical case, there is a set of off-diagonal positions that are associated to a collection of column indices:

Definition 15. *To qubit-injective $\beta : [r] \to [2n]$ associate the sets*

$$T^{moves}(\beta) =$$
$$\left\{(i,j) : i \in Im(\beta), j < i, q_n(j) \notin \left\{q_n(\beta(1)), \ldots, q_n(\beta(\beta^{-1}(i)-1))\right\}\right\} \tag{66}$$
$$T^{rev}(\beta) = \{(i,j) : (2n+1-j, 2n+1-i) \in T^{moves}(\beta)\} \tag{67}$$
$$T^{tcr}(\beta) = T^{moves}(\beta) \cup T^{rev}(\beta) \tag{68}$$

$T^{moves}(\beta)$ is the set of index pairs (i,j) such that $S_{2n,i,j}$ can potentially be used during the elimination algorithm described below when the column indices of the pivots are given by β. $T^{rev}(\beta)$ is the elementwise reversal of $T^{moves}(\beta)$; since $S_{2n,i,j} = S_{2n,2n+1-j,2n+1-i}$, it is natural to consider also this set. $T^{tcr}(\beta)$ is the transitive, closed under reversal set that defines the associated group $\mathcal{B}_{2n}(T^{tcr}(\beta))$. Relevant properties of $T^{tcr}(\beta)$ are shown below:

Lemma 7. *If $\beta : [r] \to [2n]$ is qubit-injective, then $T^{tcr}(\beta)$ is transitive and closed under reversal.*

If β' is qubit-injective and extends β, then $T^{tcr}(\beta) \subset T^{tcr}(\beta')$.

Proof. $T^{tcr}(\beta)$ is closed under reversal by construction.

Now, take any $(i,j),(j,k) \in T^{tcr}(\beta)$ and consider cases:

Case 1: $(i,j),(j,k) \in T^{moves}(\beta)$. Then, $i,j \in Im(\beta)$ with $\beta^{-1}(j) > \beta^{-1}(i)$. Moreover, $i > j > k$ and

$$q_n(k) \notin \{q_n(\beta(1)), \ldots, q_n(\beta(\beta^{-1}(j)-1))\} \supset \{q_n(\beta(1)), \ldots, q_n(\beta(\beta^{-1}(i)-1))\} \tag{69}$$

Then, $(i,k) \in T^{moves}(\beta)$.

Case 2: $(i,j),(j,k) \in T^{rev}(\beta)$. Then, $(2n+1-k, 2n+1-j), (2n+1-j, 2n+1-i) \in T^{moves}(\beta)$, so $(2n+1-k, 2n+1-i) \in T^{moves}(\beta)$, so $(i,k) \in T^{rev}(\beta)$.

Case 3: $(i,j) \in T^{moves}(\beta), (j,k) \in T^{rev}(\beta)$. Then $i > j > k$ and $i, 2n+1-k \in Im(\beta)$. Consider subcases:

1. If $\beta^{-1}(2n+1-k) \geq \beta^{-1}(i)$, then $(i,k) \in T^{moves}(\beta)$.
2. If $\beta^{-1}(2n+1-k) < \beta^{-1}(i)$, then $(2n+1-k, 2n+1-i) \in T^{moves}(\beta)$, so $(i,k) \in T^{rev}(\beta)$.

Case 4: $(i,j) \in T^{rev}(\beta), (j,k) \in T^{moves}(\beta)$. Then, $j, 2n+1-j \in Im(\beta)$. This is a contradiction, because β is qubit injective. This case does not occur.

Combining the cases implies that $T^{tcr}(\beta)$ is transitive.

Finally, if β' extends β, then $T^{moves}(\beta') \supset T^{moves}(\beta)$, so $T^{tcr}(\beta') \supset T^{tcr}(\beta)$. $\square$

Now, consider Gaussian elimination using symplectic moves on the right. The typical step of this algorithm is:

Lemma 8. *Let $\alpha : [r] \to [m]$ be increasing and $\beta : [r] \to [2n]$ be qubit-injective. Let $A \in \mathbb{F}_2^{m \times 2n}$ be a stabilizer parity check matrix that is (α, β) partially reduced, but not fully reduced. Let (α', β') extend (α, β) so that $(\alpha'(r+1), \beta'(r+1))$ is the position of the next pivot in the left and down order. Let $u \in \mathbb{F}_2^m$, $v \in \mathbb{F}_2^{2n}$ be such that*

$$Ae_{2n,\beta'(r+1)} = e_{m,\alpha'(r+1)} + u \tag{70}$$

$$e_{m,\alpha'(r+1)}^\top A = e_{2n,\beta'(r+1)}^\top + v^\top \tag{71}$$

and let

$$A' = G_{m,u,\alpha'(r+1)} A S_{2n,\beta'(r+1),v^\top} \tag{72}$$

Then,

1. $G_{m,u,\alpha'(r+1)} \in \mathcal{L}_m\left(T^{Left}(\alpha')\right)$.

2. β' *is qubit injective.*

3. $S_{2n,\beta'(r+1),v^\top} \in \mathcal{B}_{2n}\left(T^{tcr}(\beta')\right)$.

4. A' *is a stabilizer parity check matrix and is* (α',β') *partially reduced.*

5. (α',β',u,v,A') *can be computed from* (A,α,β) *in time* $O(mn)$.

Proof. **Part 1:** Since A is (α,β) reduced, u is supported on a subset of $\{j : j > \alpha'(r+1)\}$.

Parts 2 and 3: Since A is (α,β) reduced, and since $A\Omega_{2n}A^\top = 0$, the following holds for the columns of A:

$$\forall i \in [r]: \quad Ae_{2n,\beta(i)} = e_{m,\alpha(i)} \quad \wedge \quad Ae_{2n,2n+1-\beta(i)} = 0 \tag{73}$$

Therefore, $q_n(\beta'(r+1)) \notin \{q_n(\beta(1)),\dots q_n(\beta(r)\}$, and v is supported on a subset of $\{i : i < \beta'(r+1), q_n(i) \notin \{q_n(\beta(1)),\dots,q_n(\beta(r))\}\}$.

Part 4: First,

$$\begin{aligned} A'\Omega_{2n}(A')^\top &= G_{m,u,\alpha'(r+1)}AS_{2n,\beta'(r+1),v^\top}\Omega_{2n}S^\top_{2n,\beta'(r+1),v^\top}A^\top G^\top_{m,u,\alpha'(r+1)} \\ &= G_{m,u,\alpha'(r+1)}A\Omega_{2n}A^\top G^\top_{m,u,\alpha'(r+1)} = 0 \end{aligned} \tag{74}$$

Next, check the rows: for $i \le \alpha'(r+1)$,

$$\begin{aligned} &e^\top_{m,i}A' = e^\top_{m,i}G_{m,u,\alpha'(r+1)}AS_{2n,\beta'(r+1),v^\top} = e^\top_{m,i}AS_{2n,\beta'(r+1),v^\top} \\ &= \begin{cases} 0 & \text{if } i \notin Im(\alpha') \\ e^\top_{m,\alpha(k)}AS_{2n,\beta'(r+1),v^\top} = e^\top_{2n,\beta(k)}S_{2n,\beta'(r+1),v^\top} = e^\top_{2n,\beta(k)} & \text{if } i = \alpha(k), k \in [r] \\ (e^\top_{2n,\beta'(r+1)} + v^\top)S_{2n,\beta'(r+1),v^\top} = e^\top_{2n,\beta'(r+1)} & \text{if } i = \alpha'(r+1) \end{cases} \end{aligned} \tag{75}$$

Next, check the columns: for $i \in [r+1]$

$$\begin{aligned} A'e_{2n,\beta'(i)} &= G_{m,u,\alpha'(r+1)}AS_{2n,\beta'(r+1),v^\top}e_{2n,\beta'(i)} = G_{m,u,\alpha'(r+1)}Ae_{2n,\beta'(i)} \\ &= \begin{cases} G_{m,u,\alpha'(r+1)}(e_{m,\alpha'(r+1)} + u) = e_{m,\alpha'(r+1)} & \text{if } i = r+1 \\ G_{m,u,\alpha'(r+1)}e_{m,\alpha(i)} = e_{m,\alpha(i)} & \text{if } i \le r \end{cases} \end{aligned} \tag{76}$$

