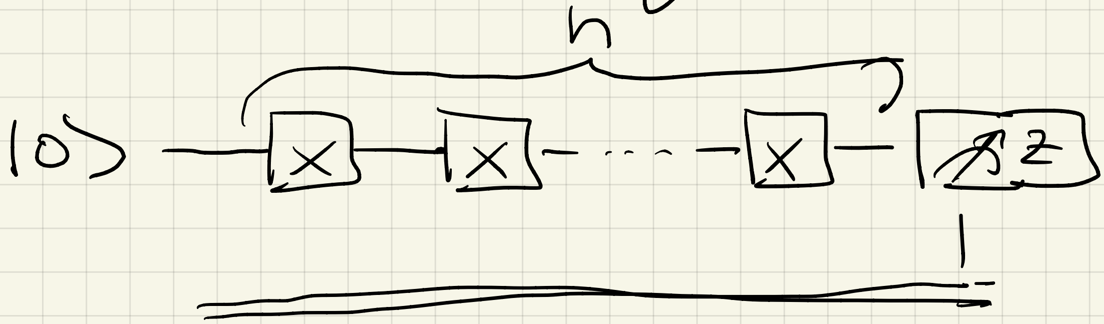
Lecture 2

Noisy Simulations

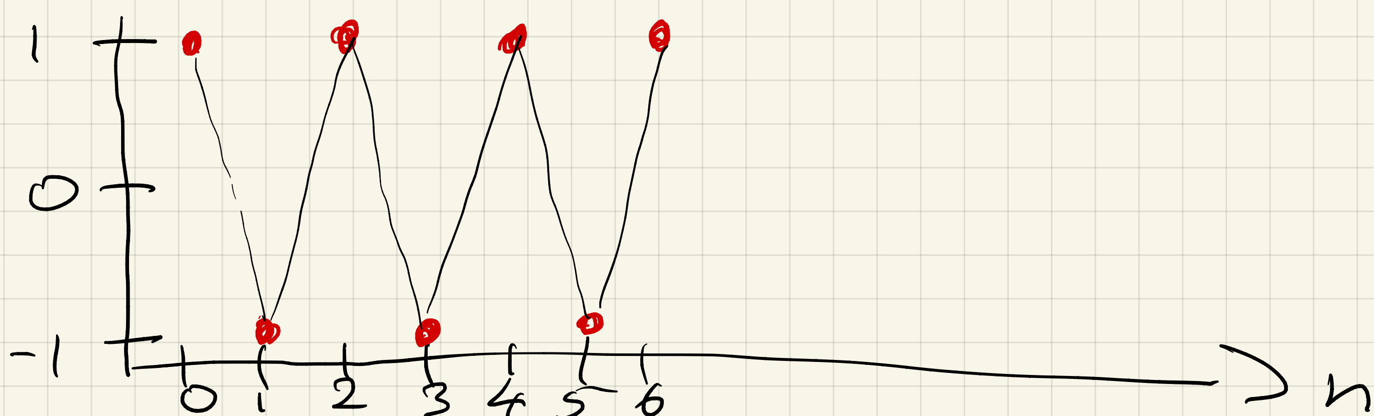
- From yesterdays lecture saw that noise can severly affect the results we can obtain from quantum computers.

- Goal of todays lecture is to understand the origin of this noise better and find ways to mitigate this noise.

- To start, consider a very simple circuit



- What do we expect from this circuit?



- In practice, no quantum computer is perfect, and there are 3 kinds of errors that a computer can make:

1) Coherent error:
The operations performed are not exact, so instead of applying X-gate, applies some slightly different $\tilde{X}$ gate

2) Sampling error:
Quantum mechanics is probabilistic, and one therefore has statistical uncertainties

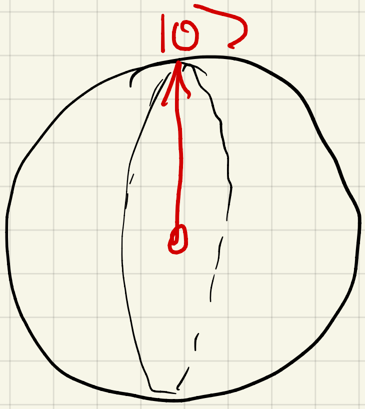
3) Measurement error:
The probability to measure a pure state $|\psi\rangle = |0,1\rangle$ in $|0,1\rangle$ state is not 100%

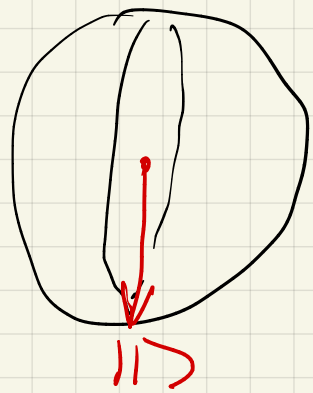
4) Incoherent error:
The quantum computer is coupled to environment, which will lead to decoherence.

