An Overview of

Optimal Single-Shot Decoding of Quantum Codes

A. Cumitini, S. Tinelli, B. Matuz, F. Lazaro, L. Barletta

Presented by Borishan Ghosh

Consider a faulty channel that communicates a bit with success probability (1 - ϵ)

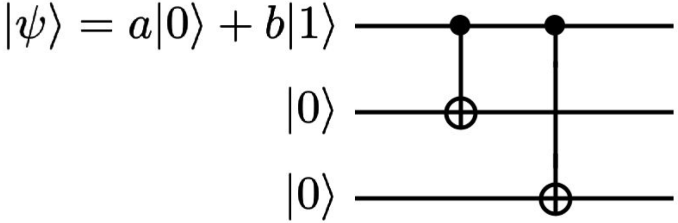
$$\mathrm{BSC}(\epsilon) = \begin{cases} b, \ p = 1 - \epsilon \\ \neg b, \ p = \epsilon \end{cases} \qquad b \in \{0, 1\}$$

A simple way to counteract the error could be to make 3 copies of the given data and finally correct for the noise in circuit by taking the maximum of the redundant copies.

$$\bar{0} = 000, \quad \bar{1} = 111$$

The repeated (i,e. coded) digits are called *codewords*, and a notion of distance between strings can be decided by the number of locations where they differ, this is the *Hamming Distance*.

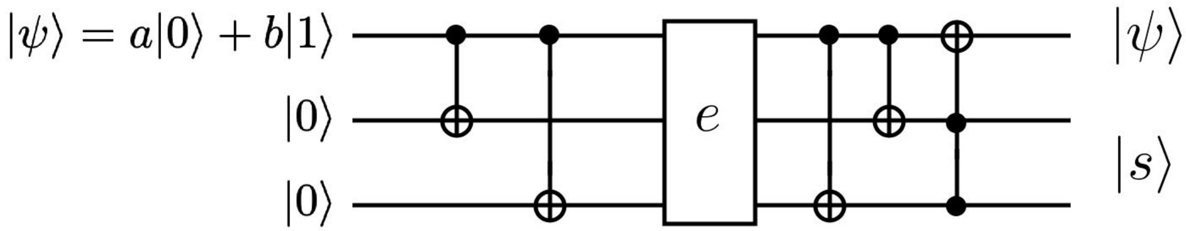
Because of the *No-Cloning Theorem* we cannot take a similar approach with qubits, instead we may choose to encode them as so,



$$a|000\rangle + b|111\rangle$$

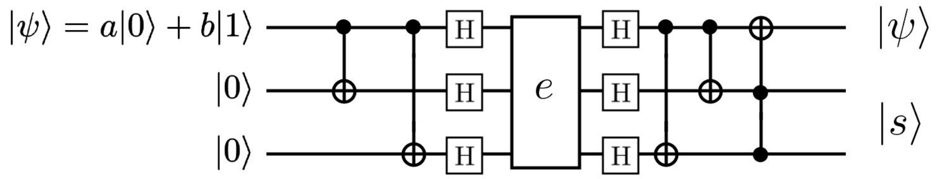
Decoherence may occur in many forms. We first consider the case of a simple bit-flip, or an X gate applied at random to one of the qubits.

The following circuit corrects for a random bit-flip, and returns the original qubit alongside the syndome bits corresponding to the bit-flip.



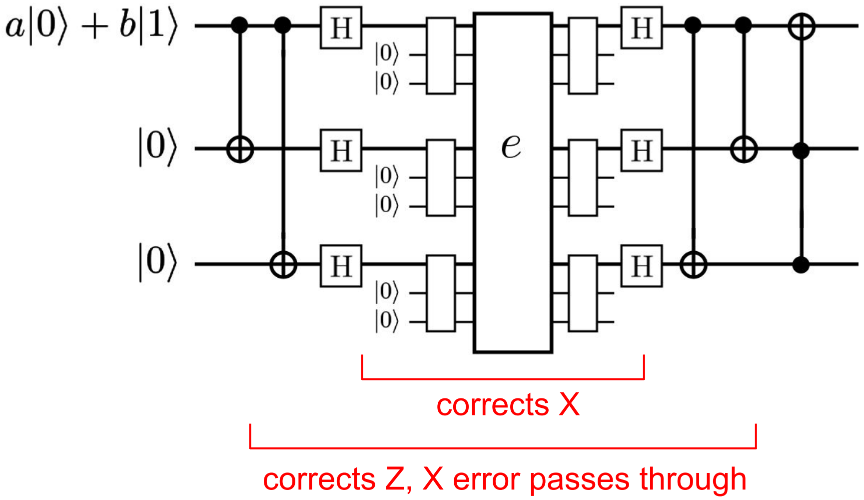
However, phase-flips are unaffected by the circuit and pass through, we can accommodate for the same by padding the error channel with Hadamard gates. These have the effect of changing the Z-basis flips to X-basis.

