HQS QUANTUM SIMULATIONS

Mitigation methods

An Overview

Oliver Hahn, Sebastian Zanker, Michael Marthaler

August 12, 2019

Contents

1 Quantum algorithms and error mitigation

In this summary we discuss four mitigation methods. Two of the methods are compatible with any arbitrary quantum algorithm, two mitigation methods are specific to material simulation:

- Extrapolation: Useful for any quantum algorithm.
- Probabilistic error cancellation: Useful for any quantum algorithm.
- Quantum subspace expansion: Only useful for material simulation.
- Symmetry verification: In principle it has been developed for material simulation but could potentially also be useful for algorithms like QAOA.

The error mitigation will be implemented as a part of the PlanQK platform. A user of the PlanQK platform can then choose an algorithm and additionally an error mitigation method. Not all error mitigation methods can be combined with each other and some could be particularly useful for some hardware platforms. If possible the PlanQK platform will include a guidance for costumers, which error mitigation should be combined with which hardware.

In the coming sections we will summarize various mitigation methods which have been discussed in the literature. For each of the methods HQS will give an estimate for the strengths and the weaknesses of the method.

2 Extrapolation

2.1 Idea

The methods discussed in this section can, given a few calculations on a quantum computer, reduce the noise by extrapolating into the low noise regime. The advantage is, that these methods can be applied after the calculations are done and thus only need more measurements and not necessarily additional gates. A sketch of this idea is shown in 1. This idea first appeared in Refs. [17] and [7] and relies on the ability to tune the error probabilities or decoherence rate respectively.

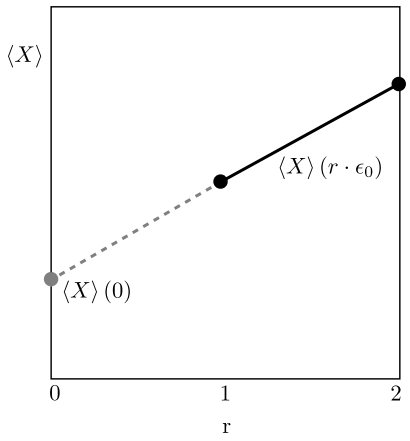


Figure 1: A schematic representation of error reduction by linear extrapolation. The noise free expectation value $\langle X \rangle (0)$ can be approximated by measuring the real noisy expectation value $\langle X \rangle (r\epsilon_0)$ for different $r \geq 1$ and subsequent linear extrapolation to $r \to 0$.

2.2 Richardson Deferred Approach & Exponential Variant

2.2.1 Description, summary and important equations

The Richardson Deferred Approach is a general method for extrapolation, not limited to quantum computing. A in-depth treatment can be found in ref. [14].
The expectation value obtained by a noisy calculation on a quantum computer can be written as an expansion in the error probability ϵ,

$$\langle X \rangle (\epsilon) = \langle X \rangle^{(0)} + \sum_{k=1}^{n} a_k \epsilon^k, \tag{1}$$

where a_k are the coefficients of a Taylor expansion and $\langle X \rangle^{(0)}$ is the noiseless result fulfilling the relation:

$$\lim_{\epsilon \to 0} \langle X \rangle (\epsilon) = \langle X \rangle^{(0)} . \tag{2}$$

A derivation using the Master equation in Lindblad-form can be found in the supplemental material of [17].
The result of a calculation on a quantum computer would give $\langle X \rangle (\epsilon_0)$, where ϵ_0 is the lowest achievable error probability. In principle we should always be able to increase the

error probability by increasing the time to implement and operation and therefore we run the algorithm with different decoherence rates:

$$\epsilon_i = c_i \epsilon_0. \tag{3}$$

From this definition we obviously see that $c_0 = 1$. For each run we obtain $\langle X \rangle (\epsilon_i) = \langle X \rangle (c_1 \epsilon_0)$ and we use these results to construct another expansion:

$$\langle \tilde{X} \rangle^{(n)} (\epsilon) = \sum_{i=0}^{n} \mu_i \langle X \rangle (c_i \epsilon). \tag{4}$$