- Show simple circuit in Qiskit.

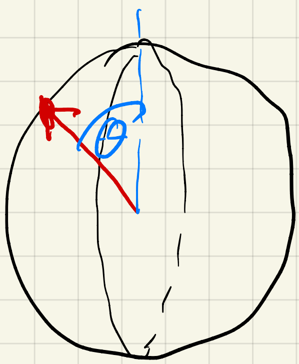
Coherent errors

- How does a QC implement X-gate?
- Know that any 1-qubit state lies on Bloch-sphere





- Can move states on Bloch-sphere by applying some pulses to my qubits
- One possibility is x-rotation $R_x(\Theta)$



- $\hat{X}|0\rangle = |1\rangle$, $\hat{X}|1\rangle = |0\rangle$
- This is the same as $R_x(\pi)$

- To be precise $\hat{X} = i\,\hat{R}_x(\pi)$
- So when applying an $\hat{X}$ gate, tell QC "Apply R_x rotation with angle π"
- But can never do things precisely, so will typically do $\hat{R}_x(\pi+\varepsilon) \equiv \hat{\tilde{X}}$
- Rotating by $\pi+\varepsilon$ is same as first rotating by π and then rotating by ε

$$\Rightarrow \tilde{X} = R_x(\varepsilon)\,\hat{X}$$

- Using simple math

$$R_x(\varepsilon) = \exp\left(\frac{-i\varepsilon}{2}\hat{X}\right)$$

$$= 1 - \frac{i\varepsilon}{2}\hat{X} + \frac{1}{2!}\left(\frac{i\varepsilon}{2}\right)^2 \hat{X}^2 + \frac{1}{3!}\left(\frac{i\varepsilon}{2}\right)^3 \hat{X}^3 + \dots$$

$$= 1 - \frac{i\varepsilon}{2}\hat{X} + \frac{1}{2!}\left(\frac{i\varepsilon}{2}\right)^2 + \frac{1}{3!}\left(\frac{i\varepsilon}{2}\right)^3 \hat{X} + \dots$$

$$= \cos\left(\frac{\varepsilon}{2}\right) I - i\sin\left(\frac{\varepsilon}{2}\right)\hat{X}$$

$$\Rightarrow \tilde{X} = \cos\left(\frac{\varepsilon}{2}\right)\hat{X} - i\sin\left(\frac{\varepsilon}{2}\right) I$$

In general

$$\tilde{X}^n = R_x^n(\varepsilon)\,\hat{X}^n$$

$$= R_x(n\varepsilon)\,\hat{X}^n$$

$$= \cos\left(\frac{n\varepsilon}{2}\right)\hat{X}^n - i\sin\left(\frac{n\varepsilon}{2}\right)\hat{X}^{n+1}$$

- Show results with different values of ε.

Sampling error

- Know that quantum mechanics is probabilistic, so only recover observables probabilistically.
- Say $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

$$\Rightarrow |\langle 0|\psi\rangle|^2 = \frac{1}{2} \quad , \quad |\langle 1|\psi\rangle|^2 = \frac{1}{2}$$

- But if only doing a single measurement will find either $|0\rangle$ or $|1\rangle$ when measuring.
- Probabilities only recovered statistically
- All know that statistical uncertainty scales as $\frac{1}{\sqrt{N}}$

- In general

$$|\psi\rangle = \alpha|0\rangle + \sqrt{1-\alpha^2}\,|1\rangle$$

$$\Rightarrow \sigma = \sqrt{\frac{\alpha^2(1-\alpha^2)}{N}}$$

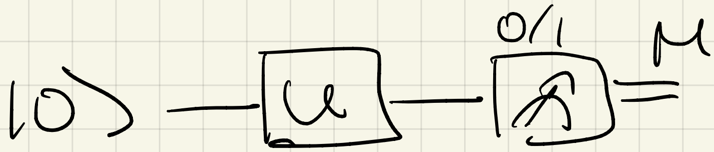
$\Rightarrow$ - need large number of measurements

