# *Polylog-LDPC Capacity Achieving Codes for the Noisy Quantum Erasure Channel*



**Citation:** S. Lloyd, P. Shor and K. Thompson, "Polylog-LDPC Capacity Achieving Codes for the Noisy Quantum Erasure Channel," in IEEE Transactions on Information Theory, vol. 65, no. 11, pp. 7584-7595, Nov. 2019, doi: 10.1109/TIT.2019.2925100.

**As Published:** 10.1109/TIT.2019.2925100

**Publisher:** Institute of Electrical and Electronics Engineers (IEEE)

**Persistent URL:** https://hdl.handle.net/1721.1/126678

**Version:** Original manuscript: author's manuscript prior to formal peer review

# polylog-LDPC Capacity Achieving Codes for the Noisy Quantum Erasure Channel

Seth Lloyd [*], Peter Shor [†], and Kevin Thompson[‡]

July 11, 2018

**Abstract**

We provide *poly*log sparse quantum codes for correcting the erasure channel arbitrarily close to the capacity. Specifically, we provide $[[n, k, d]]$ quantum stabilizer codes that correct for the erasure channel arbitrarily close to the capacity if the erasure probability is at least 0.33, and with a generating set $\langle S_1, S_2, \ldots S_{n-k}\rangle$ such that $|S_i| \leq \log^{2+\zeta}(n)$ for all $i$ and for any $\zeta > 0$ with high probability. In this work we show that the result of Delfosse et al. [5] is tight: one can construct capacity approaching codes with weight almost $O(1)$.

## 1 Introduction

Graph States [22, 12] are a very useful set of stabilizer states that have found many applications in quantum information theory. They are important components of measurement based computation schemes [21], and quantum error correcting codes [22, 7, 4, 19, 8, 11, 13, 9, 15] to name a few. One reason they are so ubiquitous in quantum computing is that they have an easy to understand entanglement structure [12]. In some sense they provide a "standard form" for stabilizer states [4, 21], since any stabilizer state is locally equivalent to a graph state, and any stabilizer quantum code is equivalent to a quantum graph code[7].

Indeed, given a graph state over a composite quantum system, it is well known how to compute the entanglement entropy between any partition of the subsystems [12]. In this paper Hein shows that the entanglement entropy for a two set partition of a graph state depends in a natural way on the rank of the "cut" matrix separating the two sets. This idea was generalized to qudits and explored to give conditions for secret sharing in [13]. Our construction leverages exactly this property of graph states to produce quantum codes with interesting properties.

In particular, we supply quantum stabilizer codes that approach the capacity of the quantum erasure channel[1] and are log-LDPC (Low Density Parity Check). This means that the stabilizer group of the code can be generated by elements of the Pauli group that are only polylogarithmic with the total number of qubits. We stress that the codes described make little progress toward the so called quantum LDPC conjecture[2, 24, 10][1]. This conjecture posits the existence of locally generated codes with linear minimum distance. Our codes are indeed generated by stabilizers which are only

[*]Department of Mechanical Engineering, MIT
[†]Department of Mathematics, MIT
[‡]School of Engineering and Applied Science, Harvard

[1]See the second reference and the references it contains for many other LDPC constructions

logarithmic in their length, but we expect the distance of these codes to also only be logarithmic in the number of qubits, by the arguments presented in [15].

The codes described in this paper have very poor adversarial distance ($O(\log(n))$), however the adversarial errors that destroy the encoded information are very unlikely. This feature allows us to correct for the quantum erasure channel at capacity. It is known that this cannot be achieved by stabilizer codes of constant weight [5], we demonstrate that it is possible with a stabilizer code with generators that are only $O(\log^{2+\zeta}(n))$ for any $\zeta > 0$.

It is interesting to contrast the quantum case with the classical. Classical LDPC codes with linear distance have been known for some time [6]. Even random classical LDPC codes can be shown to have linear distance generically. Constructions of Quantum LDPC codes with linear distance are not known, despite much research in this area[2, 24, 10]. In both the classical and quantum regime there are bounds demonstrating that the capacity of the erasure channel cannot be achieved with LDPC codes[5][3]. For classical codes, there are known constructions[18] with 'barely' diverging parity check weight that can correct for the erasure channel very near the capacity. We provide such a construction for the quantum setting. Note further that Reed-Muller codes have recently been shown to achieve the capacity of both the classical and quantum erasure channels [16, 17].[2] To contrast this work with ours, note that Reed-Muller codes have locality which is linear in the number of qubits.

Our proof techniques are along the lines of standard 'typicality' arguments in information theory, along with some simple observations about graph states and the rank of random matrices. We construct a quantum graph code by randomly sampling an Erdos-Renyi graph and constructing the corresponding graph state. In addition, we randomly sample the parity check matrix for a classical code, and use these to construct a quantum graph code. We send this randomly sampled code through the erasure channel, and apply a result from[14], as well as standard typicality arguments to show our scheme succeeds with high probability.

## 2 Notation

Graphs will be denoted by the pair $(V, E)$ where $V$ is the vertex set of the graph and $E$ is the edge set. $V$ is some set of the form $\{1, \ldots n\}$. As is standard, the edge set $E$ will be of the form $\{(i, j)\}$ where each $i, j \in V$. We will study undirected graphs, so it is important to treat $(i, j)$ as unordered. For a vertex $i \in V$ we denote the neighbor set as usual $N(i) = \{j \in V \,|\, (i, j) \in E\}$. Denote the adjacency matrix $A$. We will also denote the "top half" of the adjacency matrix as $A_{top}$. If $i \geq j$ then $A_{top,ij} = A_{ij}$, otherwise $A_{top,ij} = 0$.

Let $K$ be a subset of the vertices. We will denote the cut matrix between $K$ and $V \backslash K$ as $A_{cut}$. $A_{cut}$ is a $|V \setminus K| \times |K|$ matrix such that if $A_{ij} = 1$ for $j \in K$ and $i \in V \setminus K$ then $A_{cut}$ of the corresponding entry is 1. Otherwise $A_{cut}$ is zero.

All vectors in $\mathbb{F}_2^n$ will be column vectors by default, although occasionally we will refer to a row vector with parenthesis, and a concatenation of two vectors with parenthesis.

For classical codes, we will use the standard $[n, k, d]$ notation. We will say some code $C$ is a $[n, k, d]$ code if $C$ is a $k$ dimensional subspace of $\mathbb{F}_2^n$ such that $\min_{x \in C} |x| = d$, where $|\cdot|$ denotes the Hamming weight. We will denote a code $C$ as a $[n, k, d, w]$ code if the code is a $[n, k, d]$ code and its linear dual can be generated by words of weight at most $w$. We will use the analogous $[[n, k, d, w]]$ for quantum

[2] In the later case, it was shown that CSS codes where the $X$ and $Z$ checks are Reed-Muller codes can correct for the quantum erasure channel.

stabilizer codes. Here $w$ stands for the largest weight of a generator for the stabilizer group. If we have a code $C$ on $n$ bits, and some subset $K \subseteq [n]$ we will use the notation $C_K$ to denote the code $C$ restricted to the bits $K$.

Suppose $w = (w_1, w_2, \ldots w_n)$ is some binary vector in $\mathbb{F}_2^n$, and suppose $\{S_1, S_2, \ldots S_n\}$ is some set of Pauli group elements over $n$ qubits. We will use the notation:

$$\prod_{i \in w} S_i := \prod_{i \,|\, w_i = 1} S_i \tag{1}$$

For matrices $A$ and $B$, we will use the standard notation $[A, B] := AB - BA$. The capital letters $X$, $Y$ and $Z$ will be reserved for the Pauli matrices:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \; Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \; Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2}$$

In an $n$ qubit system, we will use $X_i$ to denote a Pauli $X$ operator acting on qubit $i$, and similarly for $Y$ and $Z$. For a binary vector $v \in \mathbb{F}_2^n$, we define:

$$X_v := \prod_{i : v(i) = 1} X_i \tag{3}$$

and similarly for $Y$ and $Z$

All logarithm functions will be natural logarithm by default.

For convenience, we will include a definition of a Bernoulli random variable to refer to later:

$$\text{Bern}(p) = \begin{cases} 1 \text{ with probability } p \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

and we will denote the binary entropy function as $h(x)$:

$$h(x) := -x \log_2(x) - (1 - x) \log_2(1 - x) \tag{5}$$

# 3 Definitions

For us, perhaps the most important definition is that of a graph state. We will only state the definition, for examples we invite the reader to examine any of several comprehensive reviews [12, 11]. It may not be clear that the following definitions are equivalent a priori, proofs of equivalence can be found in the stated references.

