Quantum Information

(ELEC-C9440)

Lecture 9

Matti Raasakka
University Lecturer
Micro and Quantum Systems group
Dept. of Electronics and Nanoengineering
School of Electrical Engineering

$\langle \mu | Q a^{\dagger} | S - 1 \rangle$

Micro and Quantum Systems

Aalto University
School of Electrical
Engineering

Topics

- Simulation of quantum systems with a quantum computer
- Error mitigation for NISQ machines

NISQ era

- Most quantum algorithms (e.g. Shor, Grover,…) require many thousands of near-perfect qubits to solve practically relevant problems. Error correction needed to push the error rates down. => A lot of extra overhead, thousands of physical qubits per one logical qubit. Likely not possible in the very near future.
- In near future we only have a few hundred to few thousand noisy qubits at our disposal. We are currently in the **"noisy intermediate-scale quantum" (NISQ) era.**
- **Can we still do something useful with NISQ machines?**
- Some practical applications that have been proposed for NISQ machines:
  - **Simulation of quantum systems**
  - Optimization and machine learning (next lecture)
- Even though full error correction is out of the question, there are less resource intensive ways to reduce errors. => **Error mitigation**.
- Preskill, "Quantum computing in the NISQ era and beyond", *Quantum* **2, 79** (2018), https://doi.org/10.22331/q-2018-08-06-79

Digital quantum simulation

Digital quantum simulation

Nielsen, Chuang section 4.7

- Time-evolution of a quantum system governed by the Schrödinger equation

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\varphi(t)\rangle = \hat{H}|\varphi(t)\rangle \quad \Leftrightarrow \quad |\varphi(t+s)\rangle = e^{-\frac{i}{\hbar}\hat{H}s}|\varphi(t)\rangle$$

- The resource requirements of exactly simulating a generic quantum system on a classical computer grow **exponentially** in the size of the system. E.g., describing the density matrix of n qubits takes $4^n - 1$ real numbers. For quantum particles, the Hilbert space dimensionality grows exponentially in the number of particles.
  => **Use quantum computers to simulate quantum systems.**
- Analog simulation also possible, but here we concentrate on **digital** simulation with a universal quantum computer.
- Two types of dynamical quantum systems most often considered:
  - Quantum many-particle systems in continuous space
  - Finite quantum systems on a lattice (e.g. models for magnetization, lattice QCD,…)

Trotterization of time–evolution

- For simulation, we can split the continuous time into small time steps of size Δt, and update the quantum state one step at a time up to some error $O(\Delta t^k)$.
- In many cases the Hamiltonian is a sum of different terms $\widehat{H} = \sum_L \widehat{H}_L$, where each term is easy to simulate by itself.
- **Trotter formula**: $\lim_{n\to\infty}(e^{iAt/n}e^{iBt/n})^n = e^{i(A+B)t}$ (see also Baker-Campbell-Hausdorff formula)
- Trotter formula allows to simulate each term in the Hamiltonian separately, up to a **second-order error** in Δt:

$$\underline{e^{i(A+B)\Delta t} = e^{iA\Delta t}e^{iB\Delta t} + O(\Delta t^2)}$$

$$e^{i(A+B)\Delta t} = e^{iA\Delta t}\, e^{iB\Delta t} \times e^{-i[A,B]\Delta t^2} \times \ldots$$