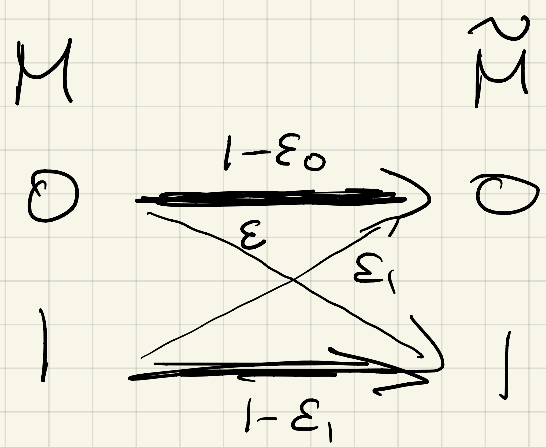
- error maximal if $\alpha = \frac{1}{\sqrt{2}}$.

- Show with more or less shots

Measurement error

- Assume we have circuit

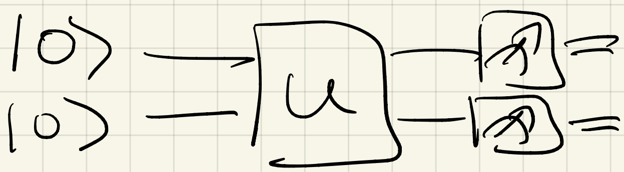


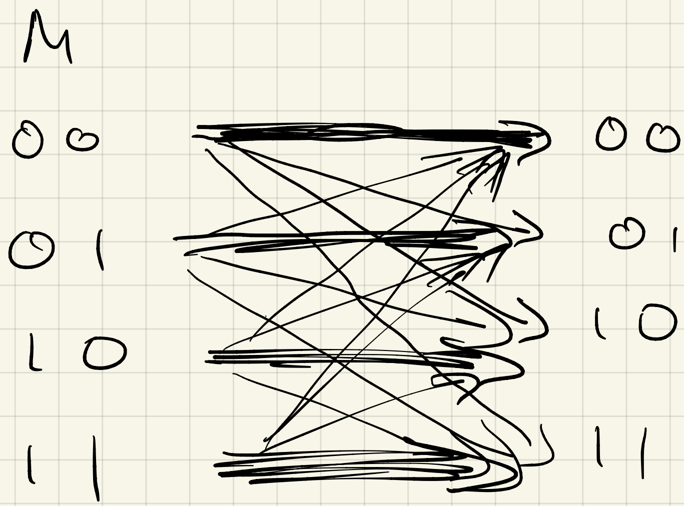


2 numbers describe measurement error of each qubit.

$\Rightarrow 2n_Q$ numbers

- In general, can have measurement errors that depend on all qubits





$\Rightarrow 2^{2n_q}$ numbers

- Show results for different measurement error values

Incoherent error

- Incoherent errors arise from interactions with environment.

- The combined system of $|\psi\rangle$ and $|e\rangle$ undergoes unitary evolution

- A general way of writing unitary evolution of $|\psi\rangle \otimes |e\rangle$ can be written as

$$|\psi\rangle \otimes |e\rangle \xrightarrow{U} \sum_k [E_k |\psi\rangle] \otimes |e_k\rangle \;, \quad \sum_k E_k^\dagger E_k = 1$$

- Density matrix transforms as

$$\rho = |\psi\rangle\langle\psi| \otimes |e\rangle\langle e| \xrightarrow{U} \sum_{k,l} [E_k |\psi\rangle\langle\psi| E_l^\dagger] \otimes |e_k\rangle\langle e_l|$$

- After tracing out environment

$$\rho_\psi \to \rho_\psi = Tr_e[\rho]$$
$$= \sum_{k,l} [E_k |\psi\rangle\langle\psi| E_l^\dagger] \langle e_l | e_k\rangle$$
$$= \sum_k [E_k |\psi\rangle\langle\psi| E_k^\dagger]$$

- Note that from derivation above, k runs over all basis state in $|e\rangle$.

- But we know that operators E_k act on state $|\psi\rangle$ so are $d \times d$ matrices. $\Rightarrow$ number restricted by dimension of $|\psi\rangle$

- $E_k = \sum_i C_{ik} B_i \quad , \quad Tr[B_i B_j] = d\,\delta_{ij}$

$$\Rightarrow \sum_k E_k \rho E_k^\dagger = \sum_{ij} \left(\sum C_{ik} C_{jk}^*\right) B_i \rho B_j^\dagger$$

$$= \sum_{ij} d_{ij} B_i \rho B_j^\dagger$$

$$d_{ij} = \sum_m u_{im} p_m u_{jm}^* \quad , \quad A_m = \sum_i u_{im} B_i$$

$$= \sum_{m=1}^{d^2} p_m A_m \rho A_m^\dagger$$

- From original unitarity, $\sum_m p_m = 1$

d : dim of Hilbert space

- All interactions with environment can be parametrized by d^2 Krauss operators.