Part 5: Computation of (α',β',u,v) can clearly be done in time $O(mn)$. For the computation of A', use the sparsity of $G_{m,u,\alpha'(r+1)}, S_{2n,\beta'(r+1),v^\top}$. For any $x \in \mathbb{F}_2^m$, $G_{m,u,\alpha'(r+1)}x$ can be computed in time $O(m)$; since A has $2n$ columns, this gives time $O(mn)$ to compute $G_{m,u,\alpha'(r+1)}A$. Similarly, for any $y \in \mathbb{F}_2^{2n}$, $y^\top S_{2n,\beta'(r+1),v^\top}$ can be computed in time $O(n)$; for m rows, this gives time $O(mn)$ to compute $G_{m,u,\alpha'(r+1)}AS_{2n,\beta'(r+1),v^\top}$. ∎

Repeated application of the elimination step in Lemma 8 produces a matrix decomposition for stabilizer parity check matrices that is in fact a canonical form. To prove uniqueness of the decomposition, a modification of Lemma 5 is needed:

Lemma 9. *Let $\beta : [r] \to [2n]$ be qubit-injective and let $R \in \mathcal{B}_{2n}\left(T^{tcr}(\beta)\right)$. If there exist lower triangular $L \in \mathcal{L}_m\left(Pairs(m)\right)$ and increasing $\alpha : [r] \to [m]$ such that $L\Pi(\alpha, \beta) = \Pi(\alpha, \beta)R$, then $R = I$.*

Proof. Take any $(i, j) \in T^{tcr}(\beta)$. The goal is to show $e_{2n,i}^{\top} R e_{2n,j} = 0$. Consider cases.

Case 1: $(i, j) \in T^{moves}(\beta)$**.** Then, $i = \beta(k)$ for some $k \in [r]$ and

$$e_{2n,i}^{\top} R e_{2n,j} = e_{m,\alpha(k)}^{\top} \Pi(\alpha, \beta) R e_{2n,j} = e_{m,\alpha(k)}^{\top} L\Pi(\alpha, \beta) e_{2n,j} \tag{77}$$

Consider subcases:

1. If $j \notin Im(\beta)$, then $\Pi(\alpha, \beta)e_{2n,j} = 0$.
2. If $j = \beta(l)$ for some $l \in [r]$, then $l > k$ and $e_{m,\alpha(k)} L e_{m,\alpha(l)} = 0$ because α is increasing and L is lower triangular.

Case 2: $(i, j) \in T^{rev}(\beta)$**.** Then, $(2n + 1 - j, 2n + 1 - i) \in T^{moves}(\beta)$ and $2n+1-j = \beta(k)$ for some $k \in [r]$. Apply the identity $R^{\top} = \Omega_{2n} R^{-1} \Omega_{2n}$, which holds for any element of $\mathcal{SP}_{2n}$:

$$\begin{aligned} e_{2n,i}^{\top} R e_{2n,j} &= e_{2n,j}^{\top} R^{\top} e_{2n,i} = e_{2n,j}^{\top} \Omega_{2n} R^{-1} \Omega_{2n} e_{2n,i} = e_{2n,2n+1-j}^{\top} R^{-1} e_{2n,2n+1-i} \\ &= e_{m,\alpha(k)} \Pi(\alpha, \beta) R^{-1} e_{2n,2n+1-i} = e_{m,\alpha(k)} L^{-1} \Pi(\alpha, \beta) e_{2n,2n+1-i} \end{aligned} \tag{78}$$

Now, proceed as in case 1 to conclude that this is zero. ☐

Now, all the ingredients for a canonical form for stabilizer parity check matrices are in place. To state the theorem, it is convenient to introduce the following terminology:

Definition 16. *Say that (r, α, β, L, R) is $(m, 2n)$ stabilizer allowed if $r \leq \min(m, n)$, $\alpha : [r] \to [m]$ is increasing, $\beta : [r] \to [2n]$ is qubit-injective, $L \in \mathcal{L}_m\left(T^{Left}(\alpha)\right)$, $R \in \mathcal{B}_{2n}\left(T^{tcr}(\beta)\right)$.*

Theorem 5. 1. *For every stabilizer parity check matrix $A \in \mathbb{F}_2^{m \times 2n}$, Gaussian elimination with symplectic moves on the right produces $(m, 2n)$ stabilizer allowed (r, α, β, L, R) such that*

$$A = L\Pi(\alpha, \beta)R \tag{79}$$

Moreover, the algorithm runs in time $O(mnr)$.

2. If (r, α, β, L, R) and $(r', \alpha', \beta', L', R')$ are $(m, 2n)$ stabilizer allowed and if

$$L\Pi(\alpha, \beta)R = L'\Pi(\alpha', \beta')R' \tag{80}$$

then

$$(r, \alpha, \beta, L, R) = (r', \alpha', \beta', L', R') \tag{81}$$

Proof. **Part 1** follows by repeated application of lemma 8.

Part 2: First, Lemma 4 implies $\Pi(\alpha, \beta) = \Pi(\alpha', \beta')$. Next, Lemma 9 implies $R = R'$. Finally, Lemma 6 implies $L = L'$. □

4.3 Canonical form for the symplectic and Clifford groups

For an element of $\mathcal{SP}_{2n}$, it is natural to apply Gaussian elimination with symplectic moves both on the left and on the right, in order to preserve the symplectic products of the rows and of the columns. This algorithm produces a canonical form for $\mathcal{SP}_{2n}$ in time $O(n^3)$. The tableau version of the algorithm produces a canonical form for $\mathcal{C}_n$ in time $O(n^3)$. Details follow.

Since elements of $\mathcal{SP}_{2n}$ are invertible, they have a pivot in every row. Therefore, the increasing function α that was used previously to describe the row indices of pivots becomes superfluous. Thus, for injective $\beta : [r] \to [2n]$, let

$$\Pi(\beta) = \sum_{i=1}^{r} e_{2n,i} e_{2n,\beta(i)}^\top \tag{82}$$

and say that $A \in \mathcal{SP}_{2n}$ is β partially reduced if rows $1, \ldots, r$ and columns $\beta(1), \ldots, \beta(r)$ of A coincide with the corresponding rows and columns of $\Pi(\beta)$.

As in the case of stabilizer parity check matrices, the symplectic constraint can be used to extract additional information about β reduced elements of $\mathcal{SP}_{2n}$:

Lemma 10. *Take $r \leq n$ and let $\beta : [r] \to [2n]$ be injective. Let $A \in \mathcal{SP}_{2n}$ be β partially reduced, i.e.*

$$\forall i \in [r]: \quad Ae_{2n,\beta(i)} = e_{2n,i} \quad \wedge \quad e_{2n,i}^\top A = e_{2n,\beta(i)}^\top \tag{83}$$

Then, β is qubit-injective and the following rows and columns of A are also reduced:

$$\forall i \in [r]: \quad Ae_{2n,2n+1-\beta(i)} = e_{2n,2n+1-i} \quad \wedge \quad e_{2n,2n+1-i}^\top A = e_{2n,2n+1-\beta(i)}^\top \tag{84}$$

Proof. β is qubit-injective because rows $1, \ldots, r$ of A have pairwise symplectic products 0.