By comparing this expansion with the expansion in eq. (1) we find that the coefficients have to fulfill the following relations for $k = 1...n$

$$\left.\begin{array}{l} \sum_{i=0}^{n} \mu_i = 1 \\ \sum_{i=0}^{n} \mu_i c_i^k = 0 \end{array}\right\} \rightarrow \mu_i = \prod_{m \neq i} \frac{c_m}{c_i - c_m}, \tag{5}$$

to extrapolate to $\langle \tilde{X} \rangle^{(n)} (\epsilon) \approx \langle X \rangle^{(0)}$. By choosing the coefficients this way, the error to the n-th order cancels out.
A common series chosen for the rescaling parameter[14] is

$$c_i = 2^i, \tag{6}$$

for $i = 0, 1, 2, ...$ defining the order of the extrapolation by the number of points used.
A measure for the stability of this method is given by [14]:

$$\mu = \sum_{i=0}^{n} |\mu_i| \tag{7}$$

being as close to one as possible. Therefore depends heavily on the progression of c_i.

For algorithmic errors (Trotter errors) the Richardson approach can be applied as well by expanding by $\frac{1}{N}$ with N being the number of steps [11].

Exponential version of the extrapolation method
The idea of extrapolation can be extended [4] by choosing an exponential ansatz:

$$\langle X \rangle (\epsilon_0) = e^{\epsilon_0} \langle X \rangle (0), \tag{8}$$
$$\langle X \rangle (r\epsilon_0) = e^{r\epsilon_0} \langle X \rangle (0). \tag{9}$$

This is motivated by an argument presented in [4]. Namely that the impact of noise is not proportional to a polynomial but to an exponential decay in the error rate. Usually $r = 2$ is chosen.

A comparison to the Richardson deferred approach can be seen in figure 2a. For small decoherence rates both extrapolation methods lead to an improvement of the energies, but the exponential method is very unstable for intermediate rates, therefore it is better to use the Richardson deferred approach instead.

2.2.2 Strengths

- Very flexible
- does not need additional quantum resources
- does not depend on the exact knowledge of the error model
- post processing
- can easily be used with other methods

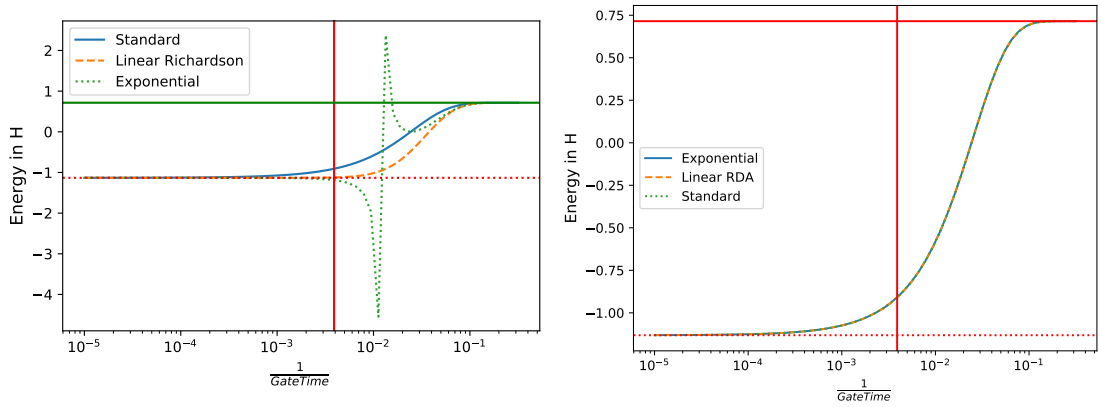
2.2.3 Weaknesses

- only improves results for small errors
- higher orders can lead to unstable behavior
- depends on progression c_i
- can not be implemented easily during the optimization loop (entwined VQE)