**Definition 1.** *[Graph state] Given some graph $G = (V, E)$ with no self loops, associate the vertices of the graph $G$ to the numbers $[1, 2, \ldots n]$, we define the graph state $|G\rangle$ according to three equivalent definitions:*

1. *Let $A_{top}$ be as defined in the notations section. We define:*

$$|G\rangle = \frac{1}{2^{n/2}} \sum_{x \in \mathbb{F}_2^n} (-1)^{x^T A_{top} x} |x\rangle \tag{6}$$

2. *Define the following Pauli group elements:*

$$\forall i \in V \;\; S_i := X_i \prod_{j \in N(i)} Z_j \tag{7}$$

*The graph state $|G\rangle$ is defined as the unique state stabilized by all $S_i$.*

3. *Let $CP_{ij}$ be the standard controlled phase operation between qubits $i$ and $j$. Let $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. $|G\rangle$ can be defined as:*

$$|G\rangle = \prod_{(i,j) \in E} CP_{ij} |+\rangle \otimes \ldots \otimes |+\rangle \tag{8}$$

Graph states satisfy an important orthogonality property:

**Lemma 1.** *[13] Let $x \in \mathbb{F}_2^n$ be some nonzero binary string. Then,*

$$\langle G|Z_x|G\rangle = 0 \tag{9}$$

*Proof.* Since $Z$ operators commute with controlled phase operators, according to definition 1 we calculate:

$$\langle G|Z_x|G\rangle = \tag{10}$$

$$\langle +| \ldots \langle +| \prod_{ij} CP_{ij} Z_x \prod_{ij} CP_{ij} |+\rangle \ldots |+\rangle \tag{11}$$

$$= \langle +| \ldots \langle +|Z_x|+\rangle \ldots |+\rangle = 0 \tag{12}$$

Where we used $CP_{ij}CP_{ij} = \mathbb{I}$ to get from eq. (11) to eq. (12). □

With this definition in hand we can define a quantum graph code.

**Definition 2** ([23])**.** *Given a graph $G$ and a $[n, k, d]$ classical code $C$ over $\mathbb{F}_2$, we define a graph code $(G, C)$ as the linear span of quantum states of the form:*

$$Z_c|G\rangle \tag{13}$$

*where $c$ is any binary code word in $C$.*

Now we will present a few simple facts regarding this definition.

**Lemma 2.** *Let $(G, C)$ be as defined in definition 2 and suppose it has parameters $[[n, k_Q, d_Q, w_Q]]$. Suppose the code $C$ has parameters $[n, k, d, w]$. Denote the maximum vertex degree in $G$ as $K_{max}$ and the minimum vertex degree as $K_{min}$.*

1. *$k = k_Q$*

2. *[15] Let $\{h_1, \ldots h_{n-k}\}$ be some minimal weight generating set for the code $C^\perp$, and define:*

$$g_j := \prod_{i \in h_j} S_i \tag{14}$$

*The code $(G, C)$ is a stabilizer code code with stabilizer generators $\{g_j\}$ for all $j$.*

3. *[15]* $w_Q \leq wK_{max}$

4. *[15]* $d_Q \leq K_{min}$

*Proof.* We can see item 1 immediately from lemma 1 and definition 2.

We can prove 2 as follows. We claim that the set $\{g_j\}$ form a complete, minimal set of generators for the stabilizer group for the code defined as the span of all $Z_c|G\rangle$. Independence of these operators follows from independence of the binary vectors $h_i$. They also all clearly commute since $[S_i, S_j] = 0$ for all $i$ and $j$. It remains to show that these operators stabilize the code. The code $(G, C)$ is the span of states of the form $Z_c|G\rangle$ where $c \in C$. Suppose $g_j$ has the form:

$$\prod_{i \in h_j} S_i = (-1)^{\phi} Z_w \left( \prod_{j \in h_k} X_j \right) \tag{15}$$

for some binary vector $w$. We have just rewritten the operator so that the $X$ operators are on the right and the $Z$ operators on the left. Potentially we have introduced a phase $\phi \in \{0, 1\}$.

$$\prod_{i \in h_j} S_i Z_c|G\rangle = (-1)^{\phi} Z_w \left( \prod_{i \in h_j} X_i \right) Z_c|G\rangle \tag{16}$$

Since $h_j \in C^{\perp}$,

$$\left[ \prod_{i \in h_j} X_i, Z_c \right] = 0 \tag{17}$$

so

$$(-1)^{\phi} Z_w \left( \prod_{i \in h_j} X_i \right) Z_c|G\rangle = \tag{18}$$

$$Z_c (-1)^{\phi} Z_w \left( \prod_{i \in h_j} X_i \right) |G\rangle = Z_c|G\rangle$$

Item 3 follows from item 2. The stabilizers are given, and the weight of each stabilizer is upper bounded (by definition) by the maximum vertex degree times the maximum weight of $h_j$.

We sketch the proof of item 4. Focusing on the bit in the graph with smallest degree, an adversary can "disconnect" this vertex from the rest of the graph by enacting a unitary on this bit and its neighbors. Then, the adversary can induce undetectable phase on the code by acting a Pauli on the disconnected bit. Hence the adversarial distance is upper bounded by the minimal vertex degree.

□

The classical erasure channel "erases" any particular message bit with probability $p$. This corresponds to replacing that symbol with a special erasure symbol $e$. On the other side of the channel, the user knows exactly which bits were erased and knows the other bits were unaltered by the channel. The quantum erasure channel is the natural quantum analog of this channel. Each transmitted bit is replaced with a special erasure state $|e\rangle$. The state $|e\rangle$ is orthogonal to both $|0\rangle$ and $|1\rangle$ so that the user can unambiguously determine which bits were erased and which bits were unaffected by the channel.

**Definition 3.** *The erasure channel with probability p will be denoted as $\mathcal{E}_p$. It acts on density matrices of qubits as:*

$$\mathcal{E}_p(\rho) = (1-p)\rho + p|e\rangle\langle e| \tag{19}$$

*where the state $|e\rangle$ is outside the qubit space ($\langle e|0\rangle = 0 = \langle e|1\rangle$).*

It has been well known for some time that the capacity of this channel for communicating quantum information is $1-2p$:

**Lemma 3.** *[1] For any $R < 1-2p$, and for any $\delta > 0$, there exists a family of subspaces $\{V_n\}_{n=1}^{\infty}$ satisfying $\{V_n\} \subset [\mathbb{C}^2]^{\otimes n}$ and $\log_2(\dim(V_n)) = \lfloor Rn \rfloor$ as well as a family of decoding operations $\{\mathcal{R}_n\}_{n=1}^{\infty}$ such that for all n sufficiently large:*

$$\forall |\psi\rangle \in V_n, \left|\mathcal{R}_n\left(\mathcal{E}_p^{\otimes n}\left(|\psi\rangle\langle\psi|\right)\right) - |\psi\rangle\langle\psi|\right| < \delta \tag{20}$$

*in trace distance.*

*In addition, any family of codes with rate exceeding $1-2p$ is not correctable after being sent through the erasure channel[20][25].*

# 4 Preliminaries

With these definitions in hand, we can present the lemmas we will need to prove that our construction approaches the erasure channel. First, we will need several properties of graph states:

**Lemma 4.** *Let $G = (V, E)$ be a graph on n vertices. Further, let $K$ be some subset of the bits with complement $V \setminus K$. Let $A_{cut}$ be the cut matrix across the cut $(K, V \setminus K)$, let $A_K$ be the upper half of the adjacency matrix restricted to the bits $K$, and let $G'$ be the subgraph induced by the vertices $V \setminus K$.*

1. *[13]Suppose a user measures the bits $K$ in the computational basis. Every bit string $y$ is equally likely and after the user measures and gets bit string $y$, the resulting quantum state is:*

$$\langle y|_K|G\rangle = \frac{1}{2^{|K|/2}}(-1)^{y^T A_K y} Z_{A_{cut}y}|G'\rangle \tag{21}$$

2. *[12] Denote the entanglement entropy of the graph state across the cut $(K, V\setminus K)$ as $E^{(K,V\setminus K}(|G\rangle)$. Then,*

$$E^{(K,V\setminus K)}(|G\rangle) = \mathrm{rank}_{\mathbb{F}_2}(A_{cut}) \tag{22}$$

In addition we will also need the following technical result on the rank of randomly chosen matrices:

**Theorem 1.** *[Rank of Sparse Matrices, [14]] Let $A$ be a binary $\alpha n \times n$ matrix with $\alpha < 1$a constant. Suppose each entry of $A$ is sampled independently according to Bern $\left(\frac{w\log(n)}{n}\right)$. The expected number of linear combinations of rows that add to $\mathbf{0}$, or the expected number of critical sets is*

$$1 + O\left(\frac{1}{n^{w-1}}\right) \tag{23}$$

*Excluding the "all zeros" linear combination or the empty set, the number of critical sets is:*

$$O\left(\frac{1}{n^{w-1}}\right) \tag{24}$$

*Proof.* Slight modifications of the proof of Lemma 3.3.2 in [14] yield the result. Note for reference that

$$1 - \frac{2w\log(n)}{n} \leq 1 - \frac{2\log(n)}{n} \tag{25}$$

so the upper bounds present in the proof follow immediately. □

Hence, by Markov's inequality, if a matrix $A$ is sampled as described above, then the probability that $A$ is not full rank is upper bounded by $O\left(\frac{1}{n^{w-1}}\right)$. Combined with the Rank Nullity theorem from linear algebra, this theorem says that if we randomly sample a parity check matrix from this ensemble, then the rows will be linearly independent with high probability. Hence, with high probability a random $m \times n$ parity check matrix $H$ will yield a code with dimension $n - m$. We will need a similar result on the distance of a code sampled from this ensemble. We show that such a code has linear distance with high probability, and that the code retains this linear distance after the erasure channel:

**Lemma 5** (Linear distance)**.** *Let $H$ be a $\alpha n \times n$ binary matrix with $\alpha < 1$, where each entry is chosen uniformly at random according to i.i.d. Bern$(q)$ where $q = \frac{w\log(n)}{n}$. Let $C$ be the code with $H$ as its parity check matrix. Further, let $d_C$ be the distance of the code $C$.*

*Let $K$ be the first $\beta n$ bits (corresponding to the first $\beta n$ columns of $H$) and assume that $\alpha > \beta$. Denote the code $C$ restricted to the bits $V \setminus K$ as $C_{V \setminus K}$. If $\varepsilon'$ satisfies:*

$$(1-\beta)h\left(\frac{\varepsilon'}{1-\beta}\right) < \alpha - \beta \quad \text{and} \quad h(\varepsilon') < \alpha \tag{26}$$

*then,*

$$d_{C_{V \setminus K}} > \varepsilon' n \tag{27}$$

*with probability at least*

$$1 - O\left(\frac{1}{n^{w\alpha - 2}}\right) \tag{28}$$

*Proof.* See Appendix. □

We will also require the following simple result on the sum of binary random variables:

**Lemma 6** ([6] Lemma 4.1)**.** *Let $\{B_i\}$ be a set of $k$ independent Bern$(p)$ random variables. Then we have that*

$$\mathbb{P}\left[\sum_{i=1}^{k} B_i = 1 \ mod\ 2\right] = \frac{1-(1-2p)^k}{2} \tag{29}$$

This establishes the following corollary:

**Corollary 1.** *Let $\{b_i\}$ be a set of $k$ vectors of random variables such that all entries of each $b_i$ are independent $Bern(p)$ random variables. For any fixed word $c \in \mathbb{F}_2^n$ such that $|c| = g$,*

$$\mathbb{P}\left[\sum_i b_i = c\right] = \frac{1}{2^n}\left[1 + (1-2p)^k\right]^{n-g}\left[1 - (1-2p)^k\right]^g \tag{30}$$

We will need the "Principle of Implicit Measurement" to analyze the correcting properties of our graph codes. It is standard in quantum information theory, and it states that "lost" quantum systems can be assumed to be measured in some basis of our choosing. The user can treat the system as though it was measured, but does not know the outcome.

**Theorem 2.** *[Principle of Implicit measurement] Let $|\psi\rangle = \sum_{i,j} c_{i,j}|i\rangle_K|j\rangle_{V\setminus K}$ be a pure quantum state on $n = |V|$ qubits. If the qubits $K$ are lost, then the reduced density matrix on qubits $V \setminus K$ is given by the partial trace:*

$$\rho_{V\setminus K} = Tr_K(|\psi\rangle\langle\psi|) = \sum_i \left(\langle i|_K \otimes \mathbb{I}_{V\setminus K}\right)|\psi\rangle\langle\psi|\left(|i\rangle_K \otimes \mathbb{I}_{V\setminus K}\right) \tag{31}$$

Lastly, we will state the Chernoff and Markov bounds for later use:

**Lemma 7.** *[Chernoff and Markov Bounds] Let $X$ be a random variable. Then, we have:*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(e^{t\cdot X})}{e^{t\cdot a}} \qquad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a} \tag{32}$$

We will make use of a simple corollary to the Chernoff bound:

**Corollary 2.** *Let $X = \sum_{i=1}^n X_i$ be a random variable with average $\mathbb{E}(X) = \mu$ such that the variables $X_i$ are independent. Then, it holds that:*

$$\mathbb{P}\left(|X - \mu| \geq \gamma\mu\right) \leq 2e^{-\frac{\gamma^2\mu}{3}} \tag{33}$$

# 5 polylog-LDPC Codes

## 5.1 Coset Measurement

For our recovery scheme, we will require the notion of a "coset" measurement. Such a measurement allows us to distinguish between different labeled graph states that are in different cosets of $C$. Consider $C$ as a subgroup of $\mathbb{F}_2^n$. For each vector $e \in \mathbb{F}_2^n$ we have the coset $C + e$. Let $\{C + e\}$ be the set of all cosets of the code $C$. Define the following subspaces:

$$V_e = \text{span}_{c\in C} Z_{c+e}|G\rangle \tag{34}$$

Further, define $P_e$ to be the projection onto this subspace. We define the observable:

$$\mathcal{O} = \sum_{\{C+e\}} \lambda_e P_e \tag{35}$$

where the sum goes over distinct cosets of $C$ and for $e \neq e'$ we have $\lambda_e \neq \lambda_{e'}$. It is well known that distinct cosets are disjoint, so $P_e P_{e'} = 0$ for $e \neq e'$. Hence, such an observable is well defined and serves to distinguish cosets.

### 5.2 Recovery Operation

Suppose we send the state $|\psi\rangle \in (G, C)$ through an erasure channel and $pn$ bits are erased. Denote the dimension of the code $C$ as $Rn$, and the set of erased bits $K$. Construct the following $[(R+p)n] \times (1-p)n$ matrix $F$. Let the first $Rn$ rows of $F$ be the generators of the code $C$ restricted to the non erased bits, and let the remaining rows be the transpose of the cut matrix between the erased bits and the non-erased bits:

$$F = \begin{array}{c} \overset{(1-p)n}{\longleftrightarrow} \\ \left[\begin{array}{c} C_{V\setminus K} \\ A_{cut}^T \end{array}\right] \end{array} \begin{array}{l} \\ \updownarrow (R+p)n \end{array} \tag{36}$$

Further, let $G'$ be the subgraph of $G$ induced by the vertex set $V \setminus K$.

**Lemma 8.** *If the matrix $F$ is full rank over $\mathbb{F}_2$, then there is a quantum operation $\mathcal{R}$ which recovers from the erasure.*

*Proof.* Denote the quantum state before the channel as:

$$|\psi\rangle = \sum_{c \in C} b_c Z_c |G\rangle \tag{37}$$

Suppose without loss of generality that the erased bits, $K$, are the first $pn$ bits (or rearrange the bits so that this is the case). For each codeword $c \in C$, decompose $c$ as the concatenation of its value on the erased bits with its value on the non erased bits: $c = (c_K, c_{V\setminus K})$. By hypothesis, and by lemma 4, after the erasure channel, we are left with the following density matrix on the subsystems $V \setminus K$:

$$\rho = \sum_{j \in \mathbb{F}_2^{|K|}} \rho_j |\phi_j\rangle\langle\phi_j| \tag{38}$$

where:

$$\sum \rho_j = 1 \quad \text{and} \quad |\phi_j\rangle = \sum_{c \in C} b_c (-1)^{j \cdot c_K} Z_{c_{V\setminus K} + A_{cut} j} |G'\rangle_{V\setminus K} \tag{39}$$

If $F$ is full rank, then the coset measurement of $C_{V\setminus K}$ can be used to determine the string $j$. Indeed, if $F$ is full rank then each element of the range of $A_{cut}$ belongs to a different coset. So, measuring which coset the graph state falls into will determine the string $j$. Note further that such a measurement will not disturb the encoded information since for fixed $j$ the states $\{Z_{c_{V\setminus K} + A_{cut} j}\}$ all lie in a particular coset of $C_{V\setminus K}$.