Next, use $A\Omega_{2n}A^\top = A^\top \Omega_{2n} A = \Omega_{2n}$ and $e_{2n,\beta(i)}^\top A^\top = e_{2n,i}^\top, A^\top e_{2n,i} = e_{2n,\beta(i)}$ as follows:

$$Ae_{2n,2n+1-\beta(i)} = A\Omega_{2n}e_{2n,\beta(i)} = \Omega_{2n}(A^\top)^{-1}e_{2n,\beta(i)} = \Omega_{2n}e_{2n,i} = e_{2n,2n+1-i} \tag{85}$$

$$e_{2n,2n+1-i}^\top A = e_{2n,i}^\top \Omega_{2n} A = e_{2n,i}^\top (A^\top)^{-1}\Omega_{2n} = e_{2n,\beta(i)}^\top \Omega_{2n} = e_{2n,2n+1-\beta(i)}^\top \tag{86}$$

This completes the proof. $\square$

This lemma shows that the pivots in the first n rows uniquely determine all the rest and motivates the following:

Definition 17. *For $r \leq n$ and qubit-injective $\beta : [r] \to [2n]$, let*

$$\Pi^{sym}(\beta) = \sum_{i=1}^{r} \left(e_{2n,i} e_{2n,\beta(i)}^{\top} + e_{2n,2n+1-i} e_{2n,2n+1-\beta(i)}^{\top} \right) \tag{87}$$

As before, Gaussian elimination with symplectic moves both on the left and on the right takes a partially reduced matrix and simplifies it further. The typical step of this algorithm is:

Lemma 11. *Take $r < n$, qubit-injective $\beta : [r] \to [2n]$, and β partially reduced $A \in \mathcal{SP}_{2n}$. Let β' extend β to $[r+1]$ so that $(r+1, \beta'(r+1))$ is the position of the next pivot in the left and down order. Let $u, v \in \mathbb{F}_2^{2n}$ be such that*

$$A e_{2n,\beta'(r+1)} = e_{2n,r+1} + u \tag{88}$$

$$e_{2n,r+1}^{\top} A = e_{2n,\beta'(r+1)} + v^{\top} \tag{89}$$

and let

$$A' = S_{2n,u,r+1} A S_{2n,\beta'(r+1),v^{\top}} \tag{90}$$

Then:

1. *$S_{2n,u,r+1} \in \mathcal{B}_{2n}\left(Pairs\left(2n\right)\right)$.*
2. *β' is qubit injective.*
3. *$S_{2n,\beta'(r+1),v^{\top}} \in \mathcal{B}_{2n}\left(T^{tcr}\left(\beta'\right)\right)$.*
4. *$A' \in \mathcal{SP}_{2n}$ is β' partially reduced.*
5. *(β', u, v, A') can be computed from (A, β) in time $O(n^2)$.*

Proof. **Parts 1, 2 and 3:** Since A is β partially reduced, Lemma 10 implies $q_n(\beta'(r+1)) \notin \{q_n(\beta(1), \ldots, q_n(\beta(r))\}$, u is supported on a subset of $\{j : r+1 < j < 2n+1-r\}$, and v is supported on a subset of $\{j : j < \beta'(r+1), q_n(j) \notin \{q_n(\beta(1)), \ldots q_n(\beta(r))\}\}$.

Part 4: A' is a product of three elements of $\mathcal{SP}_{2n}$, so $A' \in \mathcal{SP}_{2n}$. Next, check the rows: for $i \in [r+1]$:

$$\begin{aligned} e_{2n,i}^{\top} A' &= e_{2n,i}^{\top} S_{2n,u,r+1} A S_{2n,\beta'(r+1),v^{\top}} = e_{2n,i}^{\top} A S_{2n,\beta'(r+1),v^{\top}} \\ &= \begin{cases} e_{2n,\beta(i)}^{\top} S_{2n,\beta'(r+1),v^{\top}} = e_{2n,\beta(i)}^{\top} & \text{if } i \leq r \\ (e_{2n,\beta'(r+1)}^{\top} + v^{\top}) S_{2n,\beta'(r+1),v^{\top}} = e_{2n,\beta'(r+1)}^{\top} & \text{if } i = r+1 \end{cases} \end{aligned} \tag{91}$$

Next, check the columns: for $i \in [r+1]$:

$$\begin{aligned}A' e_{2n,\beta'(i)} &= S_{2n,u,r+1} A S_{2n,\beta'(r+1),v^\top} e_{2n,\beta'(i)} = S_{2n,u,r+1} A e_{2n,\beta'(i)} \\ &= \begin{cases} S_{2n,u,r+1} e_{2n,i} = e_{2n,i} & \text{if } i \leq r \\ S_{2n,u,r+1}(e_{2n,r+1} + u) = e_{2n,r+1} & \text{if } i = r+1 \end{cases}\end{aligned} \tag{92}$$

Part 5: As before, the sparsity of $S_{2n,u,r+1}$ and $S_{2n,\beta'(r+1),v^\top}$ implies that multiplication by them can be performed in time $O(n^2)$. □

Now, all ingredients are in place to show that Gaussian elimination with symplectic moves on both sides produces a canonical form for $\mathcal{SP}_{2n}$ in time $O(n^3)$. To state the theorem, it is convenient to introduce the following terminology:

Definition 18. *Say that (β, L, R) is $\mathcal{SP}_{2n}$ allowed if $\beta : [n] \to [2n]$ is qubit-injective, $L \in \mathcal{B}_{2n}(Pairs(2n))$ and $R \in \mathcal{B}_{2n}(T^{tcr}(\beta))$.*

Theorem 6. *1. For every $A \in \mathcal{SP}_{2n}$, Gaussian elimination with symplectic moves on both sides produces $\mathcal{SP}_{2n}$ allowed (β, L, R) such that*

$$A = L \Pi^{sym}(\beta) R \tag{93}$$

Moreover, the algorithm runs in time $O(n^3)$.

2. If (β, L, R) and (β', L', R') are $\mathcal{SP}_{2n}$ allowed and if

$$L \Pi^{sym}(\beta) R = L' \Pi^{sym}(\beta') R' \tag{94}$$

then

$$(\beta, L, R) = (\beta', L', R') \tag{95}$$

Proof. **Part 1** follows from repeated application of Lemma 11.

Part 2: Lemma 4 implies $\Pi^{sym}(\beta) = \Pi^{sym}(\beta')$, so $\beta = \beta'$. Then, Lemma 9, applied to the first n rows of $(L')^{-1} L \Pi^{sym}(\beta) = \Pi^{sym}(\beta) R' R^{-1}$, gives $R = R'$. $L = L'$ follows. □

5 Finite blocklength bounds for stabilizer codes and Pauli noise

The preparatory subsection 5.1 establishes a property of the uniform distribution over the symplectic group that is used to derive the achievability bound. Another preparatory subsection 5.2 establishes notation necessary to state the bounds. Next, subsection 5.3 states and proves achievability and converse bounds on $\varepsilon_{errorguess}(p_{UV}, r)$ and $R_{errorguess}(p_{UV}, \epsilon)$ for arbitrary p_{UV}. Subsection 5.4 and subsection 5.5 show how the general bounds can be computed efficiently in the cases of the qubit erasure channel and the qubit depolarizing channel respectively. Moreover, it is seen there that the achievability and converse bounds on the rate differ by $O(1/n)$ for n qubits.

5.1 The uniform distribution over symplectic matrices

The following lemma will be useful later on:

Lemma 12. *Let C be uniformly distributed over $\mathcal{SP}_{2n}$, and let u be a non-zero vector in $\mathbb{F}_2^{2n}$. Then, Cu is uniformly distributed over non-zero vectors.*

Proof. Take non-zero vectors v, v'. Take symplectic matrix C' such that $C'v = v'$. Then, $\mathbb{P}(Cu = v) = \mathbb{P}(C'Cu = v') = \mathbb{P}(Cu = v')$, because $C'C$ is also uniformly distributed over symplectic matrices. □

The proof of Lemma 12 uses the fact that $\mathcal{SP}_{2n}$ acts transitively on $\mathbb{F}_2^{2n} \backslash \{0\}$. Previous expositions of the hashing bound also implicitly establish and use this transitive action; see for example [3, Section 3.2.3]. A short proof based on the symplectic Gaussian moves is as follows:

Lemma 13. *Let u, v be non-zero vectors in $\mathbb{F}_2^{2n}$. Then, there exists $C \in \mathcal{SP}_{2n}$ such that $Cu = v$.*

Proof. If there is i such that $e_{2n,i}^\top u = e_{2n,i}^\top v = 1$, then let $u' = u - e_{2n,i}$, $v' = v - e_{2n,i}$. Then, Lemma 3 implies that $S_{2n,v',i} S_{2n,u',i} u = S_{2n,v',i} e_{2n,i} = v$.