$$HZH = X$$



The two error correcting circuits can be now concatenated to correct for both X and Z random flips.

This is *Shor's 9-bit Error Correcting Code.*



To recap,

1. We introduce redundancy to counteract noise
2. Measure the syndrome unique to each bit/phase flip
3. Apply corrections (X / Z gates) correspondingly

Just correcting for X and Z gates also solves for Y = iXZ, where i is the imaginary phase that is irrelevant during measurement.

Since any unitary transformation can be thus decomposed into Pauli matrices, **this circuit corrects for any arbitrary single bit error.**

$$U = \alpha I + \beta X + \gamma Y + \delta Z$$

However there is an error associated with each syndrome measurement as well. We must look for fault tolerant circuits or other clever ways of implementing the same

A [n,k] error correcting code is one that codes k-bit binary strings to n-bit strings. The *Hamming weight* (denoted d in the figure) of the n-dimensional subspace is the minimum hamming distance between 2 codewords.

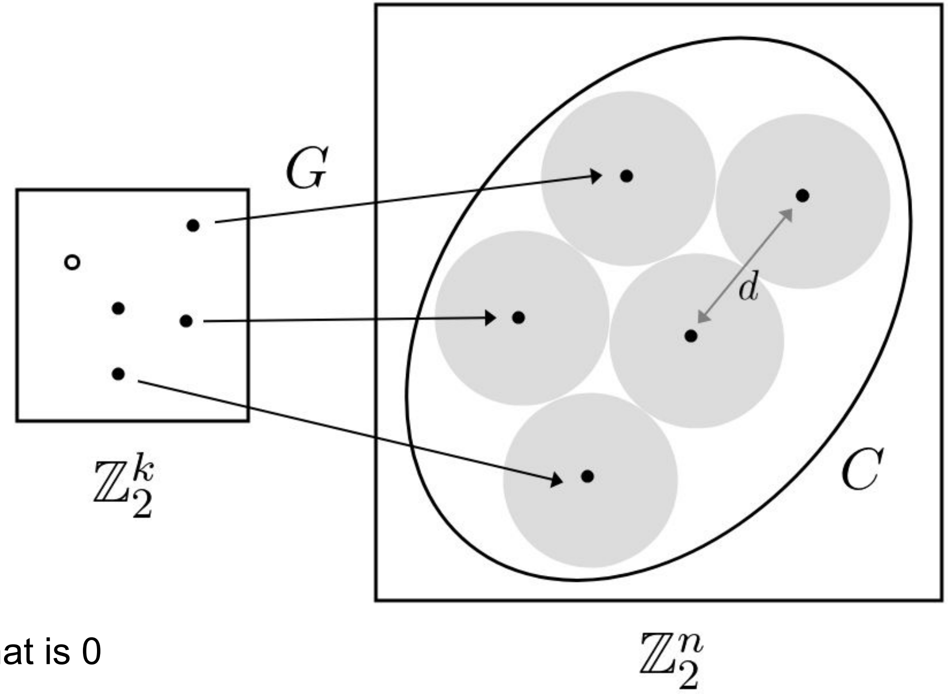
$$d = \min_{a,b \in C} \Delta(a, b), \quad C \subseteq \mathbb{Z}_2^n$$

The generator G is a n × k matrix such that

$$C = \{Ga \mid a \in \{0,1\}^k\}$$

The parity check matrix H is a (n-k) × n matrix that is 0 for every valid codeword.

$$C = \{b \in \mathbb{Z}_2^n \mid Hb = 0\}$$



The subspace C is then the **nullspace of the parity check matrix H**, and **the range of matrix G.**