Let us describe this more formally. Suppose that $A_{cut} j$ falls into a particular coset of $C_{V\setminus K}$:

$$A_{cut} j = c''_{V\setminus K} + e \tag{40}$$

for some fixed $c''_{V\setminus K} \in C_{V\setminus K}$. If $P_{e'}$ is the projector onto a different coset, then $P_{e'}|\phi_j\rangle = 0$:

$$P_{e'}|\phi_j\rangle = \sum_{c,c' \in C} \sigma_{(c,c',e',j)} \langle G'| Z_{A_{cut} j + c_{V\setminus K} + c'_{V\setminus K} + e'} |G'\rangle \tag{41}$$

for some operators $\sigma_{(c,c',e',j)}$. Since $e'$ corresponds to a distinct coset, the string $A_{cut} j + c_{V\setminus K} + c'_{V\setminus K} + e'$ is always in some nonzero coset, so by lemma 1, $P_{e'}|\phi_j\rangle = 0$.

The projector $P_e$ has the following effect on the state $|\phi_j\rangle$:

$$P_e|\phi_j\rangle = \left[\sum_{c'_{V\setminus K}} Z_{c'_{V\setminus K}+e}|G'\rangle\langle G'|Z_{c'_{V\setminus K}+e}\right]|\phi_j\rangle \tag{42}$$
$$= \sum_{\substack{c'_{V\setminus K}, c_{V\setminus K}: \\ c'_{V\setminus K}+e=c_{V\setminus K}+A_{cut}j}} b_c(-1)^{j\cdot c_K} Z_{c'_{V\setminus K}+e}|G'\rangle$$

The relation $c'_{V\setminus K}+e = c_{V\setminus K}+A_{cut}j$ fixes $c'_{V\setminus K}$ given $c_{V\setminus K}$. It is easy to see that we get back exactly the state $|\phi_j\rangle$.

To complete our description of the recovery operation $\mathcal{R}$ we need to give the unitary that can recover the original quantum state $|\psi\rangle$ given the state $|\phi_j\rangle$ and given the string $j$. By appending extra copies of the state $|+\rangle$, we can achieve the state:

$$\sum_{c\in C} b_c(-1)^{j\cdot c_K} Z_{c_{V\setminus K}+A_{cut}j}|+\ldots+\rangle_K|G'\rangle_{V\setminus K} \tag{43}$$

Now we an apply controlled phase operations (using definition 1) to transform the state to:

$$\sum_{c\in C} b_c(-1)^{j\cdot c_K} Z_{c_{V\setminus K}+A_{cut}j}|G\rangle_V \tag{44}$$

Now we can apply the unitary $Z_{A_{cut}j}$ to transform the state to:

$$\sum_{c\in C} b_c(-1)^{j\cdot c_K} Z_{c_{V\setminus K}}|G\rangle \tag{45}$$

Since

$$\langle G|Z_{c_{V\setminus K}}Z_{c'_{V\setminus K}}|G\rangle = 0 = \langle G|Z_cZ_{c'}|G\rangle \tag{46}$$

the unitary that sends $Z_{c_{V\setminus K}}|G\rangle \to Z_c|G\rangle$ is well defined. We can apply it to obtain:

$$\sum_{c\in C} b_c(-1)^{j\cdot c_K} Z_c|G\rangle \tag{47}$$

Finally we can apply a unitary that is diagonal in the $Z_c|G\rangle$ basis to remove the phase and recover $|\psi\rangle$ □

To leverage the previous result we have to show that the matrix $F$ will be full rank with high probability:

**Theorem 3.** *For any $0.33 < p < \frac{1}{2}$, let $R < 1-2p$. Also fix $q = \frac{w\log(n)}{n}$ for some constant $w > 1$. Let $H$ be a randomly sampled $(1-R)n \times n$ matrix where all $H_{ij}$ are distributed according to i.i.d Bern$(q)$ random variables. Denote the code with parity check matrix $H$ as $C$. Let $G$ be a randomly sampled graph where we begin with an empty graph and add each edge independently with probability $q$.*

*In analogy to the erasure channel, suppose each column is 'erased' with probability $p$. I.e. we start with the empty set $K \subseteq [n]$ and add each bit independently to $K$ with probability $p$. Let $F$ be the matrix defined in section 5.2 given the subset $K$. The probability that $F$ is not full rank satisfies:*

$$p_e = O\left(\frac{1}{n^{w-1}} + \frac{1}{n^{2(pw-1)}}\right) \tag{48}$$

*Proof.* Let $K$ denote the set of erased bits. By corollary 2, we can assume that $|K| = p'n \in [(p - \delta')n, (p + \delta')n]$ with probability exponentially close to 1 for any $\delta' > 0$. Since the erased bits are independent of the code, we can assume without loss of generality that $K$ consists of the first $p'n$ bits and randomly sample our code. Let us define the constant $\delta$ such that $R + \delta = 1 - 2p$. The matrix $F$ is as defined in the previous lemma:

$$F = \begin{array}{c} (R+p')n \left\updownarrow \right. \\ \end{array} \overset{\overset{(1-p')n}{\longleftrightarrow}}{\begin{bmatrix} C_{V\setminus K} \\ A_{cut}^T \end{bmatrix}} \tag{49}$$

where $C_{V\setminus K}$ are the generators of the code $C$ restricted to the non-erased bits, and $A_{cut}$ is the $|V \setminus K| \times |K|$ cut matrix across the cut $(K, V \setminus K)$. The matrix $F$ fails to to full rank if and only if one the the following events occurs:

1. $A_1$- The rows of $C_{V\setminus K}$ are linearly dependent
2. $A_2$- The rows of $A_{cut}^T$ are linearly dependent
3. $A_3$- A linear combination of the rows of $A_{cut}^T$ produce some word in $C_{V\setminus K}$

We will produce upper bounds on the probabilities of each of these events, and use the union bound to find an upper bound on the probability that $F$ is not full rank.

For $A_1$, we will argue using the randomly generated parity check matrix of the classical code. Recall that the parity check matrix $H$ is a $[(1 - R)n] \times n$ matrix such that each entry is 1 with probability $q$, and 0 otherwise. Denote the codewords $c \in C$ as $c = (a, b)$ where $a$ is supported on the erased bits and $b$ is supported on the non-erased bits. Further, let $\{(a_i, b_i)\}$ be a minimal set of generators for the code. Now observe the following equivalence: The set of vectors $\{b_i\}$ is linearly dependent if and only if there is some nonzero $c = (a, b) \in C$ with b=0. Observe that, if $H$ is full rank on the first $p'n$ bits, then there is no non-zero $c \in C$ such that $c = (a, b)$ with $b = 0$. Hence, the probability $\mathbb{P}(A_1)$ is less than or equal to the probability that the first $p'n$ columns of $H$ are linearly dependent.

The first $p'n$ columns of $H$ correspond to a random $[(1 - R)n] \times p'n = (2p + \delta)n \times p'n$ matrix where each entry is 1 with probability $q$. Note that, conditioned on the erased bits, we can treat this submatrix as independently sampled as in theorem 1. If $\delta'$ is small compared to $p$, then this submatrix has more rows than columns by some constant fraction of $n$. Hence, we derive the following upper bound using theorem 1

$$\mathbb{P}(A_1) = O\left(\frac{1}{n^{w-1}}\right) \tag{50}$$

We can bound $\mathbb{P}(A_2)$ directly using theorem 1. $A_{cut}^T$ is a randomly sampled $p'n \times (1 - p')n$ binary matrix where each entry is 1 independently with probability $\frac{w\log(n)}{n}$. Since $p < \frac{1}{2}$, there are more rows of $A_{cut}^T$ than there are columns by some constant fraction of $n$ if we take $\delta'$ small enough. Hence:

$$\mathbb{P}(A_2) = O\left(\frac{1}{n^{w-1}}\right) \tag{51}$$

Bounding $\mathbb{P}(A_3)$ requires the technical result proven in the appendix. Let $B_1$ be the event '$d_{C_{V\setminus K}} \leq \varepsilon' n$' where $\varepsilon' = H^{-1}(p)$. We write:

$$\mathbb{P}(A_3) = \mathbb{P}(A_3 \cap B_1) + \mathbb{P}(A_3 \cap \neg B_1) \tag{52}$$

Where we have used the negation symbol $\neg$ to indicate the complement. It is easy to check that we have met the conditions required for lemma 5 for small enough $\delta'$:

$$(1-p')H\left(\frac{H^{-1}(p)}{1-p'}\right) < 2p+\delta-p-\delta' \tag{53}$$

and

$$H(\varepsilon') = p < 2p+\delta \tag{54}$$

So, we can upper bound:

$$\mathbb{P}(A_3 \cap B_1) \leq \mathbb{P}(B_1) = O\left(\frac{1}{n^{(2p+\delta)w-2}}\right) = O\left(\frac{1}{n^{2(pw-1)}}\right) \tag{55}$$