If there is no i such that $e_{2n,i}^\top u = e_{2n,i}^\top v = 1$, then let $i \neq j$ be such that u has an i-th but not a j-th component and v has a j-th but not an i-th component. Then, $S_{2n,j,i} u$ has a j-th component, which reduces this case to the previous one. □

5.2 Some additional notation

Take a distribution p_{UV} of a random vector U in $\mathbb{F}_2^{2n}$ and another discrete random variable V. For each value v that V can take, sort the vectors $u \in \mathbb{F}_2^{2n}$ in decreasing order according to the probability of the event $U = u, V = v$, resolving ties arbitrarily. Summarize this in a funciton ξ: for each v, $\xi(1, v)$, $\xi(2, v)$, …, $\xi(2^{2n}, v)$ are the 2^{2n} vectors in $\mathbb{F}_2^{2n}$ sorted so that

$$p_{UV}(\xi(1, v), v) \geq p_{UV}(\xi(2, v), v) \geq \cdots \geq p_{UV}(\xi(2^{2n}, v), v) \tag{96}$$

Let random variable J taking values in $\left[2^{2n}\right]$ be such that for each j and v, the event $J = j, V = v$ is the same as the event $U = \xi(j, v), V = v$.

5.3 General bounds

Using the notation in section 5.2, it is possible to formulate the following general bounds:

Theorem 7. *For every distribution p_{UV} and rate $r = k/n$, let*

$$\varepsilon_{errorguess}^{converse}(p_{UV}, r) = \mathbb{P}(J > 2^m) \tag{97}$$

$$\varepsilon_{errorguess}^{achievability}(p_{UV}, r) = \mathbb{P}(J > 2^m) + \mathbb{E}\left(\mathbb{I}(J \leq 2^m)(J-1)2^{-m}\right) \tag{98}$$

where $m = n - k$ and $\mathbb{I}(\cdot)$ is the indicator of an event. Then,

$$\varepsilon^{converse}_{errorguess}(p_{UV}, r) \le \varepsilon_{errorguess}(p_{UV}, r) \le \varepsilon^{achievability}_{errorguess}(p_{UV}, r) \tag{99}$$

Moreover, let

$$R^{achievability}_{errorguess}(p_{UV}, \epsilon) = \max\left\{r : \varepsilon^{achievability}_{errorguess}(p_{UV}, r) \le \epsilon\right\} \tag{100}$$

$$R^{converse}_{errorguess}(p_{UV}, \epsilon) = \min\left\{r : \varepsilon^{converse}_{errorguess}(p_{UV}, r) > \epsilon\right\} \tag{101}$$

Then,

$$R^{achievability}_{errorguess}(p_{UV}, \epsilon) \le R_{errorguess}(p_{UV}, \epsilon) < R^{converse}_{errorguess}(p_{UV}, \epsilon) \tag{102}$$

Proof. First, consider the converse bound on the optimal error probability. Without loss of generality, the optimal strategy is deterministic, so there is a fixed symplectic matrix C and the decoder D is a fixed mapping of pairs (s, v) of syndrome and side information to vectors $u \in \mathbb{F}_2^{2n}$. For every v, if U takes a value outside the image of D, then an error occurs. In the best case, for every v, the 2^m possible syndromes are mapped to the 2^m most likely vectors $\xi(1, v)$, $\ldots$, $\xi(2^m, v)$. Then, the probability of error is at least

$$\sum_v \sum_{j=2^m+1}^{2^{2n}} p_{UV}(\xi(j, v), v) = \mathbb{P}(J > 2^m) \tag{103}$$

Now, consider the achievability bound on the optimal error probability. Let C be uniformly distributed over symplectic matrices. Let the decoder D proceed as follows: try the vectors $\xi(1, v)$, $\xi(2, v)$, $\ldots$ in order until one matches the syndrome. Conditional on $J = j, V = v$, the union bound and Lemma 12 imply that the probability that for some $i < j$, $\xi(i, v)$ produces the same syndrome as $\xi(j, v)$ is at most $\min(1, (j-1)2^{-m})$. This proves the achievability bound.

Finally, bounds on the optimal error probability lead to bounds on the optimal rate.

If $\varepsilon^{achievability}_{errorguess}(p_{UV}, r) \le \epsilon$, then there is a strategy of rate r and error probability at most ϵ, so $R_{errorguess}(p_{UV}, \epsilon) \ge r$. The best lower bound that can be obtained in this way is $R^{achievability}_{errorguess}(p_{UV}, \epsilon)$.

If $\varepsilon^{converse}_{errorguess}(p_{UV}, r) > \epsilon$, then no strategy of rate r has error probability at most ϵ, so $R_{errorguess}(p_{UV}, \epsilon) < r$. The best upper bound that can be obtained in this way is $R^{converse}_{errorguess}(p_{UV}, \epsilon)$. □

5.4 The qubit erasure channel

As a first example, consider the qubit erasure channel (15). Let p_{UV} correspond to n independent erasure channels with parameter δ. The side information V is a vector of n independent *Bernoulli*(δ) random variables; it specifies the erased qubits. Conditional on $V = v$, U is uniformly distributed over the $2^{2|v|}$ vectors supported on the erased qubits, and J is uniformly distributed on $\left[2^{2|v|}\right]$.

These observations show that for the erasure channel, the general bounds in Theorem 7 can be expressed in terms of the cumulative distribution function of the binomial distribution. Since the CDF of the binomial can be computed efficiently and only a few calls to this function are needed, this gives an efficient algorithm for computing the bounds. Moreover, the resulting bounds on the rate differ by at most $O(1/n)$ from a simple function of n, δ and ϵ.

Theorem 8. *Let $\mathcal{F}(n,p,i)$ denote the probability that a Binomial(n,p) random variable is at most i:*

$$\mathcal{F}(n,p,i) = \sum_{i'=0}^{i} \binom{n}{i'} p^{i'} (1-p)^{n-i'} \tag{104}$$

Let p_{UV} correspond to n independent erasure channels with parameter δ. Then,

$$\begin{aligned}\varepsilon_{errorguess}^{converse}(p_{UV}, r) = \mathcal{F}\left(n, 1-\delta, n - \left\lfloor \frac{m}{2} \right\rfloor - 1\right) \\ - 2^m \left(\frac{4-3\delta}{4}\right)^n \mathcal{F}\left(n, \frac{4-4\delta}{4-3\delta}, n - \left\lfloor \frac{m}{2} \right\rfloor - 1\right)\end{aligned} \tag{105}$$

and

$$\begin{aligned}\varepsilon_{errorguess}^{achievability}(p_{UV}, r) = \left(1 + \frac{1}{2^{m+1}}\right) \mathcal{F}\left(n, 1-\delta, n - \left\lfloor \frac{m}{2} \right\rfloor - 1\right) - \frac{1}{2^{m+1}} \\ - \frac{2^m + 1}{2} \left(\frac{4-3\delta}{4}\right)^n \mathcal{F}\left(n, \frac{4-4\delta}{4-3\delta}, n - \left\lfloor \frac{m}{2} \right\rfloor - 1\right) \\ + \frac{(1+3\delta)^n}{2^{m+1}} \mathcal{F}\left(n, \frac{4\delta}{1+3\delta}, \left\lfloor \frac{m}{2} \right\rfloor\right)\end{aligned} \tag{106}$$

where $r = k/n$ and $m = n - k$.

Moreover, for fixed δ, ϵ, and for sufficiently large n,

$$R_{errorguess}(p_{UV}, \epsilon) = 1 - 2\delta + \frac{2\Phi^{-1}(\epsilon)\sqrt{\delta(1-\delta)}}{\sqrt{n}} + O\left(\frac{1}{n}\right) \tag{107}$$

where $\Phi(x) = 1/\sqrt{2\pi}\int_{-\infty}^{x} e^{-t^2/2} dt$ is the cumulative distribution function of the standard normal distribution.