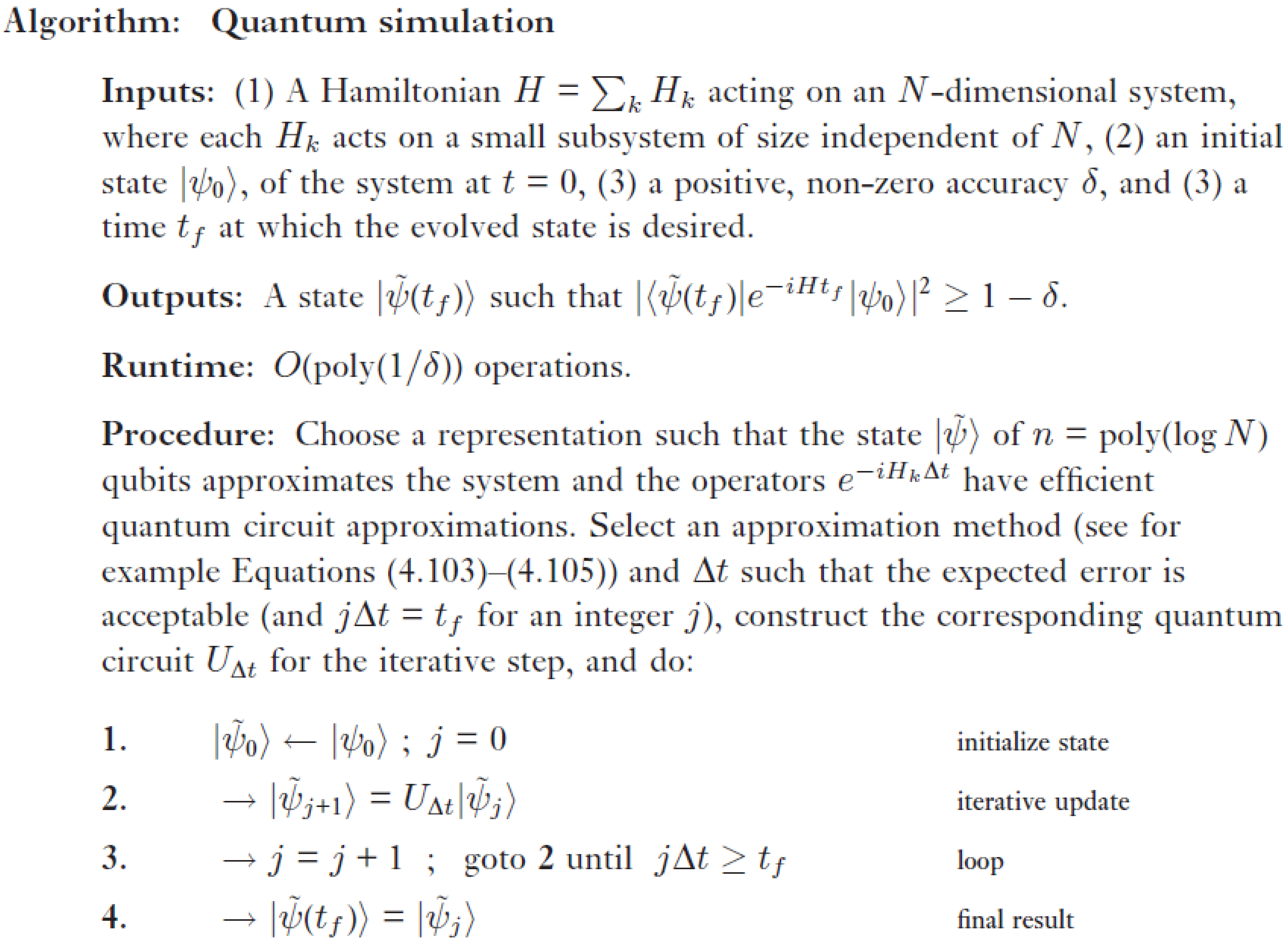
- There is even a slightly modified formula with third-order error:

$$e^{i(A+B)\Delta t} = e^{iA\Delta t/2}e^{iB\Delta t}e^{iA\Delta t/2} + O(\Delta t^3)$$

- If we can implement individual unitaries $e^{-\frac{i}{\hbar}\widehat{H}_L\Delta t}$ separately, we can simulate the whole Hamiltonian up to a small error by choosing small enough $\Delta t = t/N$!

$$e^{-\frac{i}{\hbar}\widehat{H}t} = \prod_{n=1}^{N}\prod_L e^{-\frac{i}{\hbar}\widehat{H}_L\Delta t} + O(\Delta t^2)$$

Polynomial time simulation algorithm



Algorithm: Quantum simulation

Inputs: (1) A Hamiltonian $H = \sum_k H_k$ acting on an N-dimensional system, where each H_k acts on a small subsystem of size independent of N, (2) an initial state $|\psi_0\rangle$, of the system at $t = 0$, (3) a positive, non-zero accuracy δ, and (3) a time t_f at which the evolved state is desired.

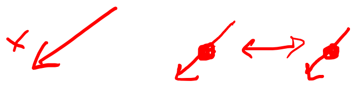
Outputs: A state $|\tilde{\psi}(t_f)\rangle$ such that $|\langle\tilde{\psi}(t_f)|e^{-iHt_f}|\psi_0\rangle|^2 \geq 1 - \delta$.

Runtime: $O(\text{poly}(1/\delta))$ operations.