2.3 Hypersurface

2.3.1 Description, summary and important equations

The extrapolation method in the previous section is equivalent to using a polynomial function to fit the values with different error strengths. The hyper-surface extrapolation[5] is like a multidimensional generalization of the Richardson Extrapolation approach and allows for any numbers of noise sources $\{\gamma_i\}$ instead of a single global source. Similar to



(a) Comparison between extrapolation methods

(b) Usage within Optimization

Figure 2: The comparison between different extrapolation methods are shown. The decoherence model used is amplitude damping and simulated was H_2. The dotted line is the lowest eigenvalue of the molecular Hamiltonian and the green horizontal line the eigenvalue for zero electrons. The red vertical line is the decoherence rate of a current IBM qubit chip. In 2a extrapolation with different methods using the results of a VQE-UCCSD calculation of H_2 is shown. The exponential variant leads to very unstable behavior. The Richardson extrapolation methods leads to always lower or equal results to the standard VQE-UCCSD calculation. With the error rate of IBM chip the noisy result can be corrected relatively well by both of the methods. In 2b the application of extrapolation methods within the optimization loop is shown. It makes no difference if the the approach is used during the optimization cycle or with the results obtained after convergence, since within one optimization cycle the same variational parameter $\vec{\theta}$ is used.

the other extrapolation methods, the calculation is repeated with varying $\{\gamma_i\}$.
For several noisy sources we can write the expectation value of the observable as

$$\langle X \rangle = X_0 + \sum_i \gamma_i X_i + \sum_i \sum_j \gamma_i \gamma_j X_{ij} \tag{10}$$

where X_0 is the noise-free observable and X_i the effect of the noise rate i on the observable, etc.

$$\begin{bmatrix} 1 & \gamma_1^1 & \dots & \gamma_i^1 & \left(\gamma_1^1\right)^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \gamma_1^j & \dots & \gamma_i^j & \left(\gamma_1^j\right)^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \langle X \rangle^1 \\ \vdots \\ \langle X \rangle^j \\ \vdots \end{bmatrix} \tag{11}$$

with γ_i^j being the i-th noise rate of the j-th repetition and $\langle X \rangle^j$ the expectation value of the j-th run.
By standard least-squares procedures the $(X_0, X_1, ...)$ are then determined, with X_0 being the desired expectation value.

2.3.2 Pros and strengths

- natural for superconducting qubits since every qubit is different
- can thus combine the calculation with different qubits/devices
- works for cases that are limited by extending the algorithm

2.3.3 Cons and weaknesses

- requires more evaluations

2.4 Important papers

Idea: [17] [7]
Experimental realization:[6]
Exponential version: [4]
Error increased differently for different gates: [5]
Extrapolation for Algorithmic errors (e.g. Trotter error): [11]

3 Probabilistic error cancellation

3.1 Idea

The idea is to represent the ideal circuit as a quasi-probabilistic linear combination of noisy ones. The method effectively realizes an inverse transformation of the error.

3.2 Description, summary and important equations

We have a set $\Omega = \{\mathcal{O}_1, .., \mathcal{O}_m\}$ of noisy operations on a noisy N-qubit device. In principle these operators can be estimated, however measuring all relevant operators reliably might be a challenge. They also need to constitute a basis in the space of trace preserving completely positive maps (tcpc) or quantum channels, such that any ideal unitary gate can be expressed as a linear combination

$$\mathcal{U}(\rho) = \sum_{\alpha} \eta_\alpha \mathcal{O}_\alpha(\rho) \tag{12}$$

where η_α gives a quasi-probability distribution.
We define the expectation value of the operator A on the state produced by a noisy circuit $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_L)$ of length L

$$E(\boldsymbol{\alpha}) = \mathrm{Tr}\left[A\mathcal{O}_{\boldsymbol{\alpha}}\left(|0\rangle\langle 0|^{\otimes n}\right)\right] \tag{13}$$

with the quasi probability decomposition

$$\mathcal{U}_\beta = \sum_{\boldsymbol{\alpha}} \eta_{\boldsymbol{\alpha}} \mathcal{O}_{\boldsymbol{\alpha}} \tag{14}$$

with coefficients $\eta_{\boldsymbol{\alpha}}$ which we will try to estimate.