Proof. For the converse bound on the error probability,

$$\mathbb{P}(J > 2^m) = \sum_{v} \mathbb{P}(V = v)\, \mathbb{P}\left(J > 2^m \middle| V = v\right) \tag{108}$$

and $\mathbb{P}\left(J > 2^m \middle| V = v\right)$ is zero when $2|v| \leq m$ and is $1 - 2^{m-2|v|}$ otherwise. This gives

$$\varepsilon_{errorguess}^{converse}(p_{UV}, r) = \mathbb{E}\left(\mathbb{I}(2|V| > m)\left(1 - 2^{m-2|V|}\right)\right) \tag{109}$$

Next,

$$\mathbb{E}\left(\mathbb{I}\left(2|V|>m\right)\right)=\mathcal{F}\left(n,1-\delta,n-\left\lfloor\frac{m}{2}\right\rfloor-1\right) \tag{110}$$

and

$$\begin{aligned}
&-\mathbb{E}\left(\mathbb{I}\left(2|V|>m\right)2^{m-2|V|}\right)\\
&\qquad=-2^m\sum_{i=0}^{n}\binom{n}{i}\delta^i4^{-i}(1-\delta)^{n-i}\mathbb{I}\left(2i>m\right)\\
&\qquad=-2^m\left(\frac{4-3\delta}{4}\right)^n\sum_{i=0}^{n}\binom{n}{i}\left(\frac{\delta}{4-3\delta}\right)^i\left(\frac{4-4\delta}{4-3\delta}\right)^{n-i}\mathbb{I}\left(2i>m\right)\\
&\qquad=-2^m\left(\frac{4-3\delta}{4}\right)^n\mathcal{F}\left(n,\frac{4-4\delta}{4-3\delta},n-\left\lfloor\frac{m}{2}\right\rfloor-1\right)
\end{aligned} \tag{111}$$

For the achievability bound on the error probability,

$$\begin{aligned}
&\mathbb{E}\left(\mathbb{I}\left(J\le 2^m\right)(J-1)2^{-m}\right)\\
&\qquad=\sum_t\mathbb{P}\left(V=v\right)\mathbb{E}\left(\mathbb{I}\left(J\le 2^m\right)(J-1)2^{-m}\middle|V=v\right)
\end{aligned} \tag{112}$$

and

$$\mathbb{E}\left(\mathbb{I}\left(J\le 2^m\right)(J-1)2^{-m}\middle|V=v\right)=\begin{cases}2^{2|v|-m-1}-2^{-m-1} & \text{if } 2|v|\le m\\ 2^{m-2|v|-1}-2^{-2|v|-1} & \text{if } 2|v|>m\end{cases} \tag{113}$$

This gives

$$\begin{aligned}
\varepsilon_{errorguess}^{achievability}\left(p_{UV},r\right)&=\mathbb{E}\left(\mathbb{I}\left(2|V|>m\right)\left(1-2^{m-2|V|-1}-2^{-2|V|-1}\right)\right)\\
&\quad+\mathbb{E}\left(\mathbb{I}\left(2|V|\le m\right)\left(2^{2|V|-m-1}-2^{-m-1}\right)\right)
\end{aligned} \tag{114}$$

Next,

$$\begin{aligned}
&\mathbb{E}\left(\mathbb{I}\left(2|V|>m\right)\left(1-2^{m-2|V|-1}-2^{-2|V|-1}\right)\right)\\
&\qquad=\mathcal{F}\left(n,1-\delta,n-\left\lfloor\frac{m}{2}\right\rfloor-1\right)\\
&\qquad\quad-\frac{2^m+1}{2}\left(\frac{4-3\delta}{4}\right)^n\mathcal{F}\left(n,\frac{4-4\delta}{4-3\delta},n-\left\lfloor\frac{m}{2}\right\rfloor-1\right)
\end{aligned} \tag{115}$$

and

$$\mathbb{E}\left(\mathbb{I}\left(2|V| \leq m\right) 2^{2|V|-m-1}\right)$$
$$= 2^{-m-1} \sum_{i=0}^{n} \binom{n}{i} \delta^i 4^i (1-\delta)^{n-i} \mathbb{I}\left(2i \leq m\right)$$
$$= 2^{-m-1}(1+3\delta)^n \sum_{i=0}^{n} \binom{n}{i} \left(\frac{4\delta}{1+3\delta}\right)^i \left(\frac{1-\delta}{1+3\delta}\right)^{n-i} \mathbb{I}\left(2i \leq m\right)$$
$$= 2^{-m-1}(1+3\delta)^n \mathcal{F}\left(n, \frac{4\delta}{1+3\delta}, \left\lfloor \frac{m}{2} \right\rfloor\right) \tag{116}$$

and

$$- 2^{-m-1}\mathbb{E}\left(\mathbb{I}\left(2|V| \leq m\right)\right) = -2^{-m-1}\mathcal{F}\left(n, \delta, \left\lfloor \frac{m}{2} \right\rfloor\right)$$
$$= -2^{-m-1}\left(1 - \mathcal{F}\left(n, 1-\delta, n - \left\lfloor \frac{m}{2} \right\rfloor - 1\right)\right) \tag{117}$$

For the converse bound on the rate, return to the expression from equation (111):

$$\mathbb{E}\left(\mathbb{I}\left(2|V| > m\right) 2^{m-2|V|}\right) = \sum_{2i>m} \binom{n}{i} \delta^i (1-\delta)^{n-i} 2^{m-2i} \tag{118}$$

The ratio of the $(i+1)$-st to the i-th term is

$$\frac{(n-i)\delta}{4(i+1)(1-\delta)} < \frac{(n-m/2)\delta}{4(m/2+1)(1-\delta)} < 1 \tag{119}$$

for $m > (2n\delta - 8(1-\delta))/(4-3\delta)$. Therefore, for m in this range the sum in (118) is upper bounded by the first binomial probability times a convergent geometric series. Therefore,

$$\varepsilon_{errorguess}^{converse}\left(p_{UV}, r\right) = \mathcal{F}\left(n, 1-\delta, n - \left\lfloor \frac{m}{2} \right\rfloor - 1\right) + O\left(\frac{1}{\sqrt{n}}\right) \tag{120}$$

Next, the Berry-Esseen theorem [4, 9] gives

$$\varepsilon_{errorguess}^{converse}\left(p_{UV}, r\right) = \Phi\left(\frac{n\delta - \lfloor m/2 \rfloor - 1}{\sqrt{n\delta(1-\delta)}}\right) + O\left(\frac{1}{\sqrt{n}}\right) \tag{121}$$

Finally, Taylor expansion of $\Phi^{-1}\left(\epsilon + O\left(1/\sqrt{n}\right)\right)$ shows that choosing

$$m = 2n\delta - 2\sqrt{n\delta(1-\delta)}\Phi^{-1}\left(\epsilon\right) + O(1) \tag{122}$$

suffices to ensure $\varepsilon_{errorguess}^{converse}\left(p_{UV}, r\right) > \epsilon$.

For the achievability bound on the rate,

$$\left(1+\frac{1}{2^{m+1}}\right)\mathcal{F}\left(n,1-\delta,n-\left\lfloor\frac{m}{2}\right\rfloor-1\right)-\frac{1}{2^{m+1}}$$
$$-\frac{2^m+1}{2}\left(\frac{4-3\delta}{4}\right)^n\mathcal{F}\left(n,\frac{4-4\delta}{4-3\delta},n-\left\lfloor\frac{m}{2}\right\rfloor-1\right)$$
$$\leq\Phi\left(\frac{n\delta-\lfloor m/2\rfloor-1}{\sqrt{n\delta(1-\delta)}}\right)+O\left(\frac{1}{\sqrt{n}}\right) \quad (123)$$

similarly to the argument for the converse bound. Moreover, the remaining term from equation (116) is

$$\mathbb{E}\left(\mathbb{I}\left(2|V|\leq m\right)\left(2^{2|V|-m-1}\right)\right)=\sum_{2i\leq m}\binom{n}{i}\delta^i(1-\delta)^{n-i}2^{2i-m-1} \quad (124)$$

The ratio of the $(i-1)$-st to the i-th term is

$$\frac{i(1-\delta)}{4(n-i-1)\delta}\leq\frac{(m/2)(1-\delta)}{4\delta(n-m/2+1)}<1 \quad (125)$$

when $m<8\delta(n+1)/(1+3\delta)$. Therefore, for m in this range, the sum in (124) is upper bounded by $O(1/\sqrt{n})$, again by the binomial probability times convergent geometric series argument. Thus, for $m<8\delta(n+1)/(1+3\delta)$,

$$\varepsilon_{errorguess}^{achievability}\left(p_{UV},r\right)\leq\Phi\left(\frac{n\delta-\lfloor m/2\rfloor-1}{\sqrt{n\delta(1-\delta)}}\right)+O\left(\frac{1}{\sqrt{n}}\right) \quad (126)$$

Therefore, choosing

$$m=2n\delta-2\sqrt{n\delta(1-\delta)}\Phi^{-1}\left(\epsilon\right)+O(1) \quad (127)$$

suffices to ensure $\varepsilon_{errorguess}^{achievability}\left(p_{UV},r\right)\leq\epsilon$. This completes the proof. □

5.5 The qubit depolarizing channel

As a second example, consider the qubit depolarizing channel (16). Consider n independent qubit depolarizing channels with parameter δ. In this case, there is no side information. The distribution of U is $\mathbb{P}\left(U=u\right)=(\delta/3)^{|u|}(1-\delta)^{n-|u|}$, where $|u|$ is the number of qubits for which the corresponding two entries of u are not 00. Conditional on $|U|=i$, J is a uniformly distributed integer between

$$|\{u:|u|\leq i-1\}|+1=4^n\mathcal{F}\left(n,\frac{3}{4},i-1\right)+1 \quad (128)$$

and

$$|\{u:|u|\leq i\}|=4^n\mathcal{F}\left(n,\frac{3}{4},i\right) \quad (129)$$

where the notation for the binomial CDF from equation (104) is used.