The dual of C is written,

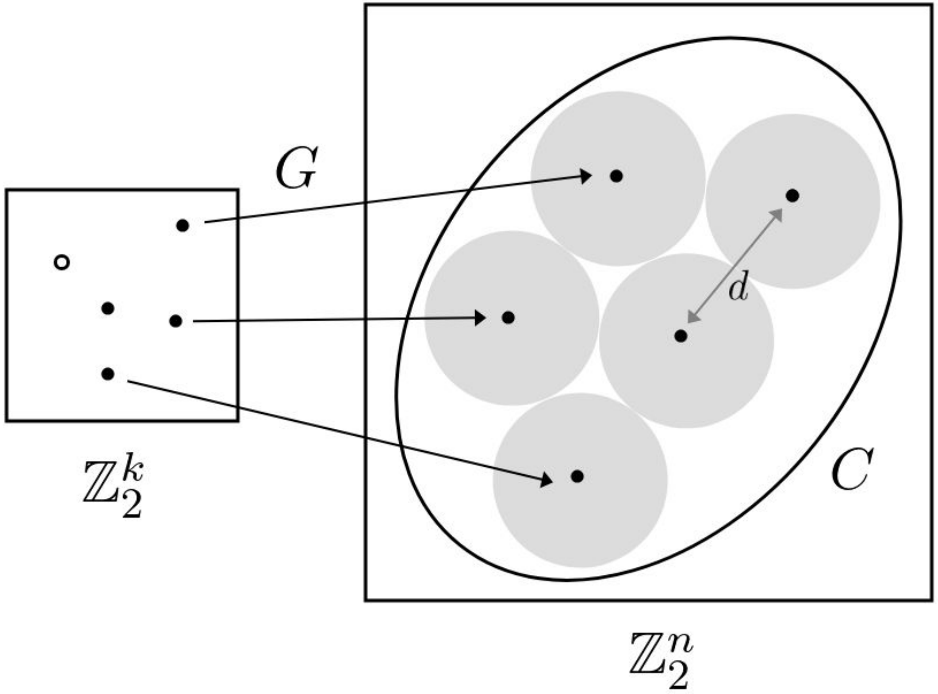
$$C^{\perp} = \{b \in \mathbb{Z}_2^n \ | b \cdot c = 0 \ \forall c \in C\}$$

Where the generator of the dual is K^T
and the parity matrix G^T

Here C can correct for at most t errors,

$$t = \lfloor \frac{d-1}{2} \rfloor$$

with its rate equal to n/k.



CSS Codes

Given, C_1 a linear code $[n, k_1]$ and C_2 a $[n, k_2]$ linear code such that $C_2 \subseteq C_1$
Where each code can correct upto t errors, we can define a $[n, k_1 - k_2]$ code $CSS(C_1, C_2)$ that is also resilient upto t qubit errors.

Take for example the (7,4) and (7,3) Hamming codes, we can construct a (7, 1) quantum code with $2^{k_1-k_2} = 2^1 = 2$ codewords, where the codewords $x_0, \ldots x_{2^{k_1-k_2}-1} \in C_1$ are picked so that,

$$x_i + x_j \notin C_2, \quad x_i \neq x_j$$

The $k_1 - k_2$ logical bits chosen are thus identified as

$$|j\rangle \mapsto |x_j + C_2\rangle = \frac{1}{|C_2|} \sum_{y \in C_2} |x_j + y\rangle$$

CSS Codes

Given, C_1 a linear code $[n, k_1]$ and C_2 a $[n, k_2]$ linear code such that $C_2 \subseteq C_1$
Where each code can correct upto t errors, we can define a $[n, k_1 - k_2]$ code $CSS(C_1, C_2)$ that is also resilient upto t qubit errors.

In our example,

$$|0\rangle_L = \frac{1}{\sqrt{8}}(|0000000\rangle + |0001111\rangle + |0110011\rangle + |1010101\rangle \\ + |0111100\rangle + |1011010\rangle + |1100110\rangle + |1101001\rangle)$$

$$|1\rangle_L = \frac{1}{\sqrt{8}}(|1111111\rangle + |1110000\rangle + |1001100\rangle + |0101010\rangle \\ + |1000011\rangle + |0100101\rangle + |0011001i\rangle + |0010110\rangle)$$

This is the 7-bit Steane Code.