- Consider a single qubit $\Rightarrow d = 2$

  $\Rightarrow$ 4 Krauss operators

- Choose as $\mathbb{1}, \hat{X}, \hat{Y}, \hat{Z}$

- $\rho \to p_1 \rho + p_x X \rho X + p_y Y \rho Y + p_z Z \rho Z$

  bit-flip bitphase-flip phase-flip

- Depolarizing channel assumes

$$p_x = p_y = p_z \equiv \frac{p}{4}$$

$$\Rightarrow \rho \to \left(1-\tfrac{3}{4}p\right)\rho + \frac{p}{4}\cdot X\rho X + \frac{p}{4} Y\rho Y + \frac{p}{4} Z\rho Z$$

$$= (1-p)\rho + \frac{p}{4}\left(\mathbb{1}\rho\mathbb{1} + X\rho X + Y\rho Y + Z\rho Z\right)$$

Take a generic density matrix $\rho = \begin{pmatrix} a & b \\ c & 1-a \end{pmatrix}$

$\mathbb{1}\rho\mathbb{1} = \begin{pmatrix} a & b \\ c & 1-a \end{pmatrix}$, $X\rho X = \begin{pmatrix} 1-a & c \\ b & a \end{pmatrix}$

$Y\rho Y = \begin{pmatrix} 1-a & -c \\ -b & a \end{pmatrix}$ $Z\rho Z = \begin{pmatrix} a & -b \\ -c & 1-a \end{pmatrix}$

$$\Rightarrow \mathbb{1}\rho\mathbb{1} + X\rho X + Y\rho Y + Z\rho Z = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\Rightarrow \rho \to (1-p)\rho + \frac{p}{2}\mathbb{1}$$

- Show result with depolarizing channel

- Show results with all channels.

Mitigating noise

- How big are typical noise levels?
- Assume incoherent errors for now
- Noise classified through probabilities in Kraus decomposition.
- Noise on single qubits $\sim 10^{-4} - 10^{-5}$
- Noise on entangling (CNOT) gates $\sim 10^{-2} - 10^{-3}$

  $\Rightarrow$ focus on CNOT gates

- Combining unitary operation with depolarizing channel gives

$$\rho \to (1-\varepsilon) U_q \rho U_q^\dagger + \varepsilon \left(\frac{\mathbb{1}_q}{2^{n_q}} \otimes \rho_{\bar{q}} \right)$$

- Apply this to CNOT gate acting on qubits k and l $\quad U_c^{(kl)}$

$$\rho \to (1-\varepsilon^{(kl)}) U_c^{(kl)} \rho U_c^{(kl)\dagger} + \varepsilon^{(kl)} \left(\frac{I_4}{4} \otimes \rho_{\bar{k}\bar{l}} \right)$$

- Apply 2 CNOT operations

$$\rho \xrightarrow{2\times} (1-\varepsilon^{(kl)})\, U_c \left[(1-\varepsilon^{(kl)}) U_c \rho U_c^{(kl)} + \varepsilon^{(kl)} \left(\frac{I_4}{4} \oplus \rho_{kl} \right) \right] U_c^{(kl)}$$

$$+ \varepsilon^{kl} \left(\frac{I_4}{4} \oplus \rho_{kl} \right)$$

$$= (1-\varepsilon^{(kl)})^2 U_c^{(kl)} U_c^{(kl)} \rho U_c^{(kl)} U_c^{(kl)}$$

$$+ (1-\varepsilon^{(kl)})\, \varepsilon^{(kl)} \frac{U_c^{(kl)} \mathbb{1}^{kl} U_c^{(kl)}}{4} \oplus \rho_{kl}$$

$$+ \varepsilon^{(kl)} \left(\frac{I_4}{4} \oplus \rho_{kl} \right)$$

$$= (1-\varepsilon^{(kl)})^2 \rho + \left[\varepsilon^{(kl)} - \varepsilon^{(kl)^2} + \varepsilon^{(kl)} \right] \left(\frac{I_4}{4} \oplus \rho_{kl} \right)$$

$$= (1-\varepsilon^{(kl)})^2 \rho + \left[1 - (1-\varepsilon^{(kl)})^2 \right] \left(\frac{I_4^{(kl)}}{4} \oplus \rho_{kl} \right)$$