These observations show that the cumulative distribution function of J can be computed in terms of the CDF of the binomial.

Lemma 14. *For every n, p, extend the binomial CDF $\mathcal{F}(n,p,\cdot)$ from equation (104) to a piecewise linear function whose graph connects the points $(-1,0)$, $(0, \mathcal{F}(n,p,0))$, $\ldots$, $(n, \mathcal{F}(n,p,n))$. This extension is an increasing bijection that maps $[-1,n]$ to $[0,1]$; let $\mathcal{F}^{-1}(n,p,\cdot)$ be the inverse function.*

With this notation, the CDF of J for n independent depolarizing channels with parameter δ is

$$\mathbb{P}(J \leq j) = \mathcal{F}\left(n, \delta, \mathcal{F}^{-1}\left(n, \frac{3}{4}, \frac{j}{4^n}\right)\right) \tag{130}$$

Proof. For each $i = 0, \ldots, n$, the events $J/4^n \leq \mathcal{F}(n, 3/4, i)$ and $|U| \leq i$ coincide. Therefore, the CDF of $J/4^n$ can be extended to a piecewise linear function $[0,1] \to [0,1]$ whose graph connects the points $(0,0)$, $(\mathcal{F}(n,3/4,0), \mathcal{F}(n,\delta,0))$, $\ldots$, $(\mathcal{F}(n,3/4,n), \mathcal{F}(n,\delta,n))$. This can equivalently be expressed as the composition of two piecewise linear functions: $\mathcal{F}^{-1}(n,3/4,\cdot)$, whose graph connects the points $(0,-1)$, $(\mathcal{F}(n,3/4,0), 0)$, $\ldots$, $(\mathcal{F}(n,3/4,n), n)$, and $\mathcal{F}(n,\delta,\cdot)$, whose graph connects the points $(-1,0)$, $(0, \mathcal{F}(n,\delta,0))$, $\ldots$, $(n, \mathcal{F}(n,\delta,n))$. $\square$

Knowing the CDF of J in terms of the CDF of the binomial gives an efficient algorithm for computing the general bounds of Theorem 7 in the case of n independent depolarizing channels. Moreover, the resulting bounds on the rate differ by at most $O(1/n)$ from a simple function of n, δ and ϵ.

Theorem 9. *Let p_U correspond to n indepedent depolarizing channels with parameter δ. Then,*

$$\varepsilon^{converse}_{errorguess}(p_U, r) = \mathcal{F}(n, 1-\delta, n-1-\ell) \tag{131}$$

and

$$\begin{aligned}\varepsilon^{achievability}_{errorguess}(p_U, r) = {} & \left(1 + \frac{1}{2^{m+1}}\right)\mathcal{F}(n, 1-\delta, n-1-\ell) - \frac{1}{2^{m+1}} \\ & + 2^{m-1}(1-\delta)^n \left(\frac{\delta}{3-3\delta}\right)^{\lfloor \ell \rfloor + 1} \\ & + \frac{(16-16\delta)^n}{2^{m+1}}\left(\frac{3-4\delta}{3-3\delta}\right)\sum_{i=0}^{\lfloor \ell \rfloor}\left(\frac{\delta}{3-3\delta}\right)^i \mathcal{F}\left(n, \frac{3}{4}, i\right)^2 \end{aligned} \tag{132}$$

where $r = k/n$, $m = n-k$ and $\ell = \mathcal{F}^{-1}\left(n, \frac{3}{4}, \frac{2^m}{4^n}\right)$.

Moreover, for fixed δ, ϵ and for sufficiently large n,

$$\begin{aligned} R_{errorguess}(p_U, \epsilon) = {} & 1 - h(\delta) - \delta \log_2(3) \\ & - \sqrt{\frac{\delta(1-\delta)}{n}}\Phi^{-1}(\epsilon)\log_2\left(\frac{\delta}{3(1-\delta)}\right) + \frac{\log_2(n)}{2n} + O\left(\frac{1}{n}\right) \end{aligned} \tag{133}$$

where $h(\cdot)$ is the binary entropy.

Proof. For the converse bound on the error probability,

$$\mathbb{P}\left(J>2^m\right)=1-\mathbb{P}\left(J\le 2^m\right)=1-\mathcal{F}\left(n,\delta,\mathcal{F}^{-1}\left(n,\frac{3}{4},\frac{2^m}{4^n}\right)\right)$$
$$=1-\mathcal{F}\left(n,\delta,\ell\right)=\mathcal{F}\left(n,1-\delta,n-1-\ell\right) \quad (134)$$

For the achievability bound on the error probability,

$$\mathbb{P}\left(J>2^m\right)+\mathbb{E}\left(\mathbb{I}\left(J\le 2^m\right)(J-1)2^{-m}\right)$$
$$=\mathbb{P}\left(J>2^m\right)-\frac{1}{2^{m+1}}\mathbb{P}\left(J\le 2^m\right)+\frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J\le 2^m\right)\left(J-\frac{1}{2}\right)\right)$$
$$=\left(1+\frac{1}{2^{m+1}}\right)\mathcal{F}\left(n,1-\delta,n-1-\ell\right)-\frac{1}{2^{m+1}}$$
$$+\frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J\le 2^m\right)\left(J-\frac{1}{2}\right)\right) \quad (135)$$

Next, partition the event $J\le 2^m$ into the disjoint events

$$4^n\mathcal{F}\left(n,\frac{3}{4},i-1\right)<J\le 4^n\mathcal{F}\left(n,\frac{3}{4},i\right),\quad i=0,\ldots,\lfloor\ell\rfloor \quad (136)$$

and

$$4^n\mathcal{F}\left(n,\frac{3}{4},\lfloor\ell\rfloor\right)<J\le 4^n\mathcal{F}\left(n,\frac{3}{4},\ell\right)=2^m \quad (137)$$

and note that for $i=0,\ \ldots,\ \lfloor\ell\rfloor$,

$$\mathbb{E}\left(\mathbb{I}\left(4^n\mathcal{F}\left(n,\frac{3}{4},i-1\right)<J\le 4^n\mathcal{F}\left(n,\frac{3}{4},i\right)\right)\left(J-\frac{1}{2}\right)\right)$$
$$=\mathbb{P}\left(4^n\mathcal{F}\left(n,\frac{3}{4},i-1\right)<J\le 4^n\mathcal{F}\left(n,\frac{3}{4},i\right)\right)$$
$$*\,\mathbb{E}\left(J-\frac{1}{2}\middle|4^n\mathcal{F}\left(n,\frac{3}{4},i-1\right)<J\le 4^n\mathcal{F}\left(n,\frac{3}{4},i\right)\right)$$
$$=\left(4^n\mathcal{F}\left(n,\frac{3}{4},i\right)-4^n\mathcal{F}\left(n,\frac{3}{4},i-1\right)\right)\left(\frac{\delta}{3}\right)^i(1-\delta)^{n-i}$$
$$*\,\frac{4^n\mathcal{F}\left(n,3/4,i-1\right)+4^n\mathcal{F}\left(n,3/4,i\right)}{2}$$
$$=\frac{(16-16\delta)^n}{2}\left(\frac{\delta}{3-3\delta}\right)^i\left(\mathcal{F}\left(n,\frac{3}{4},i\right)^2-\mathcal{F}\left(n,\frac{3}{4},i-1\right)^2\right) \quad (138)$$

and, similarly,

$$\mathbb{E}\left(\mathbb{I}\left(4^n\mathcal{F}\left(n,\frac{3}{4},\lfloor\ell\rfloor\right)<J\le 4^n\mathcal{F}\left(n,\frac{3}{4},\ell\right)\right)\left(J-\frac{1}{2}\right)\right)$$
$$=\frac{(16-16\delta)^n}{2}\left(\frac{\delta}{3-3\delta}\right)^{\lfloor\ell\rfloor+1}\left(\mathcal{F}\left(n,\frac{3}{4},\ell\right)^2-\mathcal{F}\left(n,\frac{3}{4},\lfloor\ell\rfloor\right)^2\right) \quad (139)$$