CSS Codes

Given, C_1 a linear code $[n, k_1]$ and C_2 a $[n, k_2]$ linear code such that $C_2 \subseteq C_1$ Where each code can correct upto t errors, we can define a $[n, k_1 - k_2]$ code $CSS(C_1, C_2)$ that is also resilient upto t qubit errors.

The error correction routine is as follows,

1. Create a state $|y\rangle|00\ldots0\rangle$ and compute its syndrome according to code C_1 make a measurement on the syndrome and correct for the bit-flips.
2. Apply a Hadamard Gate to each qubit, compute its syndrome according to code $C_2^{\perp}$ measure the syndrome and correct for bit-flips (the hadamard gate converts the phase errors to bit errors)
3. Apply another Hadamard Gate to recover the original state.

CSS Codes

Given, C_1 a linear code $[n, k_1]$ and C_2 a $[n, k_2]$ linear code such that $C_2 \subseteq C_1$
Where each code can correct upto t errors, we can define a $[n, k_1 - k_2]$ code $CSS(C_1, C_2)$ that is also resilient upto t qubit errors.

The error correction routine is as follows,

1. Create a state $|y\rangle|00\ldots0\rangle$ and compute its syndrome according to code C_1 make a measurement on the syndrome and correct for the bit-flips.
2. Apply a Hadamard Gate to each qubit, compute its syndrome according to code $C_2^{\perp}$ measure the syndrome and correct for bit-flips (the hadamard gate converts the phase errors to bit errors)
3. Apply another Hadamard Gate to recover the original state.

Syndrome measurements themselves may not be fault-tolerant

Which leaves a few options

1. Introduce ancilla bits and repeat syndrome measurements (Shor's Syndrome Extraction), scales linearly with code distance.
2. Carry out redundant syndrome measurements (single-shot error correction), constant number of measurement rounds.

We proceed with the same set-up as before, over a $[n_q, k_q]$ code, with a $(n_q - k_q) \times 2n_q$ check matrix,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_X & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_Z \end{bmatrix}$$

Where the error term $e = [e_X | e_Z]$ is a $2n_q$ bit string. Where the i-th element of the X-component is set to 1 for a bit-flip on the i-th position, likewise for the Z component. Consequently the syndrome is given as $s = eH^T$

Syndrome Error Probability

We can model the faulty measurement with a Binary Symmetric Channel.
Each row $\{h_1, h2, \ldots, h_{n-k}\}$ of the check matrix H corresponds to an ancillary bit that is inserted and interacting with $w(h_i)$ bits. Where, $w(h_i)$ is the hamming weight of the column. Supposing that each interaction fails independently with a probability q, the failure probability is,

$$\Pr(z_j \neq \tilde{z_j}) = \textstyle\sum_{i \text{ odd}} \displaystyle\binom{w(h_j)}{i} q^i (1-q)^{w(h_j)-i}$$

Where z_j is the redundant syndrome and $\tilde{z}_j$ the observed syndrome. The averaged error over each stabilizer is,

$$\delta = \frac{\sum_{i=1}^{m} \Pr(z_j \neq \tilde{z_j})}{m}$$

Code Construction

Before measurement the syndrome is coded with a Generator matrix,
We thus obtain the redundant syndrome,

$$z = sG_s = eH^T G_s = eH_o^T$$

The rank of the check matrix H_o is bounded to be (n-k). The size of H_o being m×n, and m > n-k implies it is overcomplete.

$$H_o^T = [H|P]$$

Where P can be found through gaussian elimination, similarly we can solve for G after deducing A from the following,

$$H^T G_s = H_o^T \implies H^T [I|A] = [H|P]^T \implies H^T A = P^T$$

Decoding

Two errors are degenerate if their modulo 2 sum is a stabilizer, when this happens, the coset thus generated are errors that can be corrected by a single unique recovery operator. Given the observed syndrome we can thus find the most probable coset using a Maximum A Posteriori decoder.

$$\hat{\mathcal{E}} = \arg\max_{\mathcal{E}} \Pr(\mathcal{E}|\tilde{z}) = \arg\max_{\mathcal{E}} \Pr(\tilde{z}|\mathcal{E})\Pr(\mathcal{E})$$