A noisy basis Ω simulates an ideal circuit β if there exists a probability distribution $P_\beta(\boldsymbol{\alpha})$ with coefficients $\sigma_\beta = \pm 1$, such that

$$\mathcal{U}_\beta = \gamma_\beta \sum_{\boldsymbol{\alpha}\in\Omega_L} P_\beta(\boldsymbol{\alpha})\sigma_\beta(\boldsymbol{\alpha})\mathcal{O}_{\boldsymbol{\alpha}}. \tag{15}$$

which is just the decomposition with

$$\gamma_\beta = \sum_{\boldsymbol{\alpha}} |\eta_{\boldsymbol{\alpha}}| \tag{16}$$

$$P_\beta(\boldsymbol{\alpha}) = \frac{|\eta_{\boldsymbol{\alpha}}|}{\gamma_\beta} \tag{17}$$

$$\sigma_\beta(\boldsymbol{\alpha}) = \mathrm{sign}(\eta_{\boldsymbol{\alpha}}). \tag{18}$$

Then we can write for the noiseless estimate

$$E^*(\boldsymbol{\beta}) = \gamma_{\boldsymbol{\beta}} \sum_{\boldsymbol{\alpha}\in\Omega_L} P_{\boldsymbol{\beta}}(\boldsymbol{\alpha})\sigma_{\boldsymbol{\beta}}(\boldsymbol{\alpha})E(\boldsymbol{\alpha}) \tag{19}$$

To compensation the errors, we have to fulfill the following relation

$$\mathcal{U}_{\boldsymbol{\beta}} = \sum_{\boldsymbol{\alpha}} \eta_{\boldsymbol{\alpha}}\mathcal{O}_{\boldsymbol{\alpha}} \tag{20}$$

$$= \sum_{\boldsymbol{\alpha}} \eta_{\boldsymbol{\alpha}}\mathcal{D}V_{\boldsymbol{\alpha}} \tag{21}$$

$$= \mathcal{D} \underbrace{\sum_{\boldsymbol{\alpha}} \eta_{\boldsymbol{\alpha}}\tilde{V}_{\boldsymbol{\alpha}}}_{\mathcal{D}^{-1}} \mathcal{U}_{\boldsymbol{\beta}} \tag{22}$$

with $\tilde{V}_{\boldsymbol{\alpha}}$ being a basis in the appropriate space. In section 3.4 we talk about appropriate basis operations.

3.3 Decomposition of Noise Free Circuit

Let's consider decomposition that can be constructed for each gate individually. $\mathcal{U}_\beta$ is a ideal gate. Let $\mathcal{O}_1, ..., \mathcal{O}_p \in \Omega$ be the list of all noisy operations whose support is contained in the support of $\mathcal{U}_\beta$.
For real variable μ_i, ν_j the decomposition has to fulfill the following requirements

$$\text{minimize} \sum_{\alpha=1}^{p} \mu_\alpha \tag{23}$$

$$\eta_\alpha \leq \mu_\alpha \tag{24}$$

$$-\eta_\alpha \leq \mu_\alpha \tag{25}$$

$$\mathcal{U}_\beta = \sum_{\alpha=1}^{p} \eta_\alpha \mathcal{O}_\alpha \tag{26}$$

