Pauli Propagation

Henry Lin

June 2025

1 Introduction

Suppose we want to calculate an expectation of an observable(real eigenvalue), then it is defined below:

$$\langle H \rangle = \langle \psi | V^{\dagger} H V | \psi \rangle$$

In the Schrödinger picture, we are evolving $|\psi\rangle$ to $V|\psi\rangle$ and calculate the expected energy.
In the Heisenberg picture, we are evolving the observable H to $H' = V^{\dagger}HV$ and calculate the expected energy.
These two ways of calculating the expected energy are the same because of associative property. We will be focusing on the Heisenberg picture for the rest of this article.

Clifford Gates consists of $\{H, CX, CZ, X, Y, I\}$. If the circuit only consists of Clifford gates, then by Gottesman-Knill theorem, we can simulate the circuit efficiently on a classical hardware.

Pauli propagation is not limited by the number of qubits or the level of entanglements.

1.1 Pauli Weight

The maximum number of qubits on which each Pauli string produced act, i.e. the maximum number of non-identity operators in any Pauli string produced by the mapping. This is, in general, referred to as Pauli weight. Low Pauli weight means a Pauli string that has a lot of identities.

1.2 Pauli String

The number of different Pauli strings resulting from the mapping. The number of Pauli strings directly impacts the cost of implementing a VQE. As such, one should always prefer to have the lowest number of Pauli strings to measure. In general, we see that this number of strings scales O(n4) for molecular Hamiltonian, and with the number of edges for lattice models.

2 Clifford Gates

any Pauli word is mapped to another Pauli word. This means $H'P_l \in P_l$, if $H, V \in Cl$. Cl is the Clifford group.

3 Non-Clifford Gates

On to some non-Clifford gates, namely Pauli rotation gates. They have the important property of mapping a Pauli word to two Pauli words whenever the rotation generator and the transformed Pauli word do not commute. And if the rotation generator commutes with the transformed Pauli word, the gate will of course leave the word unchanged.

3.1 Example

As an example, we can compute the action of RZ on the Pauli X operator $H' = R_Z^\dagger(\theta) X R_Z(\theta)$, here, we have Z as the generator of Pauli Z rotation group in su(2). We know that $[Z, X] = 2iY$ anti-commute,then this implies H' will map a Pauli word X into two Pauli words X and Y.

$$H' = R_Z^\dagger(\theta) X R_Z(\theta) = \cos(\theta) X - \sin(\theta) Y$$

4 Comments

As we can see, if we evolve P with Clifford Gates in the Heisenberg picture, then we will ending up with the same amount of Pauli gates and they will belong to the Clifford group as well. However, if we evolve it with non-Clifford gates and it is anti-commute with P, then it will split into two terms. Pictorially, this is where the split happen like a binary tree. As we go deeper into the tree, the number of terms scales exponentially, that's why we need tool such as Pauli propagation to truncate at some depth k.

They also showed that high-weight Pauli contribute little to the overall expectation value calculations. "We observe an exponentially decaying contribution of high-weight Pauli operators... The inset shows the expected contribution of all operators per weight... we observe an exponentially decaying contribution of high-weight Pauli operators."

5 Truncating the Pauli propagation

Each Hamiltonian H can be decomposed into Pauli strings, and in the worse case, the number of Pauli strings scaled 4^N which is a lot. We can use technique such as truncating the Pauli string to make the runtime more feasible.

Requirements for truncation:

- The unitary U has to be locally scrambling

- The scrambling layer has to be shallow. This will make the branching small. What I meant by branching is $U^\dagger P U$ will transform P a Pauli word into n Pauli words. We hope n is small, and to make n small, we have to make sure that scrambling layer to be shallow.

The algorithm is basically do the Heisenberg picture transformations to all the Pauli strings in the Hamiltonian and then evolve each of them in the Heisenberg picture. Define a parameter k which is the truncation parameter. we are going to delete any P' has a length greater than k. Note, we don't do truncate for single qubit rotation this means $R_p(\theta)^\dagger P R_p(\theta)$. This will transform P into more Pauli terms. However, when we transform $CNOT * P * CNOT$, that's when we are going to do truncation based on the parameter k.

6 Locally Scrambling

Locally scrambling is a concept from quantum information and many-body physics that refers to how quantum information becomes delocalized (or spread out) over a system, but only within a limited spatial region (locally), rather than over the whole system.

1. Scrambling refers to how quickly and thoroughly information about a local input (say, a qubit or small subsystem) becomes hidden or "scrambled" across many degrees of freedom in a quantum system. This makes it essentially irretrievable without accessing a large portion of the system.

2. Locally scrambling refers to this process happening within a finite or small region of the system. The information becomes entangled with nearby degrees of freedom but hasn't yet spread out over the entire system.