Procedure: Choose a representation such that the state $|\tilde{\psi}\rangle$ of $n = \text{poly}(\log N)$ qubits approximates the system and the operators $e^{-iH_k\Delta t}$ have efficient quantum circuit approximations. Select an approximation method (see for example Equations (4.103)–(4.105)) and Δt such that the expected error is acceptable (and $j\Delta t = t_f$ for an integer j), construct the corresponding quantum circuit $U_{\Delta t}$ for the iterative step, and do:

```
1.   $|\tilde{\psi}_0\rangle \leftarrow |\psi_0\rangle$ ; $j = 0$                 initialize state
2.     $\rightarrow |\tilde{\psi}_{j+1}\rangle = U_{\Delta t}|\tilde{\psi}_j\rangle$          iterative update
3.     $\rightarrow j = j + 1$ ;  goto 2 until $j\Delta t \geq t_f$      loop
4.     $\rightarrow |\tilde{\psi}(t_f)\rangle = |\tilde{\psi}_j\rangle$                final result
```

Simple simulation example

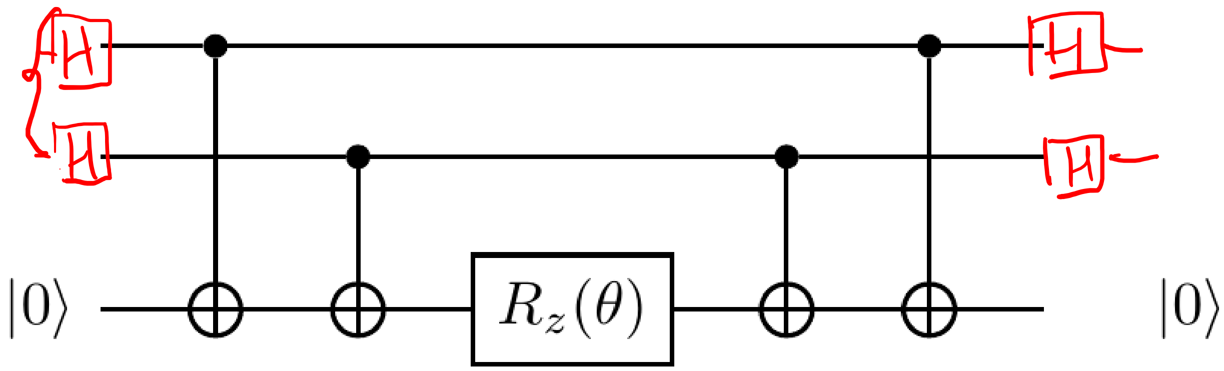


- Simulate for time T a pair of spin-$\frac{1}{2}$ systems with the Hamiltonian
$$H = \alpha(X_1 + X_2) + \beta Z_1 Z_2$$
starting with some initial state $|\varphi(0)\rangle$.

- The two terms in the Hamiltonian do not commute, so must use **Trotterization**. Split the time-interval into N steps of duration $\Delta t = T/N$. Approximate:
$$e^{-\frac{i}{\hbar}\hat{H}T} = \prod_{n=1}^{N} e^{-\frac{i}{\hbar}\beta Z_1 Z_2 \Delta t} e^{-\frac{i}{\hbar}\alpha X_2 \Delta t} e^{-\frac{i}{\hbar}\alpha X_1 \Delta t} + O(\Delta t^2)$$

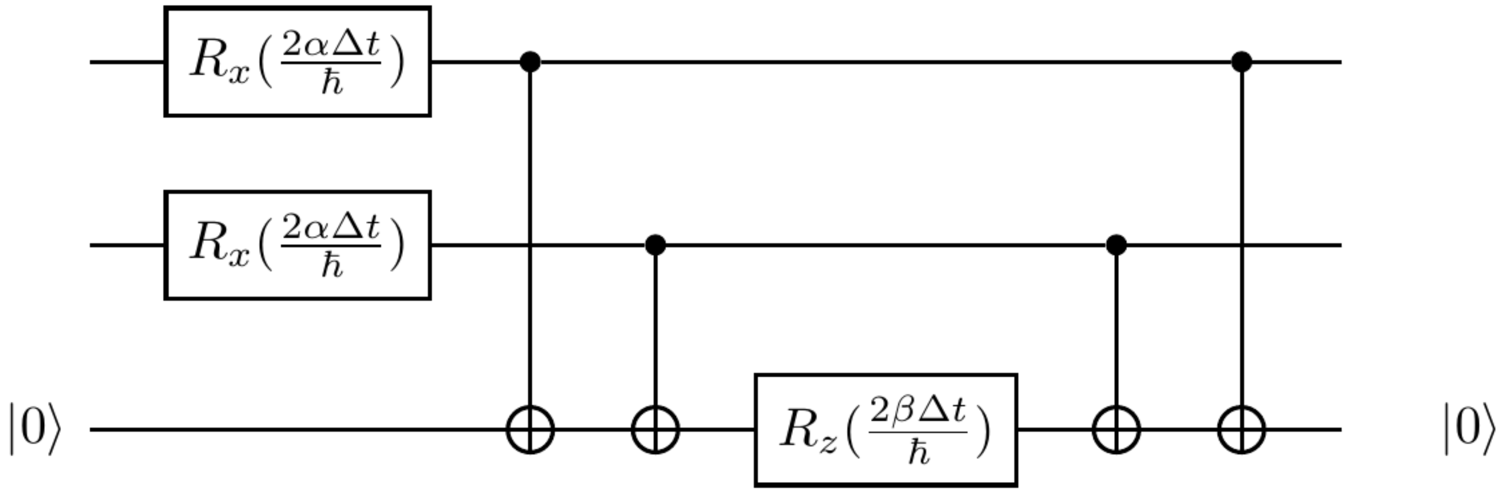
- Here, $e^{-\frac{i}{\hbar}\alpha X_i \Delta t}$ are simple single-qubit rotations around the X-axis, $R_x\left(\frac{2\alpha\Delta t}{\hbar}\right)$.