Now we need to find an upper bound on $\mathbb{P}(A_3 \cap \neg B_1)$. Let $c_{V\setminus K}$ be some word in $C_{V\setminus K}$ of weight $g$, and let $v$ be any word in $\mathbb{F}_2^{p'n}$ of weight $k$. By corollary 1, we write:

$$\frac{b(k)}{2^{(1-p')n}} := \mathbb{P}\left(A_{cut}v = c_{V\setminus K}\right) = \frac{1}{2^{(1-p')n}}\left(1+\left(1-\frac{2w\log(n)}{n}\right)^k\right)^{(1-p')n-g}\left(1-\left(1-\frac{2w\log(n)}{n}\right)^k\right)^g \tag{56}$$

It is sufficient for an upper bound to analyze the word $c_{V\setminus K}$ of smallest weight, since clearly this expression is decreasing with $g$ increasing. Since $d_{C_{V\setminus K}} > \varepsilon' n$ we can assume $g = \varepsilon' n$. Let us define the intervals $I_1 = \left[1, \frac{zn}{\log(n)}\right]$ and $I_2 = \left[\frac{zn}{\log(n)}, p'n\right]$ for some large constant $z$ to be determined. For fixed $g$, we will first provide an upper bound for this function inside the interval $I_1$. Let us analyze the derivative of $b(k)$ with respect to $k$:

$$b'(k) = \left[1+\left(1-\frac{2w\log(n)}{n}\right)^k\right]^{(1-p')n-g-1}\left[1+\left(1-\frac{2w\log(n)}{n}\right)^k\right]^{g-1}\log\left(1-\frac{2w\log(n)}{n}\right)\times \tag{57}$$
$$\left\{-\left(1+\left(1-\frac{2w\log(n)}{n}\right)^k\right)g + ((1-p')n-g)\left(1-\left(1-\frac{2w\log(n)}{n}\right)^k\right)\right\}$$

We can set this expression equal to zero and solve. We obtain:

$$k^1_{max} = \frac{\log\left(\frac{(1-p')n-2g}{(1-p')n}\right)}{\log\left(1-\frac{2w\log(n)}{n}\right)} = \frac{\log\left(1-\frac{2\varepsilon'}{1-p'}\right)}{\log\left(1-\frac{2w\log(n)}{n}\right)} \tag{58}$$

The function $b'(k)$ is peaked around $k^1_{max}$, or $b'(k) \geq 0$ for $k \leq k^1_{max}$ and $b'(k) \leq 0$ for $k \geq k^1_{max}$. Expanding the expression for $k^1_{max}$, we can see that $k^1_{max} \in I_1$, so $b(k^1_{max})$ provides an upper bound for $b(k)$ in this interval. We calculate:

$$\frac{b(k^1_{max})}{2^{(1-p')n}} = \frac{1}{2^{(1-p')n}}\left[2-2\frac{\varepsilon'}{1-p'}\right]^{(1-p')n-\varepsilon' n}\left[\frac{2\varepsilon'}{1-p'}\right]^{\varepsilon' n} \tag{59}$$
$$= 2^{\wedge}\left[-\varepsilon' n + \log_2\left(1-\frac{\varepsilon'}{1-p'}\right)((1-p')n-\varepsilon' n) + \log_2\left(\frac{2\varepsilon'}{1-p'}\right)\varepsilon' n\right]$$

In the second interval $I_2$, the best upper bound we can obtain is:

$$\frac{b\left(\frac{zn}{\log(n)}\right)}{2^{(1-p')n}} \leq \frac{1}{2^{(1-p')n}}\left[1+\left(1-\frac{2w\log(n)}{n}\right)^{\frac{zn}{\log(n)}}\right]^{(1-p')n} \tag{60}$$

For large enough $n$,

$$\leq \frac{1}{2^{(1-p')n}}\left[1+e^{-2wz}\right]^{(1-p')n} =: \frac{b(k^2_{max})}{2^{(1-p')n}} \tag{61}$$

We are interested in finding an upper bound for the probability that any $v \in \mathbb{F}_2^{p'n}$ maps to any $c_{V\setminus K} \in C_{V\setminus K}$ under $A_{cut}$. For this we can employ the union bound:

$$\mathbb{P}(\exists v \in \mathbb{F}_2^{p'n}, \exists c_{V\setminus K} \in C_{V\setminus K} : A_{cut}v = c_{V\setminus K} \cap \neg B_1) \leq \sum_{\substack{v\in\mathbb{F}_2^{p'n} \\ c_{V\setminus K}\in C_{V\setminus K}}} \mathbb{P}[A_{cut}v = c_{V\setminus K} \cap \neg B_1] \tag{62}$$

Under our assumptions, $c_{V\setminus K} > \varepsilon' n$, we can further upper bound this expression by:

$$\leq 2^{Rn}\left(\sum_{|v|\in I_1}\frac{b(k^1_{max})}{2^{(1-p')n}} + \sum_{|v|\in I_2}\frac{b(k^2_{max})}{2^{(1-p')n}}\right) \tag{63}$$

where we used the fact that the code $C_{V\setminus K}$ has at most $2^{Rn}$ many words. We then note that there are at most $2^{o(n)}$ many terms in the first sum. The upper bound we obtain is:

$$\frac{2^{Rn}2^{o(n)}b(k^1_{max})}{2^{(1-p')n}} + \frac{2^{Rn}2^{p'n}}{2^{(1-p')n}}\left[1+e^{-2wz}\right]^{(1-p')n} \tag{64}$$

For large enough $z$ and small enough $\delta'$ the second term is exponentially small with $n$. The first term is exponentially small if:

$$R - \varepsilon' + \log_2\left(1-\frac{\varepsilon'}{1-p'}\right)\left((1-p')-\varepsilon'\right) + \log_2\left(\frac{2\varepsilon'}{1-p'}\right)\varepsilon' < 0 \tag{65}$$

If

$$g(p) := 1 - 2p - \varepsilon' + \log_2\left(1-\frac{\varepsilon'}{1-p}\right)\left((1-p)-\varepsilon'\right) + \log_2\left(\frac{2\varepsilon'}{1-p}\right)\varepsilon' < 0 \tag{66}$$

Then we can make $\delta'$ small enough that eq. (65) holds. We have found computationally $g(p) < 0$ for $0.33 < p < \frac{1}{2}$ (recall that we set $\varepsilon' = h^{-1}(p)$). □

**Corollary 3** (Quantum Erasure Channel)**.** *For any $0.33 < p < \frac{1}{2}$, let $R < 1 - 2p$. Let $q = \frac{w\log(n)}{n}$ for some constant $w > 1$. Let $H$ be a randomly sampled $(1-R)n \times n$ matrix where all $H_{ij}$ are distributed according to i.i.d. Bern$(q)$ random variables. Denote the code with parity check matrix $H$ as $C$. Let $G$ be a randomly sampled graph where we begin with an empty graph and add each edge independently with probability $q$.*

*The quantum code $(G, C)$ has vanishing probability of making a decoding error when sent over the quantum erasure channel $\mathcal{E}_p$. In particular, the probability of error satisfies:*

$$p_e = O\left(\frac{1}{n^{w-1}} + \frac{1}{n^{2(pw-1)}}\right) \tag{67}$$

### 5.3 Size Bounds on the code $(G, C)$

Now to provide a probabilistic estimate on the weight of the randomly chosen stabilizer code. For $i \in [1, \ldots (1-R)n]$, define the random variable $X_i$ to be the weight of the $i$th row of the parity check matrix $H$. By lemma 7, for each $i$

$$\mathbb{P}(X_i \geq \log^{1+\zeta}(n)) \leq \frac{\mathbb{E}(e^{t \cdot X_i})}{e^{t \cdot a}} = \frac{e^{n\left(\frac{w\log(n)}{n}\right)(e^t-1)}}{e^{t\log^{1+\zeta}(n)}} = \frac{n^{w(e^t-1)}}{n^{t\log^{\zeta}(n)}} \tag{68}$$

So, we can calculate via the union bound:

$$\mathbb{P}\left(\cup_i \{X_i \geq \log^{1+\zeta}(n)\}\right) \leq \frac{n e^{n\left(\frac{w\log(n)}{n}\right)(e^t-1)}}{e^{t\log^{1+\zeta}(n)}} = \frac{n^{w(e^t-1)+1}}{n^{t\log^{\zeta}(n)}} \tag{69}$$

Now for $i \in [1, 2, \ldots n]$ define the random variable $Y_i$ to be the number of neighbors of a vertex $i$ in the randomly generated graph $G$. The same analysis yields:

$$\mathbb{P}\left(\cup_i \{Y_i \geq \log^{1+\zeta}(n)\}\right) \leq \frac{n^{w(e^t-1)+1}}{n^{t\log^{\zeta}(n)}} \tag{70}$$

If all $X_i$ and $Y_i$ are less than $\log^{1+\zeta}(n)$, then the maximum weight of a generator of the stabilizer quantum code is upper bounded by $\log^{2+2\zeta}(n)$.

$$\mathbb{P}(\text{code } (G, C) \text{ is a } [[n, k, d, j]] \text{ code with } j \geq \log^{2+2\zeta}(n)) \leq \frac{2n^{w(e^t-1)+1}}{n^{t\log^{\zeta}(n)}} \tag{71}$$

If we take $t = O(1)$ we obtain vanishing probability with $n$ for any $\zeta > 0$.