Let $\{\mu_\alpha, \eta_\alpha\}$ be the optimal solution. Then $\mu_\alpha = |\eta_\alpha|$ (otherwise objective function can be decreased). If we define the following properties as before

$$\gamma_\beta = \sum_{\alpha=1}^{p} \mu_\alpha \tag{27}$$

$$P_\beta(\alpha) = \frac{\mu_\alpha}{\gamma_\beta} \tag{28}$$

$$\sigma_\beta(\alpha) = \text{sgn}(\eta_\alpha) \tag{29}$$

then we can rewrite $\mathcal{U}_\beta$ as sampling

$$\mathcal{U}_\beta = \sum_{\alpha=1}^{p} P_\beta(\alpha)\sigma_\beta(\alpha)\mathcal{O}_\alpha \tag{30}$$

γ_β^2 determines the simulation overhead. The ideal circuit can then be written as a product of this gate-wise decompositions. Usually $\mathcal{U}_\beta$ contains at max only two-qubit gates, which allows us to restrict the decomposition on acting on at most two qubits. If not all gates can be measured, the following protocol can be used

1. Start from ideal circuit $\mathcal{O}_\alpha = \mathcal{U}_\beta$
2. Modify $\mathcal{O}_\alpha$ by adding a basis operation after each single qubit gate with probability p_2. The gate is unchanged with p_1
3. Modify $\mathcal{O}_\alpha$ by adding a two-qubit basis operation after reach two-qubit gate with p_2

The resulting circuit is then implemented on a noisy device and linear combine several variants of the circuit.

Example: Depolarizing Noise

We define $\mathcal{D}_k$ as a k-qubit depolarizing channel.

$$\mathcal{D}_k(\rho) = (1 - \frac{3}{4}p)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z) \tag{31}$$

Then we define the set of noisy operations as

$$\mathcal{O}_\alpha = \mathcal{D}_k \mathcal{P}\mathcal{U} \tag{32}$$

with $\mathcal{U}$ being an ideal gate and $\mathbb{P}$ a k-qubit Pauli tcpc map.
Given the requirements defined above, the requirements are equivalent to implement the inverse map

$$\mathcal{D}_1^{-1} = \eta_1 \mathbb{1} + \eta_2 X + \eta_3 Y + \eta_4 Z \tag{33}$$

TABLE I. Sixteen basis operations. Gates $[R_x]$ and $[R_y]$ can be derived from $[H]$ and $[S]$, and other operations can be derived from $[\pi]$, $[R_x]$, and $[R_y]$.

1	$[\mathbb{1}]$ (no operation)
2	$[\sigma^x] = [R_x]^2$
3	$[\sigma^y] = [R_x]^2[R_z]^2$
4	$[\sigma^z] = [R_z]^2$
5	$[R_x] = [(1/\sqrt{2})(\mathbb{1} + i\sigma^x)] = [H][S]^3[H]$
6	$[R_y] = [(1/\sqrt{2})(\mathbb{1} + i\sigma^y)] = [R_z]^3[R_x][R_z]$
7	$[R_z] = [(1/\sqrt{2})(\mathbb{1} + i\sigma^z)] = [S]^3$
8	$[R_{yz}] = [(1/\sqrt{2})(\sigma^y + \sigma^z)] = [R_x][R_z]^2$
9	$[R_{zx}] = [(1/\sqrt{2})(\sigma^z + \sigma^x)] = [R_z][R_x][R_z]$
10	$[R_{xy}] = [(1/\sqrt{2})(\sigma^x + \sigma^y)] = [R_x]^2[R_z]$
11	$[\pi_x] = [\frac{1}{2}(\mathbb{1} + \sigma^x)] = [R_z]^3[R_x]^3[\pi][R_x][R_z]$
12	$[\pi_y] = [\frac{1}{2}(\mathbb{1} + \sigma^y)] = [R_x][\pi][R_x]^3$
13	$[\pi_z] = [\frac{1}{2}(\mathbb{1} + \sigma^z)] = [\pi]$
14	$[\pi_{yz}] = [\frac{1}{2}(\sigma^y + i\sigma^z)] = [R_z]^3[R_x]^3[\pi][R_x]^3[R_z]$
15	$[\pi_{zx}] = [\frac{1}{2}(\sigma^z + i\sigma^x)] = [R_x][\pi][R_x]^3[R_z]^2$
16	$[\pi_{xy}] = [\frac{1}{2}(\sigma^x + i\sigma^y)] = [\pi][R_x]^2$