Rearrange slightly to conclude that

$$\frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J \le 2^m\right)\left(J - \frac{1}{2}\right)\right) = \frac{(16-16\delta)^n}{2^{m+1}}\left(\frac{\delta}{3-3\delta}\right)^{\lfloor \ell \rfloor + 1}\mathcal{F}\left(n, \frac{3}{4}, \ell\right)^2 + \frac{(16-16\delta)^n}{2^{m+1}}\left(\frac{3-4\delta}{3-3\delta}\right)\sum_{i=0}^{\lfloor \ell \rfloor}\left(\frac{\delta}{3-3\delta}\right)^i \mathcal{F}\left(n, \frac{3}{4}, i\right)^2 \quad (140)$$

and use $\mathcal{F}(n, 3/4, \ell) = 2^m/4^n$ in the first term.

For the converse bound on the rate, note that the Berry-Esseen theorem implies

$$\varepsilon_{errorguess}^{converse}\left(p_U, r\right) = \Phi\left(\frac{n\delta - 1 - \ell}{\sqrt{n\delta(1-\delta)}}\right) + O\left(\frac{1}{\sqrt{n}}\right) \quad (141)$$

Next, Taylor expansion of $\Phi^{-1}\left(\epsilon + O\left(1/\sqrt{n}\right)\right)$ shows that choosing

$$\ell = n\delta - \sqrt{n\delta(1-\delta)}\Phi^{-1}(\epsilon) + O(1) \quad (142)$$

suffices to ensure $\varepsilon_{errorguess}^{converse}\left(p_U, r\right) > \epsilon$. According to Lemma 15 below, this choice of ℓ corresponds to

$$m = n(h(\delta) + \delta\log_2(3)) + \sqrt{n\delta(1-\delta)}\Phi^{-1}(\epsilon)\log_2\left(\frac{\delta}{3(1-\delta)}\right) - \frac{1}{2}\log_2(n) + O(1) \quad (143)$$

For the achievability bound on the rate, (150) in Lemma 15 combined with the Berry-Esseen theorem and Taylor expansion of $\Phi^{-1}\left(\epsilon + O(1/\sqrt{n})\right)$ imply that choosing

$$\ell = n\delta - \sqrt{n\delta(1-\delta)}\Phi^{-1}(\epsilon) + O(1) \quad (144)$$

suffices to ensure that

$$\begin{aligned}
&\varepsilon_{errorguess}^{achievability}\left(p_U, r\right) \\
&\quad = \mathbb{P}\left(J > 2^m\right) - \frac{1}{2^{m+1}}\mathbb{P}\left(J \le 2^m\right) + \frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J \le 2^m\right)\left(J - \frac{1}{2}\right)\right) \\
&\quad \le \mathbb{P}\left(J > 2^m\right) + \frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J \le 2^m\right)\left(J - \frac{1}{2}\right)\right) \\
&\quad \le \mathcal{F}\left(n, 1-\delta, n-1-\ell\right) + O\left(\frac{1}{\sqrt{n}}\right) \le \epsilon
\end{aligned} \quad (145)$$

As before, Lemma 15 implies that this choice of ℓ corresponds to

$$m = n(h(\delta) + \delta\log_2(3)) + \sqrt{n\delta(1-\delta)}\Phi^{-1}(\epsilon)\log_2\left(\frac{\delta}{3(1-\delta)}\right) - \frac{1}{2}\log_2(n) + O(1) \quad (146)$$

This completes the proof. □

The next lemma establishes estimates that are used in the proof of Theorem 9.

Lemma 15. *For fixed δ, ϵ and for sufficiently large n, if*

$$\mathcal{F}\left(n, \frac{3}{4}, \ell\right) = \frac{2^m}{4^n} \tag{147}$$

and

$$\ell = n\delta - \sqrt{n\delta(1-\delta)}\Phi^{-1}(\epsilon) + O(1) \tag{148}$$

then

$$m = n(h(\delta) + \delta \log_2(3)) + \sqrt{n\delta(1-\delta)}\Phi^{-1}(\epsilon) \log_2\left(\frac{\delta}{3(1-\delta)}\right) \\ - \frac{1}{2}\log_2(n) + O(1) \tag{149}$$

Moreover,

$$\frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J \leq 2^m\right)\left(J - \frac{1}{2}\right)\right) \leq O\left(\frac{1}{\sqrt{n}}\right) \tag{150}$$

where the implied constant does not depend on ℓ.

Proof. **Claim 1:**

$$\binom{n}{\lfloor \ell \rfloor}\left(\frac{3}{4}\right)^{\lfloor \ell \rfloor}\left(\frac{1}{4}\right)^{n-\lfloor \ell \rfloor} \leq \mathcal{F}\left(n, \frac{3}{4}, \ell\right) \leq \frac{3}{2}\binom{n}{\lceil \ell \rceil}\left(\frac{3}{4}\right)^{\lceil \ell \rceil}\left(\frac{1}{4}\right)^{n-\lceil \ell \rceil} \tag{151}$$

Proof of Claim 1: The lower bound follows immediately. For the upper bound, note that

$$\mathcal{F}\left(n, \frac{3}{4}, \ell\right) \leq \sum_{i=0}^{\lceil \ell \rceil}\binom{n}{i}\left(\frac{3}{4}\right)^{i}\left(\frac{1}{4}\right)^{n-i} \tag{152}$$

and the ratio of the $(i-1)$-st to the i-th term in the sum is

$$\frac{\binom{n}{i-1}\left(\frac{3}{4}\right)^{i-1}\left(\frac{1}{4}\right)^{n-i+1}}{\binom{n}{i}\left(\frac{3}{4}\right)^{i}\left(\frac{1}{4}\right)^{n-i}} = \frac{i}{3(n-i+1)} \leq \frac{2\delta}{3(1-2\delta)} \leq \frac{1}{3} \tag{153}$$

for the given range of ℓ and for $\delta \leq 1/4$ (which can be assumed without loss of generality because the hashing bound is negative already at $\delta = 1/5$). Therefore, the sum can be upper bounded by the last term times a convergent geometric series.

Claim 2:

$$\log_2 \mathcal{F}\left(n, \frac{3}{4}, \ell\right) = -n\mathbf{D}\left(\frac{\ell}{n}\middle\|\frac{3}{4}\right) - \frac{1}{2}\log_2(n) + O(1) \tag{154}$$

where $\mathbf{D}(\cdot\|\cdot)$ is the relative entropy.

Proof of Claim 2: Follows from (151) and the estimate [14, Chapter 10, Lemma 7]:

$$\log_2 \binom{n}{\lambda n} = nh(\lambda) - \frac{1}{2}\log_2(n) - \frac{1}{2}\log_2(\lambda(1-\lambda)) + O(1) \tag{155}$$

Note also that for the given range of ℓ, the $\log_2(\lambda(1-\lambda))$ can be absorbed in the $O(1)$. Similarly, the differences between ℓ, $\lfloor \ell \rfloor$ and $\lceil \ell \rceil$ amount to no more than $O(1)$.