- The entangling unitary $e^{-\frac{i}{2}\theta Z_1 Z_2}$ can be implemented by the following circuit:



Simple example

- All together, we apply the below circuit $N = T/\Delta t$ times to the initial state to reach the final state at time $t = T$ with error $O(\Delta t^2)$:



$R_x(\frac{2\alpha\Delta t}{\hbar})$
$R_x(\frac{2\alpha\Delta t}{\hbar})$
$|0\rangle$
$R_z(\frac{2\beta\Delta t}{\hbar})$
$|0\rangle$

- **NOTE:** This simple construction can be easily generalized to any Hamiltonian, which is a sum of local interaction terms, i.e., tensor product of only a few Paulis: Any other entangling unitary $e^{-\frac{i}{2}\theta X_1 Y_2 Z_3}$ etc. can be similarly implemented with one ancilla by first transforming by 1-qubit unitaries to the Z-basis, and then applying the CNOTs to the qubits on which the unitary acts.

Simulation of one–particle Schrödinger equation

- To digitally simulate a point-particle in continuous space, we must restrict the system size and discretize the space.
- The spatial discretization scale Δx must be much smaller than any relevant wavelength for the problem. De Broglie relation: $\lambda = h/p$.
- For example, consider a particle in a 1D system of size L. We discretize the interval into $M = L/\Delta x$ smaller intervals of size Δx. The continuous wave function can then be approximated by a discretized version as

$$|\varphi\rangle = \int_{-L/2}^{L/2} \varphi(x)\, \underline{|x\rangle}\, \mathrm{d}x \;\rightarrow\; |\varphi\rangle = \sum_{m=-M/2}^{M/2} \varphi_m \underline{\underline{|m\Delta x\rangle}}$$

- $M + 1$ spatial basis states $|m\Delta x\rangle$ can be encoded as the computational basis states of $\sim \log_2(M+1)$ qubits. $|m\Delta x\rangle \leftrightarrow |m\rangle$

$|\varphi_1\rangle|\varphi_2\rangle$

- **NOTE:** The number of qubits scales linearly also in the number of particles. Need $\sim N \log_2(M+1)$ qubits to describe N particles.

Simulation of one–particle Schrödinger equation

- The one-particle Hamiltonian operator is $\hat{H} = -\frac{\hbar^2}{\underline{2m}}\hat{p}^2 + \underline{V(\hat{x})}$.
- We can again apply Trotterization by splitting the time interval into smaller intervals of duration Δt, and then implement the two terms separately:

$$\underline{e^{-\frac{i}{\hbar}\hat{H}\Delta t}} = \underline{e^{i\frac{\hbar\Delta t}{2m}\hat{p}^2}}\,\underline{e^{-i\frac{\Delta t}{\hbar}V(\hat{x})\Delta t}} + O(\underline{\Delta t}^2)$$

- The **potential energy term** is efficient to implement, if $V(x)$ can be computed efficiently, as it acts diagonally in the comp. basis:

$$e^{-i\frac{\Delta t}{\hbar}V(\hat{x})}|m\rangle = e^{-i\frac{\Delta t}{\hbar}\underline{V(m\Delta x)}}|m\rangle$$