Figure 3: Table of Basis Operations

The solution minimizing $\sum_\alpha |\eta_\alpha|$ is then $\eta_1 = 1 + \frac{3\epsilon}{4(1-\epsilon)}$ and $\eta_\alpha = \frac{-p}{4(1-p)}$ with $\gamma_\beta = \frac{1+0.5p}{1-p}$. Or simpler to see

$$\mathcal{D}_1^{-1} = \left(1 + \frac{4p}{4(1-p)}\right)\rho - \frac{p}{4(1-p)}(X\rho X + Y\rho Y + Z\rho Z) \tag{34}$$

$$= \gamma_\beta[p_1\rho + p_2(X\rho X + Y\rho Y + Z\rho Z)] \tag{35}$$

with $p_1 = \frac{4-p}{2p+4}$ $p_2 = \frac{-p}{2p+4}$. The noiseless expectation value can then be implemented by measuring the terms of

$$O_0 = \mathrm{Tr}\left[O\mathcal{D}^{-1}(\rho)\right] \tag{36}$$

Two qubit gates can be built similarly as a linear combination of two-qubit Pauli maps. Amplitude Damping is shown in [17] as well.

3.4 Basis Operations

Any single noisy qubit operation is effectively 4×4 matrix, because we need superoperators to describe noisy operations. It can be expressed as a linear combination of 16 basis

operations $\mathcal{B}_i$ and an example is shown in fig. 3. These basis operations then need to be decomposed into the gate sat natural to the quantum device. This is done for in the table for the Hadamard Gate $[H] = \frac{1}{\sqrt{2}}(\sigma^x + \sigma^z)$ and the phase gate $[S] = \frac{1}{\sqrt{2}}(\mathbb{1} - +i\sigma^z)$. For small errors, these 16 basis operations should still be linearly independent, the set of basis operations is still universal. For multi-qubit operations tensor products of these basis operations are used. These procedures are taken from [4].
Another method introduced in this paper is the compensation method

$$\mathcal{O}^0 = \lambda\mathcal{O} + \sum_i q_i \mathcal{B}_i \tag{37}$$

with λ being an arbitrary real number . For linearly independent basis operations the decomposition is unique , when λ is determined (can be found by optimization).

3.5 Strengths

- Mitigates error
- Works in principle in the limit of strong noise

3.6 Weaknesses

- full knowledge of the noise model
- has an simulation overhead/ increases the variance
- works for general Markovian noise
- for practicality reason assume only errors on the qubits the noisy gate is acting on

3.7 Important papers

Idea:[17]
Experimental realization: [15] Basis set for decomposition: [4]
Paper on quasi-prob. decomposition: [12]

4 Quantum Subspace Expansion

4.1 Idea

This method has been developed for material simulation in particular as a method to be combined with the Variational Quantum Eigensolver (VQE) which uses the UCCSD Ansatz [10]. The goal of VQE-UCCSD is to find the ground state of a electronic structure of a material. The hope is that by looking beyond the ground state and expanding to the linear subspace around the noisy resulting state of the calculation, the overlap between the true VQE-UCCSD groundstate and the noisy VQE-UCCSD groundstate can be determined.