Claim 3:

$$\mathbf{D}\left(\frac{\ell}{n}\middle\|\frac{3}{4}\right) = \mathbf{D}\left(\delta\middle\|\frac{3}{4}\right) - \sqrt{\frac{\delta(1-\delta)}{n}}\Phi^{-1}(\epsilon)\log_2\left(\frac{\delta}{3(1-\delta)}\right) + O\left(\frac{1}{n}\right) \tag{156}$$

Proof of Claim 3: Follows from the Taylor expansion

$$\mathbf{D}(x\|y) = \mathbf{D}(x_0\|y) + (x-x_0)\log_2\left(\frac{x_0(1-y)}{(1-x_0)y}\right) + O\left((x-x_0)^2\right) \tag{157}$$

and the given estimate for ℓ.

Claim 4: (154) and (156) and the identity $2 - \mathbf{D}(\delta\|3/4) = h(\delta) + \delta\log_2(3)$ prove (149).

Proof of Claim 4: follows by basic rearrangement.

Claim 5:

$$\sum_{i=0}^{\lfloor \ell \rfloor}\left(\frac{\delta}{3-3\delta}\right)^i \mathcal{F}\left(n,\frac{3}{4},i\right)^2 \le \frac{27}{11}\left(\frac{\delta}{3-3\delta}\right)^{\lfloor \ell \rfloor}\mathcal{F}\left(n,\frac{3}{4},\lfloor \ell \rfloor\right)^2 \tag{158}$$

Proof of Claim 5: From (153) deduce that

$$\frac{\mathcal{F}(n,3/4,i-1)}{\mathcal{F}(n,3/4,i)} \le \frac{2\delta}{3(1-2\delta)} \tag{159}$$

then bound the sum by the last term times a convergent geometric series with ratio

$$\frac{3(1-\delta)}{\delta}\frac{4\delta^2}{9(1-2\delta)^2} \le \frac{16}{27} \tag{160}$$

for $\delta \le 1/5$ (which can be assumed without loss of generality because the hashing bound is negative at $\delta = 1/5$).

Claim 6:

$$\frac{1}{2^m}\mathbb{E}\left(\mathbb{I}\left(J \le 2^m\right)\left(J - \frac{1}{2}\right)\right)$$
$$\le \frac{(16-16\delta)^n}{2^{m+1}}\left(\frac{\delta}{3-3\delta}\right)^{\lfloor \ell \rfloor + 1}\mathcal{F}\left(n, \frac{3}{4}, \ell\right)^2\left(1 + \frac{27}{11}\cdot\frac{3-4\delta}{\delta}\right)$$
$$\le \left(\frac{1}{2} + \frac{27(3-4\delta)}{22\delta}\right)2^m(1-\delta)^{n-\lfloor \ell \rfloor - 1}\left(\frac{\delta}{3}\right)^{\lfloor \ell \rfloor + 1}$$
$$\le \left(\frac{1}{2} + \frac{27(3-4\delta)}{22\delta}\right)\frac{2^m}{4^n}\frac{\binom{n}{\lfloor \ell \rfloor + 1}\delta^{\lfloor \ell \rfloor + 1}(1-\delta)^{n-\lfloor \ell \rfloor - 1}}{\binom{n}{\lfloor \ell \rfloor + 1}\left(\frac{3}{4}\right)^{\lfloor \ell \rfloor + 1}\left(\frac{1}{4}\right)^{n-\lfloor \ell \rfloor - 1}}$$
$$\le \left(\frac{1}{2} + \frac{27(3-4\delta)}{22\delta}\right)\frac{3}{2}\binom{n}{\lfloor \ell \rfloor + 1}\delta^{\lfloor \ell \rfloor + 1}(1-\delta)^{n-\lfloor \ell \rfloor - 1}$$
$$\le O\left(\frac{1}{\sqrt{n}}\right) \tag{161}$$

where the implied constant in the last line does not depend on ℓ. This proves (150).

Proof of Claim 6: The first step follows from (158) and (140). The second and third step follow by rearrangement. The fourth step follows from (151). The last step follows from the upper bound $O(1/\sqrt{n})$ on individual binomial probabilities. □

6 Conclusion

This article defines a family of groups of lower triangular matrices and uses them to establish a canonical form for unrestricted and stabilizer parity check matrices of a given size and rank. It also shows that the canonical form for the Clifford group can be computed in time $O(n^3)$, which improves upon the previously known time $O(n^6)$. Finally, the present article establishes a finite blocklength refinement of the hashing bound for stabilizer codes and Pauli noise, and shows that this achievability bound is nearly optimal among the class of arguments that use guessing the error as a substitute for guessing the coset.

A possible direction for future work is to investigate the relation between guessing the error and guessing the coset. The concatenated coding examples of [19, 7, 10, 11] show that very noisy depolarizing channels can have positive capacity even though the hashing bound is non-positive at those values of the channel parameter. It would be interesting to see whether the techniques in the present article can be used to construct further examples of strict separation between guessing the error and guessing the coset for stabilizer codes.

Acknowledgement

This work is supported by the Bavarian Ministry of Economic Affairs, Regional Development and Energy through the project High-Efficiency Stabilizer Codes.

References

[1] Alexei Ashikhmin. Fidelity lower bounds for stabilizer and css quantum codes. *IEEE Transactions on Information Theory*, 60(6):3104–3116, June 2014.

[2] Pascal Baßler, Matthias Zipper, Christopher Cedzich, Markus Heinrich, Patrick H. Huber, Michael Johanning, and Martin Kliesch. Synthesis of and compilation with time-optimal multi-qubit gates. *Quantum*, 7:984, April 2023.

[3] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, Nov 1996.

[4] Andrew C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49:122–136, 1941.

[5] Sergey Bravyi and Dmitri Maslov. Hadamard-free circuits expose the structure of the clifford group. *IEEE Transactions on Information Theory*, 67(7):4546–4563, July 2021.

[6] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.*, 78:405–408, Jan 1997.

[7] David P. DiVincenzo, Peter W. Shor, and John A. Smolin. Quantum-channel capacity of very noisy channels. *Phys. Rev. A*, 57:830–839, Feb 1998.

[8] Ross Duncan, Aleks Kissinger, Simon Perdrix, and John van de Wetering. Graph-theoretic Simplification of Quantum Circuits with the ZX-calculus. *Quantum*, 4:279, June 2020.

[9] Carl-Gustav Esseen. On the Liapounoff limit of error in the theory of probability. *Ark. Mat. Astr. Fys.*, 28A,(9):19, 1942.

[10] Jesse Fern. Correctable noise of quantum-error-correcting codes under adaptive concatenation. *Phys. Rev. A*, 77:010301, Jan 2008.

[11] Jesse Fern and K. Birgitta Whaley. Lower bounds on the nonzero capacity of pauli channels. *Phys. Rev. A*, 78:062335, Dec 2008.

[12] Daniel Gottesman. Class of quantum error-correcting codes saturating the quantum hamming bound. *Phys. Rev. A*, 54:1862–1868, Sep 1996.

[13] Andrey Boris Khesin, Jonathan Z. Lu, and Peter W. Shor. Graphical quantum clifford-encoder compilers from the zx calculus, 2023.

[14] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The theory of error-correcting codes*, volume 16. Elsevier, 1977.

[15] Dmitri Maslov and Martin Roetteler. Shorter stabilizer circuits via bruhat decomposition and quantum circuit transformations. *IEEE Transactions on Information Theory*, 64(7):4729–4738, 2018.

[16] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, jun 2012.

[17] Dimiter Ostrev. Quantum ldpc codes from intersecting subsets. *IEEE Transactions on Information Theory*, 70(8):5692–5709, Aug 2024.

[18] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdu. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359, 2010.

[19] Peter W. Shor and John A. Smolin. Quantum Error-Correcting Codes Need Not Completely Reveal the Error Syndrome. *arXiv e-prints*, pages quant–ph/9604006, April 1996.

[20] Gilbert Strang. The algebra of elimination. In *Excursions in Harmonic Analysis, Volume 3: The February Fourier Talks at the Norbert Wiener Center*, pages 3–22. Springer, 2015.

[21] Marco Tomamichel, Mario Berta, and Joseph M. Renes. Quantum coding with finite resources. *Nature Communications*, 7(1), may 2016.

[22] Mark M. Wilde. From classical to quantum shannon theory. *ArXiv*, abs/1106.1445, 2011.