- To implement the **kinetic energy term**, we can move from the position basis to the momentum basis via the **quantum Fourier transform**. We have

$$e^{i\frac{\hbar\Delta t}{2m}\hat{p}^2} = U_{\underline{QFT}}\,e^{i\frac{\hbar\Delta t}{2m}\hat{x}^2}\,U_{QFT}^{\dagger}$$

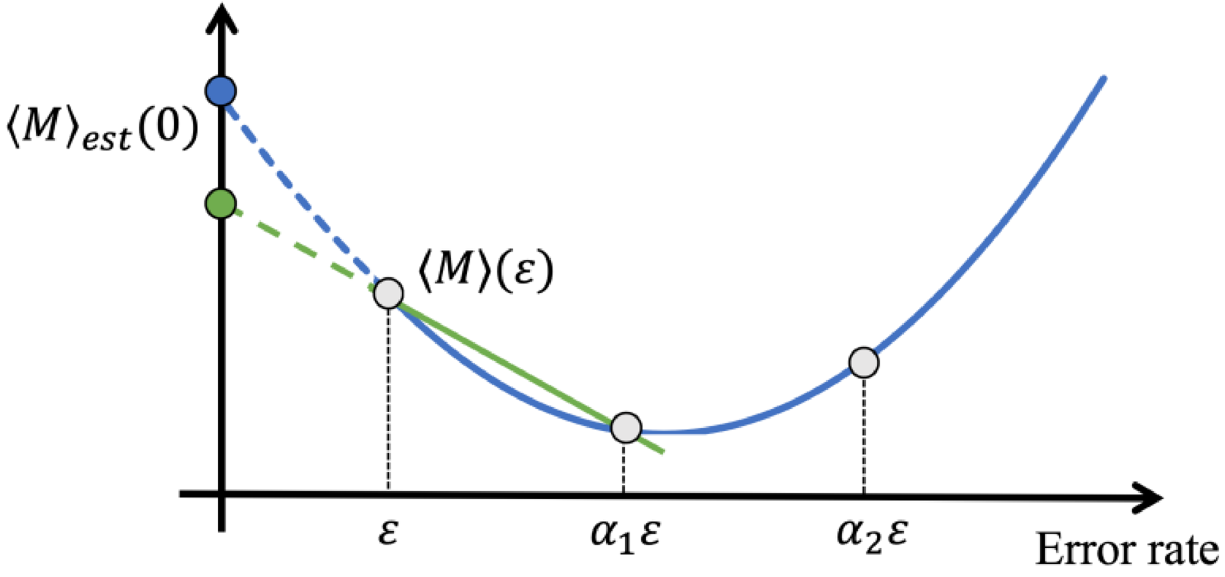
  where $e^{i\frac{\hbar\Delta t}{2m}\hat{x}^2}$ is again diagonal in the computational basis, and thus also efficient to implement.

Error mitigation

Error mitigation

Endo et al. "Hybrid quantum-classical algorithms and quantum error mitigation", *J. Phys. Soc. Jpn.*, **90**, 032001 (2021), arXiv:2011.01382.

- Current quantum computers not large enough to apply proper quantum error correction methods. Can we still do something to reduce the errors? **Yes!** => **Quantum error mitigation**
- Several different methods of various sophistication, here we only cover a couple of the simplest methods. See e.g. the cited article by Endo (chapter V) for more.
- The general idea behind **extrapolation methods**: Run the algorithm with several different error rates ε_k and extrapolate back to $\varepsilon = 0$.
- Can also be used to reduce the error in quantum simulation by simulating with different values of Δt, and then extrapolating back to $\Delta t = 0$.



Richardson extrapolation

- Often we want to measure the expectation value $\langle M \rangle$ of some observable for the output state of some quantum circuit. The result we get from the quantum computation is affected by the errors during computation.
- In general, we have many different error rates for different quantum gates and errors, but for simplicity let’s assume we have only a single physical error rate ε. In this case, we can consider the expectation value to be a (presumably differentiable) function of the error rate, $\langle M \rangle(\varepsilon)$.
- We may expand the expectation value as a polynomial series in ε around $\varepsilon = 0$,

$$\langle M \rangle(\varepsilon) = \langle M \rangle(0) + \sum_{k=1}^{n-1} M_k \varepsilon^k + O(\varepsilon^n)$$

- If we now know the values of $\langle M \rangle(\varepsilon)$ for n different values of ε, we may solve for the n unknowns $\langle M \rangle(0)$ and M_k up to an error of order ε^n. This will give us an estimate of $\langle M \rangle(0)$.