4.2 Description, summary and important equations

The linear response subspace is spanned by expanding about a reference state $|\Psi_T\rangle$, in our case the resulting variational state of the calculation.
The linear subspace is spanned by the vectors

$$c_i^\dagger c_j \, |\Psi_T\rangle \tag{38}$$

where c_j is an annihilation operator removes an electron from orbital j and $c_i^\dagger$ is a creation operator which creates an electron in orbital i. We are therefore spanning a subspace of single excitations given the reference state. To get the new groundstate energy, we have to solve the generalized eigenvalue problem

$$H_{LR}C = S_{LR}CE, \tag{39}$$

with $H_{ijkl} = \langle\Psi_T| \, c_i^\dagger c_j H c_k^\dagger c_l \, |\Psi_T\rangle$ and $S_{ijkl} = \langle\Psi_T| \, c_i^\dagger c_j c_k^\dagger c_l \, |\Psi_T\rangle$ are measured on the quantum computer, which are at maximum 4-RDM (electron densities). To increase the accuracy, virtual states can be added and be treated by perturbation theory, please refer to [16].

Fixing Symmetries

The symmetries in the subspace with symmetry operator O can be enforced by calculating H_{LR} and O_{LR}. The eigenvectors with the desired symmetry of O_{LR} can then be used to project H_{LR} to the correct symmetry-subspace and then diagonal H_{LR} there. A other way to enforce symmetries will be discussed in the next section, as well as a connection between the two methods is discussed.

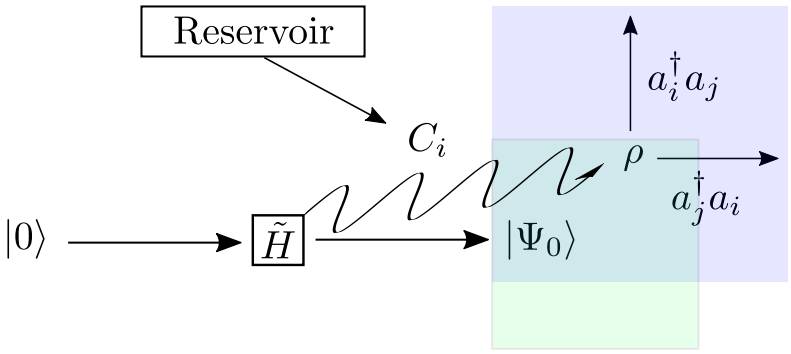


Figure 4: A schematic representation of the Quantum Subspace Expansion (QSE). The quantum device is initialized in the groundstate $|0\rangle$ and then the algorithm is started symbolized by $\tilde{H}$. Without coupling to the reservoir the state $|\Psi_0\rangle$ is obtained. Coupling to the reservoir with the collapse operators C_i leads to a different states ρ after the algorithm has run. By expanding the resulting trial state in a linear subspace, it is possible to determine a better estimation of the groundstate as well as excited states.

4.3 Strengths

- Excited states
- Physical argument: dominant contribution in a linear response theory of local time-dependent perturbations

4.4 Weaknesses

- groundstate needs to be in the subspace (e.g.H_2 with dephasing)
- higher order expansion become quickly unfeasible

4.5 Important papers

Idea: [10]
Experimental realization: [3]
Interesting ideas with QSE to increase accuracy: [16]

5 Symmetry verification

5.1 Idea

If it is our goal to create an eigenstate to a particular Hamiltonian H, we can use symmetries of the Hamiltonian to verify the result. We use the $\mathbb{Z}_2$ symmetries of the Hamiltonian after the mapping to Pauli-operators to project the state after the noisy computation in the correct eigenspace spanned by the symmetries.