- Apply 3 CNOTS

$$\rho \xrightarrow{3\times} (1-\varepsilon^{(kl)})^3 U_c \rho U_c + \left[1 - (1-\varepsilon^{(kl)})^3 \right] \left(\frac{I_4^{(kl)}}{4} \oplus \rho_{kl} \right)$$

- Apply $r = 2n+1$ CNOTS

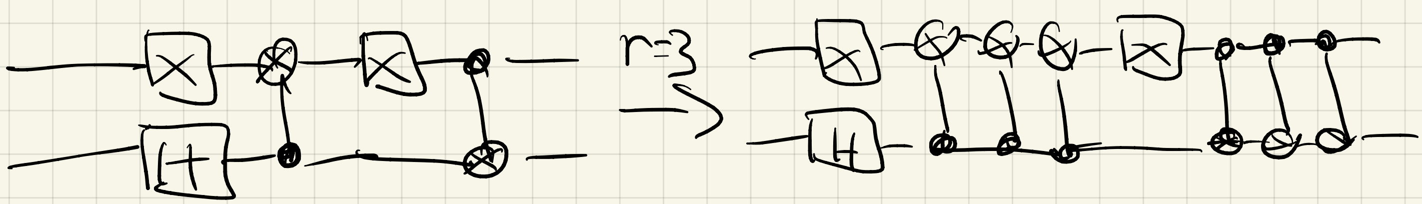
$$\rho \xrightarrow{r\times} (1-\varepsilon_{ij})^r U_c \rho U_c + \left[1 - (1-\varepsilon)^r \right] \left(\frac{I_4^{(kl)}}{4} \oplus \rho_{kl} \right)$$

- Taylor expanding around $\varepsilon = 0$

$$\rho \xrightarrow{r_x} (1 - r\,\varepsilon_i)\, U_c^{(k\ell)} \rho\, U_c^{(k\ell)} + r\,\varepsilon_i \left(\frac{I_4}{4} \oplus \rho_{k\ell} \right)$$

- Applying r_i CNOT gates is same as applying single CNOT, but with depolarizing parameter amplified by r_i

- Now consider a circuit parametrized by r (how many CNOTs for each CNOT)



- Circuit will produce density matrix $\rho(r)$
- Now measure observable M and obtain

$$\langle M(r)\rangle = \mathrm{Tr}[M\rho(r)]$$

- Use expression of $\rho(r)$ to obtain

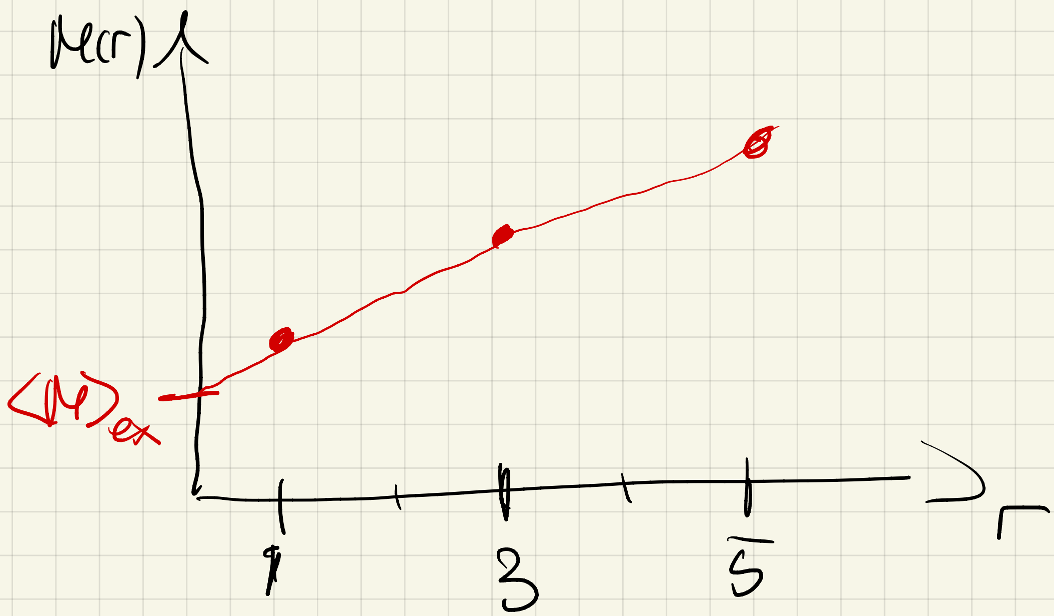
$$\langle M(r)\rangle = \left(1 - r\sum_{i=1}^{N_c} \varepsilon_i\right) \langle M\rangle_{ex} + r\sum_{i=1}^{N_c} \varepsilon_i \langle M\rangle_{dep}$$

where $\langle M\rangle_{ex}$ is expectation value in absence of noise, and $\langle M\rangle_{dep;i}$ is exp value of CNOT$_i$ is replaced by depolarizing channel