Richardson extrapolation

- Let's assume we have measured the expectation value for different error rates $\alpha_k \varepsilon$, where $1 = \alpha_0 < \alpha_1 < \cdots < \alpha_{n-1}$. Solving the equations gives us the estimate

$$\langle M \rangle_{est}(0) = \sum_{k=0}^{n-1} \beta_k \underline{\underline{\langle M \rangle(\alpha_k \varepsilon)}} = \langle M \rangle(0) + O(\varepsilon^n)$$

where $\beta_k = \prod_{i \neq k} \frac{\alpha_i}{\alpha_i - \alpha_k}$.

- The accuracy of the estimate increases with n. However, the variance also increases as

$$\mathrm{Var}[\langle M \rangle_{est}(0)] = \sum_{k=0}^{n} \beta_k^2 \mathrm{Var}[\langle M \rangle(\alpha_k \varepsilon)]$$

which is the order of $\gamma_{Ric} = \sum_k \beta_k^2$ larger than a measurement at a single error rate. Therefore, we need to make $\sim\gamma_{Ric}$ more measurements to achieve the same accuracy for the extrapolated value. γ_{Ric} grows exponentially in n, so only possible to apply Richardson extrapolation for a small n.

Exponential extrapolation

- Similar to Richardson extrapolation, but instead of polynomial expansion we use exponential expansion.
- More suitable for longer circuits for small error rate ε, since in this case the dependence on the error rate is closer to exponential. The number of errors follows binomial distribution (when assumed to be independent), which can be approximated by the Poisson distribution,

$$p_k \approx e^{-N_g \varepsilon} \frac{(N_g \varepsilon)^k}{k!}$$

probability of k errors

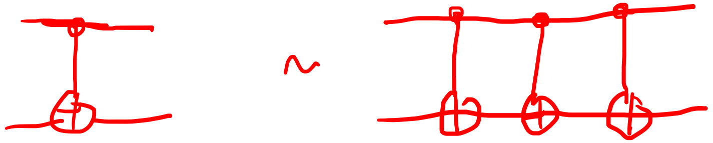
when $N_g \varepsilon = O(1)$, where N_g is the number of gates in the circuit.

- Leads to a different estimate

$$\langle M \rangle_{est}(0) = \frac{\alpha e^{N_g \varepsilon} \langle M \rangle (\varepsilon) - e^{N_g \alpha \varepsilon} \langle M \rangle (\alpha \varepsilon)}{\alpha - 1}$$

$$\gamma_{exp} = \frac{\alpha^2 e^{2 N_g \varepsilon} + e^{2 N_g \alpha \varepsilon}}{(\alpha - 1)^2}$$

Methods to increase the physical error rate



- **Repeat gates:**
  - If a single gate has error rate ε, k such gates have approximately an error rate $k\varepsilon$. If some gate G is idempotent, i.e., $\underline{G^2 = I}$ (such as Pauli gates and CNOT), we can replace G by $2n + 1$ copies of G wherever G appears in the circuit and thus change its error rate to $(\underline{2n + 1})\varepsilon$.
  - Also possible to introduce non-integer multiples by replacing gates probabilistically. If gate $\underline{G}$ is replaced by $\underline{2n + 1}$ copies with probability $\underline{p_n}$, we have $\sum_n p_n (EG)^{2n+1} \approx (1 - (2\langle n \rangle + 1)\varepsilon)G + (2\langle n \rangle + 1)\varepsilon G'$, where G' is the gate with error, thus scaling the error rate by $2\langle n \rangle + 1$.
- **Hamiltonian scaling:** For open quantum systems, the non-unitary dynamics can be enhanced by a simultaneous scaling of the system Hamiltonian and time.
- **Pauli twirling:** If the 2-qubit gate error rates are much higher than 1-qubit gate error rates, error rates of 2-qubit gates can be increased by introducing probabilistically random 1-qubit gates before and after the 2-qubit gates.