From Bayes' Rule

$$\Pr(\tilde{z}|\mathcal{E}) = \frac{\Pr(\tilde{z} \wedge \mathcal{E})}{\Pr(\mathcal{E})} = \frac{1}{\Pr(\mathcal{E})} \sum_{e \in \mathcal{E}} \Pr(\tilde{z} \wedge e)$$

$$= \frac{1}{\Pr(\mathcal{E})} \sum_{e \in \mathcal{E}} \Pr(\tilde{z}|e)\Pr(e)$$

Assuming independence of individual errors

Decoding

Two errors are degenerate if their modulo 2 sum is a stabilizer, when this happens, the coset thus generated are errors that can be corrected by a single unique recovery operator. Given the observed syndrome we can thus find the most probable coset using a Maximum A Posteriori decoder.

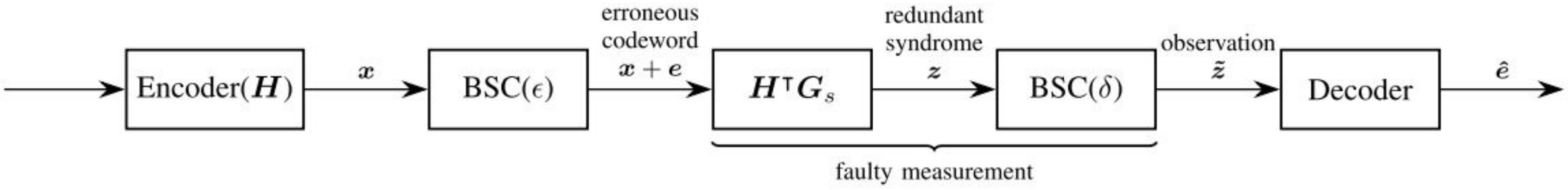
$$\hat{\mathcal{E}} = \arg\max_{\mathcal{E}} \Pr(\mathcal{E}|\tilde{z}) = \arg\max_{\mathcal{E}} \Pr(\tilde{z}|\mathcal{E})\Pr(\mathcal{E})$$

$$\Pr(\tilde{z}|\mathcal{E}) = \frac{\Pr(\tilde{z} \wedge \mathcal{E})}{\Pr(\mathcal{E})} = \frac{1}{\Pr(\mathcal{E})} \sum_{e \in \mathcal{E}} \Pr(\tilde{z} \wedge e)$$

$$= \frac{1}{\Pr(\mathcal{E})} \sum_{e \in \mathcal{E}} \Pr(\tilde{z}|e)\Pr(e)$$

$$\hat{\mathcal{E}} = \arg\max_{\mathcal{E}} \textstyle\sum_{e \in \mathcal{E}} \Pr(\tilde{z}|e)\Pr(e)$$

Decoding



For an error vector e over a BSC, with crossover probability ϵ

$$\Pr(e) = \left(\frac{\epsilon}{1-\epsilon}\right)^{w(e)} (1-\epsilon)^n$$

For a BSC with crossover possibility and Hamming distance d,

$$\Pr(\tilde{z}|e) = \Pr(\tilde{z}|z(e)) = \left(\frac{\delta}{1-\delta}\right)^{d(z(e),\tilde{z})} (1-\delta)^m$$

Using the classical MAP decoder and ignoring the degeneracy

$$\hat{e} = \underset{e \in \mathbb{F}_2^n}{\arg\max}\, \Pr(e|\tilde{z}) = \underset{e \in \mathbb{F}_2^n}{\arg\max}\, \Pr(\tilde{z}|z(e))\Pr(e)$$

Instead of computing the above for all 2^n error vectors we can do the same for the much smaller 2^{n-k} syndrome set.

$$\Pr(\tilde{z}|z(e)) = \Pr(\tilde{z}|s(e))$$

Moreover the lowest weight error vector $e^*(s)$, maximizes the above expression, we can therefore compute a one-to-one mapping between $e^*(s)$ and s.

Reformulating we have, $\hat{s} = \underset{s \in \mathbb{F}_2^{n-k}}{\arg\max} \Pr(\tilde{z}|s)\Pr(e^*(s))$