## 6 Conclusion

In this paper we have given random constructions of quantum stabilizer codes that both achieve (come arbitrarily close to) the capacity of the erasure channel, while at the same time are polylog-LDPC with high probability. We stress that these codes are interesting primarily due to the work of Delfosse et al. [5], since such a property is impossible with codes that are LDPC.

We speculate that the graph states presented have interesting entanglement properties: in a sense they are "almost" absolutely maximally entangled states [13]. An absolutely maximally entangled state is a state that is maximally entangled across any partition of the subsystems. The graph states we have described have the following property: Choose any partition of the subsystems $(A, B)$ such that $A = \alpha n$ and $B = (1-\alpha)n$ where $\alpha < \frac{1}{2}$. The quantum graph state described is maximally entangled across the partition with all but inverse polynomial probability in $n$. This is easily seen from lemma 4, and theorem 1[3]. Hence, the quantum state is maximally entangled across almost all partitions where one set in the partition is larger than the other set by some constant fraction of the qubits. Indeed, we first became interested in these states because of this property.

[3]Note that if we add an edge with probability 1/2, then we could have instead used a much simpler version of theorem 1 to prove this. We required the more complicated result of Kolchin because we needed *log*-LDPC.

## 7 Acknowledgments

The authors wish to thank Lior Eldar and Murphy Y. Niu for helpful comments on the draft.

## References

[1] Charles H. Bennett, David P. DiVincenzo, and John A. Smolin. Capacities of quantum erasure channels. *Phys. Rev. Lett.*, 78:3217–3220, Apr 1997.

[2] Sergey Bravyi and Matthew B. Hastings. Homological product codes. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 273–282, New York, NY, USA, 2014. ACM.

[3] David Burshtein, Michael Krivelevich, Simon Litsyn, and Gadi Miller. Upper bounds on the rate of ldpc codes. *IEEE Transactions on Information Theory*, 48(9):2437–2449, 2002.

[4] A. Cross, G. Smith, J. A. Smolin, and B. Zeng. Codeword stabilized quantum codes. *IEEE Transactions on Information Theory*, 55(1):433–438, Jan 2009.

[5] Nicolas Delfosse and Gilles Zémor. Upper bounds on the rate of low density stabilizer codes for the quantum erasure channel. *Quantum Info. Comput.*, 13(9-10):793–826, September 2013.

[6] Robert G. Gallager. Low-density parity-check codes, 1963.

[7] M. Grassl, A. Klappenecker, and M. Rotteler. Graphs, quadratic forms, and quantum codes. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 45–, 2002.

[8] Sylvain Gravier, Jérôme Javelle, Mehdi Mhalla, and Simon Perdrix. Quantum Secret Sharing with Graph States. In Jaroslav Nešetřil Tomáš Vojnar David Antoš Antonín Kučera, Thomas A. Henzinger, editor, *MEMICS 2012 - International Doctoral Workshop on Mathematical and Engineering Methods in Computer Science*, volume 7721 of *Lecture Notes in Computer Science (LNCS)*, pages 15–31, Znojmo, Czech Republic, October 2012. Springer. Revised Selected Papers - http://www.memics.cz/2012/.

[9] Sylvain Gravier, Jrme Javelle, Mehdi Mhalla, and Simon Perdrix. On weak odd domination and graph-based quantum secret sharing. *Theoretical Computer Science*, 598:129 – 137, 2015.

[10] M. B. Hastings. Quantum Codes from High-Dimensional Manifolds. *ArXiv e-prints*, August 2016.

[11] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. Van den Nest, and H. . Briegel. Entanglement in Graph States and its Applications. *eprint arXiv:quant-ph/0602096*, February 2006.

[12] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Phys. Rev. A*, 69:062311, Jun 2004.

[13] W. Helwig. *Multipartite Entanglement: Transformations, Quantum Secret Sharing, Quantum Error Correction*. PhD thesis, University of Toronto, 2014.

[14] V. F. Kolchin. *Random Graphs (Encyclopedia of Mathematics and its Applications)*. Cambridge University Press, 1998.

[15] Alexey A. Kovalev, Ilya Dumer, and Leonid P. Pryadko. Design of additive quantum codes via the code-word-stabilized framework. *Phys. Rev. A*, 84:062319, Dec 2011.

[16] Santhosh Kumar, Robert Calderbank, and Henry D Pfister. Reed-muller codes achieve capacity on the quantum erasure channel. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 1750–1754. IEEE, 2016.

[17] Santhosh Kumar and Henry D Pfister. Reed-muller codes achieve capacity on erasure channels. *arXiv preprint arXiv:1505.05123*, 2015.

[18] Michael G Luby, Michael Mitzenmacher, M Amin Shokrollahi, Daniel A Spielman, and Volker Stemann. Practical loss-resilient codes. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 150–159. ACM, 1997.

[19] Damian Markham and Barry C Sanders. Graph states for quantum secret sharing. *Physical Review A*, 78(4):042309, 2008.

[20] C. Morgan and A. Winter. "pretty strong" converse for the quantum capacity of degradable channels. *IEEE Transactions on Information Theory*, 60(1):317–333, Jan 2014.

[21] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

[22] D. Schlingemann and R. F. Werner. Quantum error-correcting codes associated with graphs. *Phys. Rev. A*, 65:012308, Dec 2001.

[23] Dirk Schlingemann and Reinhard F Werner. Quantum error-correcting codes associated with graphs. *Physical Review A*, 65(1):012308, 2001.

[24] J. P. Tillich and G. Zemor. Quantum ldpc codes with positive rate and minimum distance proportional to the square root of the blocklength. *IEEE Transactions on Information Theory*, 60(2):1193–1202, Feb 2014.

[25] Mark M Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, 2014.

# A Appendix

Before starting the proofs we present a few simple definitions and facts. The Gamma function is defined as:

$$\Gamma(y) := \int_0^\infty x^{y-1} e^{-x} dx \tag{72}$$

The Digamma functions is defined as:

$$\psi^{(0)}(y) := \frac{\Gamma'(y)}{\Gamma(y)} \tag{73}$$

and the $m$th order Polygamma function is defined as:

$$\psi^{(m)}(y) := \frac{d^m}{dy^m}\psi^{(0)}(y) \tag{74}$$

The Harmonic numbers are defined as:

$$H_{x-1} := \psi^{(0)}(x) + \gamma \qquad H^{(2)}_{x-1} := \frac{\pi^2}{6} - \psi^{(1)}(x) \tag{75}$$

where $\gamma$ is the Euler Mascheroni constant. At integer values, the Harmonic numbers have their usual expression:

$$H_k = \sum_{j=1}^{k} \frac{1}{j} \qquad H^{(2)}_k = \sum_{j=1}^{k} \frac{1}{j^2} \tag{76}$$

There are many important properties of these definitions, we list the properties we will need for the proofs:

**Fact 1.** *1. The Gamma function is equal to the factorial at positive integer arguments:*

$$\forall j \in \mathbb{Z},\ > 0 :\ \Gamma(j-1) = j! \tag{77}$$

*2. We can approximate $\psi^{(1)}(k)$ as $\Theta(1/k)$:*

$$\forall k > 0 :\quad \frac{1}{k} \leq \psi^{(1)}(k) \leq \frac{2}{k} \tag{78}$$

*which implies by definition:*

$$\frac{\pi^2}{6} - \frac{2}{k} \leq H^{(2)}_{k-1} \leq \frac{\pi^2}{6} - \frac{1}{k} \tag{79}$$

*3. For positive $k$:*

$$\gamma + \log(k) < H_k < \gamma + \log(k+1) \tag{80}$$

Throughout the appendix we will have some $n$ and $w$ in mind. We will denote:

$$a := 1 - \frac{2w\log(n)}{n} \tag{81}$$

The goal of the appendix is to show that codes from our ensemble have linear distance after the erasure channel. We find it useful to first prove that the code itself has linear distance with high probability:

**Lemma 9.** *Let $H$ be a $\alpha n \times n$ binary matrix with $\alpha < 1$, where each entry is chosen uniformly at random according to i.i.d. Bern$(q)$ where $q = \frac{w\log(n)}{n}$. Let $C$ be the code with $H$ as its parity check matrix. Further, let $d_C$ be the distance of the code $C$. For any constant $\varepsilon > 0$ satisfying $h(\varepsilon) < \alpha$*

$$d_C > \varepsilon n \tag{82}$$

*with probability at least:*

$$1 - O\left(\frac{1}{n^{w\alpha-2}}\right) \tag{83}$$

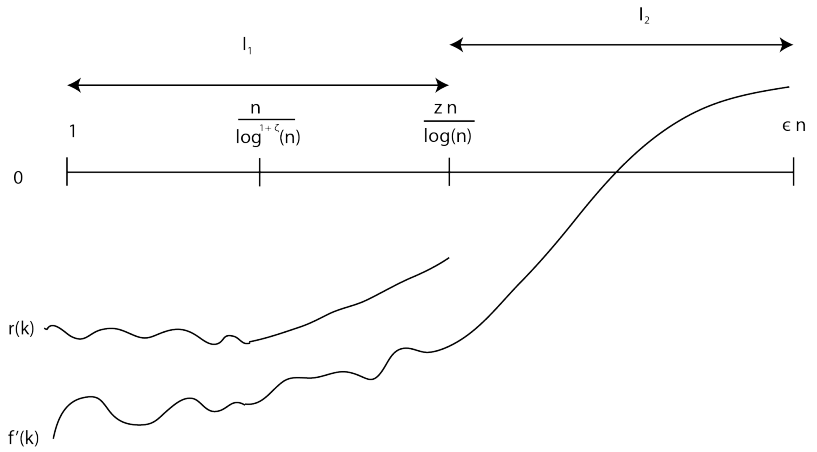


Figure 1: An illustration of our proof method. We demonstrate that $f'$ has the above form, which implies exactly one local minimum. Therefore, $f$ must be maximized at one of the endpoints (either at $k = 1$ or at $k = \varepsilon n$.

*Proof.* We use first moment methods. Let $X$ be a random variable equal to the number of subsets of columns of $H$ with size at most $\varepsilon n$ that sum to zero. $X$ can equivalently be defined as the number of words of $C$ with weight less than $\varepsilon n$. We calculate using corollary 1

$$\mathbb{E}(X) = \sum_{k=1}^{\varepsilon n} \binom{n}{k} \left(\frac{1+a^k}{2}\right)^{\alpha n} \tag{84}$$

Let us define the function:

$$f(k) := \binom{n}{k} \left(1+a^k\right)^{\alpha n} \tag{85}$$

We can make this function continuous and differientiable by substituting Gamma functions for factorials:

$$\binom{n}{k} = \frac{\Gamma(n-1)}{\Gamma(k-1)\Gamma(n-k-1)} \tag{86}$$

We will show that we can find an upper bound for $f(k)$ by examining the endpoints $k = 1$ and $k = \varepsilon n$. Define two intervals $I_1 = \left[1, \frac{zn}{\log(n)}\right]$ and $I_2 = \left[\frac{zn}{\log(n)}, \varepsilon n\right]$ where $z$ is some large constant to be determined. The property we claim follows if we can show that $f'(k) < 0$ in the interval $I_1$ and that the function $f'(k)$ has exactly one zero in the interval $I_2$. The remainder of the proof will fall into two parts. In part a we will demonstrate that $f'(k) < 0$ in $I_1$ and in part b we will demonstrate that $f'(k)$ has exactly one zero in $I_2$.

## Part a

We will further divide interval $I_1$ into two other intervals: $\left[1, \frac{n}{\log^{1+\zeta}(n)}\right]$ and $\left[\frac{n}{\log^{1+\zeta}(n)}, \frac{zn}{\log(n)}\right]$ where $\zeta$ is some small positive constant. We will provide another function $r(k)$ which upper bounds $f'$ in both of these intervals. We first demonstrate that $r(k) < 0$ in the interval $\left[1, \frac{n}{\log^{1+\zeta}(n)}\right]$. Then we will show that $r'(k)$ has a positive slope in the interval $\left[\frac{n}{\log^{1+\zeta}(n)}, \frac{zn}{\log(n)}\right]$ and has a negative endpoint at $\frac{zn}{\log(n)}$, implying $f'(k) < 0$ throughout $I_1$.

We calculate:

$$f'(k) = v(k)\left(a^k\left(\alpha n \log\left(a\right) + H_{n-k} - H_k\right) + H_{n-k} - H_k\right) \tag{87}$$

where $v(k) > 0$ We are interested in the sign of $f'(k)$, so it is sufficient to study $\frac{f'(k)}{v(k)}$. Define:

$$b(k) := a^k\left(\alpha n \log\left(a\right) + H_{n-k} - H_k\right) + H_{n-k} - H_k \tag{88}$$

For large enough $n$ we can expand to show:

$$\frac{-3w\log(n)}{n} \leq \log\left(a\right) \leq -\frac{2wlog(n)}{n} \tag{89}$$

So,

$$b(k) \leq a^k\left(H_{n-k} - H_k - 2w\alpha\log(n)\right) + H_{n-k} - H_k \tag{90}$$

By fact 1:

$$H_{n-k} \leq \log(n) + \gamma \tag{91}$$

so we have $b(k) \leq r(k)$ where:

$$r(k) := a^k\left(-2w\alpha + 2\right)\log(n) + H_{n-k} - H_k \tag{92}$$

We need to show that the function $r(k) < 0$ for all $k$ in the interval $\left[1, z\frac{n}{\log(n)}\right]$.

For any $k \in \left[1, \frac{n}{\log^{1+\zeta}(n)}\right]$ the function $r(k)$ is negative for sufficiently large $n$. Indeed for $k = \frac{n}{\log^{1+\zeta}(n)}$:

$$\lim_{n\to\infty} a^k = \lim_{n\to\infty}\left(1 - \frac{2w\log(n)}{n}\right)^{\frac{n}{\log^{1+\zeta}(n)}} = 1 \tag{93}$$

The Harmonic terms are of order $\log(n)$ at most, so the first term dominates if $w\alpha > 2$.

Now we will show that the term $r'(k) \geq 0$ for all $k$ in the interval $\left[\frac{n}{\log^{1+\zeta}(n)}, z\frac{n}{\log(n)}\right]$. We calculate:

$$r'(k) = a^k(-2\alpha w + 2)\log(n)\log\left(a\right) + H^{(2)}_{n-k} + H^{(2)}_k - \frac{\pi^2}{3} \tag{94}$$

By eq. (89):

$$r'(k) \geq \frac{4w^2\alpha\log^2(n)}{n}\left(a\right)^k + H^{(2)}_{n-k} + H^{(2)}_k - \frac{\pi^2}{3} \tag{95}$$

Now we can apply fact 1, and the fact that $\frac{1}{n-k} \leq \frac{1}{k}$ to obtain:

$$r'(k) \geq \frac{4w^2\alpha\log^2(n)}{n}\left(1 - \frac{2w\log(n)}{n}\right)^k - \frac{4}{k} \tag{96}$$

By definition of our interval, we obtain:

$$k \geq \frac{n}{\log^{1+\zeta}(n)} \Rightarrow -\frac{1}{k} \geq -\frac{\log^{1+\zeta}(n)}{n} \tag{97}$$

For large enough $n$, we have:

$$k \leq z\frac{n}{\log(n)} \Rightarrow \left(1 - \frac{2w\log(n)}{n}\right)^k \geq e^{-2wz-1} \tag{98}$$

Hence,

$$r'(k) \geq \frac{4w^2\alpha e^{-2wz-1}\log^2(n)}{n} - \frac{4\log^{1+\zeta}(n)}{n} > 0 \tag{99}$$

Once we demonstrate that $r\left(\frac{zn}{\log(n)}\right) < 0$, we will have shown that $r(k) < 0$ for all $k \in \left[1, \frac{zn}{\log(n)}\right]$. Indeed:

$$r\left(\frac{zn}{\log(n)}\right) = (a)^{\frac{zn}{\log(n)}}\left(-2w\alpha + 2\right)\log(n) + H_{n-\frac{zn}{\log(n)}} - H_{\frac{zn}{\log(n)}} \tag{100}$$

It is not hard to see that:

$$H_{n-\frac{zn}{\log(n)}} - H_{\frac{zn}{\log(n)}} = O(\log(\log(n))) \tag{101}$$

from fact 1. In addition, the term:

$$\left(1 - \frac{2w\log(n)}{n}\right)^{\frac{zn}{\log(n)}} = \Theta(1) \tag{102}$$

So, for large enough $n$, $r\left(\frac{zn}{\log(n)}\right) < 0$.