Extrapolation example

- Assume we have a quantum circuit, which prepares some state, in which we want to measure the expectation value $\langle M\rangle$ of some operator M. Assume further that the main source of error are the 2-qubit CNOT gates with error rate $\varepsilon \approx 0.01$.
- We can triple the error rate from ε to 3ε by replacing each CNOT gate in the circuit by three CNOT gates.
- Imagine we measure the values $\langle M\rangle(\varepsilon) = 0.641$ and $\langle M\rangle(3\varepsilon) = 0.658$. For the extrapolation parameters we have $\alpha_0 = 1, \alpha_1 = 3$. We can now apply Richardson extrapolation to these values to estimate the error-free expectation value:

$$\langle M\rangle_{est}(0) = \sum_{k=0}^{n-1} \beta_k \langle M\rangle(\alpha_k \varepsilon) = \frac{\alpha_1}{\alpha_1 - \alpha_0}\langle M\rangle(\alpha_0\varepsilon) + \frac{\alpha_0}{\alpha_0 - \alpha_1}\langle M\rangle(\alpha_1\varepsilon)$$

$$= \frac{3}{3-1}\langle M\rangle(\varepsilon) + \frac{1}{1-3}\langle M\rangle(3\varepsilon) = \frac{3}{2}\cdot 0.641 - \frac{1}{2}\cdot 0.658 = 0.6325$$

Quasi–probability method

- Assume a single noisy gate is described by the quantum channel $\mathcal{E} \circ \mathcal{U}$, where $\mathcal{U}$ is the noise-free gate operation and $\mathcal{E}$ is the noise.
- The idea behind quasi-probability method is that we can cancel the effect of noise by acting on the result by the inverse operation $\mathcal{E}^{-1}$, so that we recover the noise-free gate operation: $\mathcal{E}^{-1} \circ \mathcal{E} \circ \mathcal{U} = \mathcal{U}$.
- However, it is usually not possible to implement $\mathcal{E}^{-1}$ as a unitary transformation. Must decompose it into a linear combination of unitary channels, $\mathcal{E}^{-1} = \sum_i q_i \mathcal{B}_i$.
- We then have

$$\mathcal{U} = \mathcal{E}^{-1} \circ \mathcal{E} \circ \mathcal{U} = \sum_i q_i \mathcal{K}_i \,, \qquad \mathcal{K}_i = \mathcal{B}_i \circ \mathcal{E} \circ \mathcal{U}$$

$$\langle M \rangle_{\mathcal{U}(\rho)} = \sum_i q_i \langle M \rangle_{\mathcal{K}_i(\rho)}$$

where each $\langle M \rangle_{\mathcal{K}_i(\rho)}$ is a measurement with the noisy channel, but augmented with the extra operation $\mathcal{B}_i$.

Quasi-probability method

$$\langle M\rangle_{\mathcal{U}(\rho)} = \sum_i q_i \langle M\rangle_{\mathcal{K}_i(\rho)} = C_\mathcal{E} \sum_i \mathrm{sgn}(q_i)\, p_i \langle M\rangle_{\mathcal{K}_i(\rho)}$$

$$C_\mathcal{E} = \sum_i |q_i|, \qquad p_i = \frac{|q_i|}{C_\mathcal{E}}$$

- We see that $\langle M\rangle_{\mathcal{U}(\rho)}$ can be estimated by stochastic sampling of the different channels $\mathcal{K}_i$ with probabilities p_i
- Variance is amplified by $\gamma_Q = C_\mathcal{E}^2$ for one gate. For a circuit with N_g gates,

$$\gamma_{tot} = \prod_{k=1}^{N_g} C_k^2 \sim e^{2b\varepsilon N_g}, \qquad C_k = 1 + b\varepsilon$$

- Downside of the q-prob method: Must know the noise properties for the gates. Not trivial to obtain due to confounding measurement errors etc.