- Thus

$$\langle M\rangle_{ex} = \langle M(0)\rangle$$

and one can extract $\langle M\rangle_{ex}$ by extrapolating $\langle M(r)\rangle$ to zero



- So far have done linear fits, which eliminated $O(\varepsilon^2)$ errors
- In general can eliminate depolarizing error to all orders

- Review linear fit again

$$\langle M\rangle(1) = \langle M\rangle_{ex} + N_{CNOT}\,\varepsilon\left[\sum_i \langle M\rangle_{dep_i} - \langle M\rangle_{ex}\right] + O(\varepsilon^2)$$

$$\langle M\rangle(3) = \langle M\rangle_{ex} + 3N_{CNOT}\,\varepsilon\left[\sum_i \langle M\rangle_{dep_i} - \langle M\rangle_{ex}\right] + O(\varepsilon^2)$$

$$\Rightarrow \frac{3}{2}\langle M\rangle(1) - \frac{1}{2}\langle M\rangle(3) = \langle M\rangle_{ex} + O(\varepsilon^2)$$

- This can be generalized (Richardson extrapolation)

- Assume we have $\langle M\rangle(r)$ for $r = 1, 3, \ldots, r_{max}$
- Write linear combination

$$\sum_{n=0}^{n_{max}} a(n)\langle M\rangle(1+2n) \overset{!}{=} \langle M\rangle_{ex} + O(\varepsilon^{n_{max}+1})$$

- One can work out general equation

$$\begin{aligned}\langle M\rangle(r) = {} & (1-\varepsilon)^{N_c r}\langle M\rangle_{ex} \\ & + (1-\varepsilon)^{(N_c-1)r}\left[1-(1-\varepsilon)^r\right]\sum_i \langle M\rangle_{dep_i} \\ & + (1-\varepsilon)^{(N_c-2)r}\left[1-(1-\varepsilon)^r\right]^2\sum_{i_1 i_2} \langle M\rangle_{dep_{i_1 i_2}} \\ & + \ldots \\ & + \left[1-(1-\varepsilon)^r\right]^{N_c}\sum_{i_1 \ldots i_{N_c}} \langle M\rangle_{dep_{i_1 \ldots i_{N_c}}}\end{aligned}$$

$$= \langle M\rangle_{ex} - f_{N_c}(r,\varepsilon)\langle M\rangle_{ex}$$
$$+ \left[f_{N_c}(r,\varepsilon) - f_{N_c-1}(r,\varepsilon)\right]\sum_{i} \langle M\rangle_{dep\, i}$$
$$+ \left[f_{N_c}(r,\varepsilon) - f_{N_c-2}(r,\varepsilon)\right]\sum_{i_1,i_2} \langle M\rangle_{dep\, i_1 i_2}$$
$$+ \ldots$$
$$+ f_{N_c}(r,\varepsilon)^{N_c}\sum_{i_1,\ldots,i_{N_c}} \langle M\rangle_{dep\, i_1 \ldots i_{N_c}}$$

with $f_n(r,\varepsilon) = 1 - (1-\varepsilon)^{nr}$

- Important to remember that $\langle M\rangle_{ex}$ $\langle M\rangle_{dep\, i_1}$ $\langle M\rangle_{dep\, i_1\ldots i_{N_c}}$ are results of observables in noiseless circuit one does not have access to.

- But one can find coefficients $a(i)$ such that all $\langle M\rangle_{dep\, i}$ values cancel to $O(\varepsilon^{2n_{max}+1})$

- One finds

$$a(i) = \prod_{j=0,\, j\neq i}^{n_{max}} \frac{(1+2j)}{2(j-i)}$$

$$= \frac{2^{-2n_{max}}}{i!} \frac{(-1)^i}{1+2i} \frac{(1+2n_{max})!}{n_{max}!\,(n_{max}-i)!}$$

- Note that this is the same as performing a polynomial fit with degree $n_{max}-1$