We have shown that $r(k) < 0$ for all $k \in \left[1, \frac{zn}{\log(n)}\right]$. This implies $b(k) < 0$ for all $k \in I_1$ which in turn implies that $f'(k) < 0$ for these $k$.

## Part b

Now we will show that inside the interval $k \in I_2$ the function $f'(k)$ has exactly one zero. By rearranging, we can see that $f'(k) = 0$ if and only if

$$\left(H_{n-k} - H_k\right)\left(\frac{1}{a^k} + 1\right) = -\alpha n\log\left(a\right) \tag{103}$$

We define:

$$g(k) := \left(H_{n-k} - H_k\right)\left(\frac{1}{a^k} + 1\right) \tag{104}$$

and calculate:

$$g'(k) = \frac{1}{a^k}\left(\log(a)(H_k - H_{n-k}) - (1 + a^k)(\psi^{(1)}(n-k+1) + \psi^{(1)}(k+1))\right) \tag{105}$$

So we can lower bound:

$$\begin{aligned} a^k g'(k) &\geq \frac{2w\log(n)}{n}\left(H_{n-k} - H_k\right) - \frac{4}{k} \\ &\geq \frac{2w\log(n)}{n}(H_{n-k} - H_k) - \frac{4\log(n)}{zn} > 0 \end{aligned} \tag{106}$$

which is positive for $z$ large enough since $H_{n-k} - H_k = \Omega(1)$ in this interval. Since $g(k)$ is strictly increasing with $k$ and the RHS of eq. (103) is fixed, there can be at most one place where $f'(k) = 0$. We have already shown

$$f'\left(\frac{zn}{\log(n)}\right) < 0 \tag{107}$$

and it is not hard to see:

$$f'(\varepsilon n) > 0 \tag{108}$$

for large enough $n$ so $f'$ has exactly one zero in this interval.

Now back to the problem at hand, we want an upper bound on $\mathbb{E}(X)$. Our analysis implies that we can use the largest endpoint as an upper bound on $f(k)$:

$$\begin{aligned}\mathbb{E}(X) &\leq \max\left\{\varepsilon n\binom{n}{1}\left(1 - \frac{2w\log(n)}{n}\right)^{\alpha n}, \varepsilon n\binom{n}{\varepsilon n}\left(\frac{1+a^k}{2}\right)^{\alpha n}\right\} \\ &\leq \max\left\{\frac{1}{n^{w\alpha-2}}, 2^{o(n)+(h(\varepsilon)-\alpha)n}\right\}\end{aligned} \tag{109}$$

By hypothesis the second term is exponentially small. To complete the proof we use Markov's inequality. Let us define the event $A_1$ as the event '$d_C \leq \varepsilon n$'. Then, we have:

$$\mathbb{P}(A_1) = \mathbb{P}(X \geq 1) \leq \mathbb{E}(X) = O\left(\frac{1}{n^{w\alpha-2}}\right) \tag{110}$$

□

The result we are actually interested involves the code $C$ after the erasure channel. We need to show that this code still has linear distance. Suppose as in the paper that the set of bits $K$ has been erased, and say that the size of this set is $\beta n$ for some constant $\beta$. We establish the following lemma:

**Lemma 5** (Linear distance)**.** *Let $H$ be a $\alpha n \times n$ binary matrix with $\alpha < 1$, where each entry is chosen uniformly at random according to i.i.d. Bern$(q)$ where $q = \frac{w\log(n)}{n}$. Let $C$ be the code with $H$ as its parity check matrix. Further, let $d_C$ be the distance of the code $C$.*

*Let $K$ be the first $\beta n$ bits (corresponding to the first $\beta n$ columns of $H$) and assume that $\alpha > \beta$. Denote the code $C$ restricted to the bits $V \setminus K$ as $C_{V\setminus K}$. If $\varepsilon'$ satisfies:*

$$(1-\beta)h\left(\frac{\varepsilon'}{1-\beta}\right) < \alpha - \beta \quad \text{and} \quad h(\varepsilon') < \alpha \tag{26}$$

*then,*

$$d_{C_{V\setminus K}} > \varepsilon' n \tag{27}$$

*with probability at least*

$$1 - O\left(\frac{1}{n^{w\alpha-2}}\right) \tag{28}$$

*Proof.* We will use the previous lemma to show that we can expect the code $C$ to have linear distance, and condition on this event to show that the code $C_{V\setminus K}$ satisfies the same property with a weaker constant.

Again let $A_1$ be the event that $d_C \leq \varepsilon n$ where $\varepsilon > \varepsilon'$ and $H(\varepsilon) < \alpha$. Further, let $A_2$ be the event that $d_{C_{V\setminus K}} \leq \varepsilon' n$. Just as in the paper, we will use the negation symbol $\neg$ for the complement. So $\neg A_1$ is the event that $d_C > \varepsilon n$. We can write:

$$\mathbb{P}(A_2) = \mathbb{P}(A_2 \cap A_1) + \mathbb{P}(A_2 \cap \neg A_1) \tag{111}$$

By hypothesis $H(\varepsilon) < \alpha$, so by lemma 9, we can upper bound:

$$\mathbb{P}(A_2 \cap A_1) \leq \mathbb{P}(A_1) = O\left(\frac{1}{n^{w\alpha-2}}\right) \tag{112}$$

Recall from the previous lemma that the code $C$ is defined through the parity check matrix $H$ as the set of all subsets of the columns of $H$ which sum to zero. A 'bad' event in the current context is the existence of a set of columns which simultaneously sum to zero and has small weight outside the set $K$. Such an event implies the existence of a codeword which is "nearly covered up" by the erasure. We proceed by bounding the probability that such a set exists.

Let $s \subset [n]$ be some subset of the columns containing fewer than $\varepsilon' n$ many columns outside the erased set $K$, and let $Q$ be the class of all sets with this property. Let us define the event $B_s$ as 'the sum of the columns in the set $s$ is zero' (or equivalently that the membership vector of the set $s$ forms a word in the code). It is easy to see that:

$$A_2 \Rightarrow \bigcup_{s\in Q} B_s \tag{113}$$

so we have immediately that:

$$A_2 \cap \neg A_1 \Rightarrow \left(\bigcup_{s\in Q} B_s\right) \cap \neg A_1 \tag{114}$$

which implies the upper bound:

$$\mathbb{P}(A_2 \cap \neg A_1) \leq \mathbb{P}\left(\left(\bigcup_{s\in Q} B_s\right) \cap \neg A_1\right) \tag{115}$$

Now we will compute an upper bound on $\mathbb{P}((\cup_{s\in Q} B_s) \cap \neg A_1)$. Let $s$ be some subset with $l_1$ many elements in $K$ and $l_2$ many elements in $V \setminus K$.

By corollary 1, the probability is the same as before:

$$\mathbb{P}(B_s) = \left(\frac{1 + \left(1 - \frac{2w\log(n)}{n}\right)^{l_1+l_2}}{2}\right)^{\alpha n} \tag{116}$$

except that under our assumptions (namely that we are in the case $\neg A_1$) we have,

$$1 \leq l_2 \leq \varepsilon' n \tag{117}$$

and

$$(\varepsilon - \varepsilon')n \leq l_1 \tag{118}$$

so we have:

$$\mathbb{P}\left(B_s \cap \neg A_1\right) \leq \frac{1}{2^{\alpha n}} \left(1 + \left(1 - \frac{2w\log(n)}{n}\right)^{(\varepsilon-\varepsilon')n}\right)^{\alpha n} \leq \frac{1}{2^{\alpha n}} \left(1 + \frac{1}{n^{2w(\varepsilon-\varepsilon')}}\right)^{\alpha n} \tag{119}$$

Where we assumed $n$ was very large for the final inequality. Using the union bound, we can then argue:

$$\begin{aligned} \mathbb{P}\left(\left(\bigcup_{s\in Q} B_s\right) \cap \neg A_1\right) &\leq \binom{(1-\beta)n}{\varepsilon' n} \sum_{k=(\varepsilon-\varepsilon')n}^{\beta n} \binom{\beta n}{k} \frac{1}{2^{\alpha n}} \left(1 + \frac{1}{n^{2w(\varepsilon-\varepsilon')}}\right)^{\alpha n} \\ &\leq \frac{2^{o(n)+H\left(\frac{\varepsilon'}{1-\beta}\right)(1-\beta)n}}{2^{(\alpha-\beta)n}} \end{aligned} \tag{120}$$

In the last step we used the standard approximation to binomial coefficients:

$$\binom{m}{\delta m} = 2^{o(m)} 2^{H(\delta)m} \tag{121}$$

We obtain an upper bound that is exponentially small with $n$ if

$$H\left(\frac{\varepsilon'}{1-\beta}\right)(1-\beta) < \alpha - \beta \tag{122}$$

□