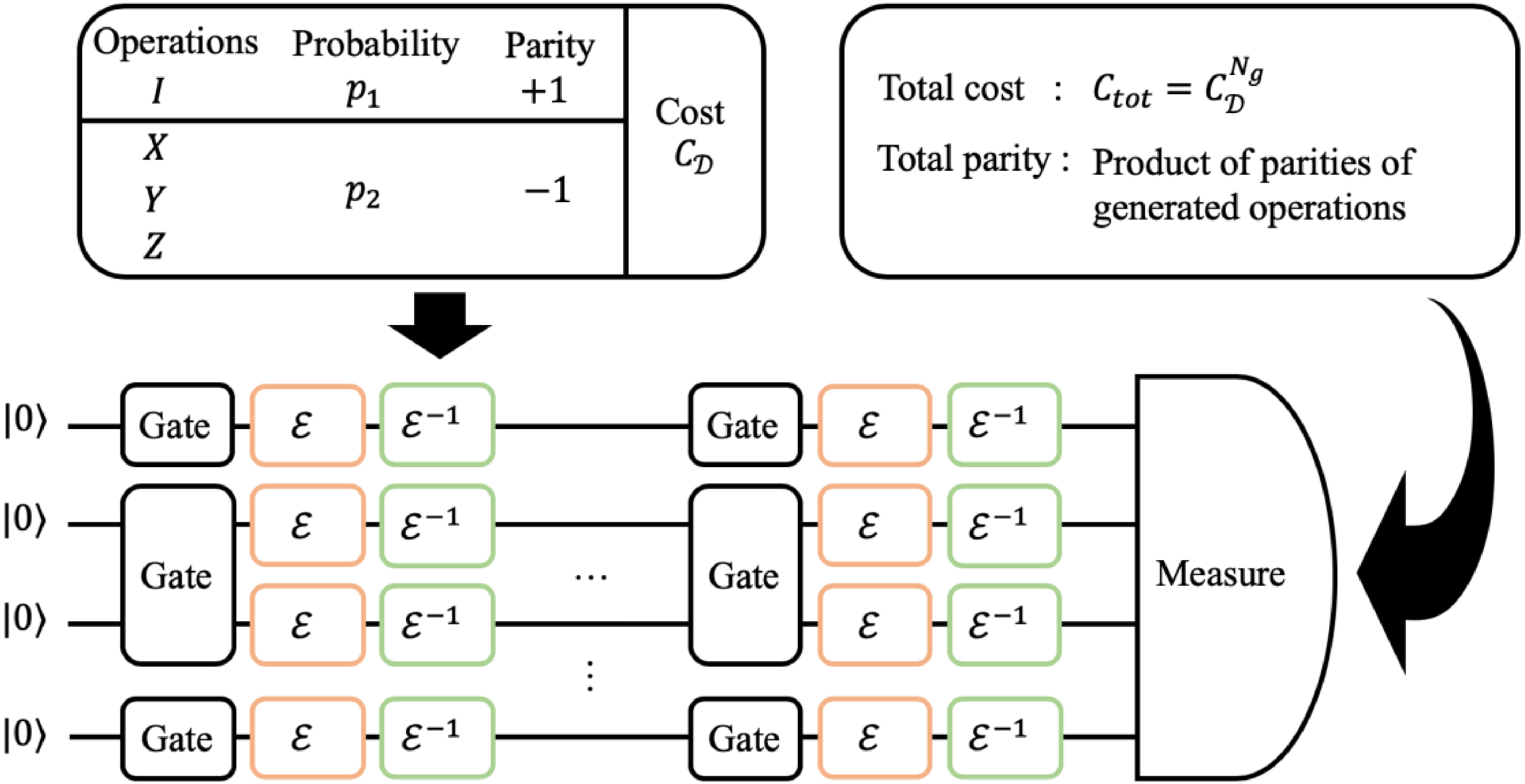


FIG. 15. Schematic of the quasi-probability method. We apply the inverse channel to cancel the noise process $\mathcal{E}$ with quasi-probability method. The inverse operation specified in the figure corresponds to the case of depolarising channel. For simplicity, we just show the case that two-qubit gate error is a tensor product of depolarising channels. We multiply the product of parities of generated operations $\mathrm{sgn}(q_{\vec{i}})$ to the outcome. We compute the average of product of the parity and the outcome, which is multiplied with the cost C_{tot} to approximate the ideal expectation value of the observable.

Noise cancellation example

- Consider depolarizing noise

$$\mathcal{D}(\rho) = (1 - \frac{3}{4}p)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z)$$

- The inverse to this channel is given by

$$\mathcal{D}^{-1}(\rho) = C_{\mathcal{D}}[p_1\rho - p_2(X\rho X + Y\rho Y + Z\rho Z)],$$

  where

$$C_{\mathcal{D}} = \frac{p+2}{2-2p}, \qquad p_1 = \frac{4-p}{2p+4}, \qquad p_2 = \frac{p}{2p+4}$$

- To cancel the depolarizing noise in gate $\mathcal{U}$, we append the gates X, Y, Z to it with probability p_2. When the state is measured, the measurement result is multiplied by the parity corresponding of the generated operation and the constant $C_{\mathcal{D}}$.
- By repeating this procedure many times, we can recover the measurement outcome with the ideal noise-free gate.

Other error mitigation methods

Several other methods, not necessarily based on extrapolation:

- Quantum subspace expansion
- Symmetry verification
- Learning-based error mitigation
- Stochastic error mitigation
- …
- Combinations of some/all of the above

- Several different quantum error mitigation methods have been shown numerically and experimentally to lead to significant improvements on computational fidelity. (See e.g. the review by Endo et al. for details.)