5.2 Description, summary and important equations

Symmetries are defined by the commutator of a symmetry operator S_i and the Hamiltonian,

$$[H, S_i] = 0. \tag{40}$$

In general the symmetries can be easily be computed by the approach described in [2]. The Hamiltonian H is block diagonal within the eigenspaces of S_i as shown in fig-5. The trial state remains within the same subspace during an ideal computation. The

symmetry operator $S_i \in \mathcal{P}_N$, with $\mathcal{P}_N$ being the N qubit Pauli group, has the eigenvalues $s_i = \pm 1$ and the projector into the eigenspace of s_i can be written as:

$$M_{s_i} = \frac{1}{2}(1 + s_i S_i), \tag{41}$$

The calculation of eigenstate of the Hamiltonian on a noisy quantum computer may shift the state ρ outside of the target eigenspace s_i of S_i. By projection back into the said eigenspace, the overlap with the unperturbed trial state can be increased. Then the state after the subspace projection is given as:

$$\rho_{s_i} = \frac{M_{s_i}\rho M_{s_i}}{\mathrm{Tr}[M_{s_i}\rho]}. \tag{42}$$

Symmetry verification can be done with an ancilla, in-line or with post-selection[1]. In this paper a method is introduced to insert symmetries by extending the Hamiltonian by one qubit and transforming the Pauli operators.

In [8] a approach with the corresponding circuits is shown to measure parities and spin/particle number with an ancilla during the calculation. If an error occurred, the run is discarded.

These symmetries commute with dephasing collapse operators and thus cannot mitigate these errors. By rotating the problem

$$\rho_R = R\rho_0 R^\dagger, \tag{43}$$
$$H_R = RHR^\dagger, \tag{44}$$
$$U_R(\theta) = RU(\theta)R^\dagger. \tag{45}$$

one is able to use symmetry verification to mitigate dephasing. But this needs to be used with caution, since in many cases makes the effect of errors worse (For big off-diagonal elements and small rates, it is often advantageous to rotate for dephasing).

5.3 Connection to QSE

For symmetry verification via post selection the formula is given by

$$\mathrm{Tr}\left[H\rho_s\right] = \mathrm{Tr}\left[H\frac{M_s\rho M_s}{\mathrm{Tr}\left[M_s\rho\right]}\right] \tag{46}$$
$$= \frac{\mathrm{Tr}\left[H\rho\right] + s\,\mathrm{Tr}\left[HS\rho\right]}{1 + s\,\mathrm{Tr}\left[S\rho\right]}\,. \tag{47}$$

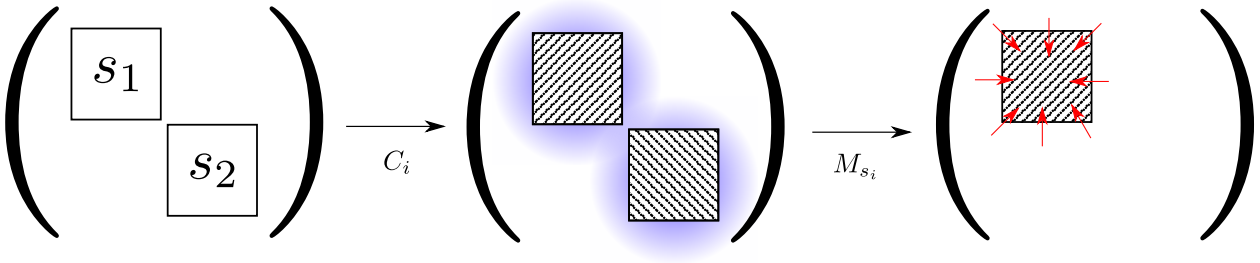


Figure 5: A schematic representation of the symmetry verification. The eigenstates of the blockdiagonal Hamiltonian can partitioned into target eigenspaces of the symmetry S_i with eigenvalue s_i symbolized by the blocks. The system will start in a certain target subspace of s_i and the calculation without decoherence will stay in the target subspace. Decoherence symbolized by the collapse operators C_i will move the states outside of the correct block during the computation. By projecting the state back with the projection operator M_{s_i} into the correct subspace s_i of the symmetry S_i, the overlap with the true unperturbed state is increased.

Here we dropped the index i. For QSE with expansion operators $\{E_i\}$ and

$$H_{ij} = \mathrm{Tr}\left[H E_i \rho E_j^\dagger\right] \tag{48}$$

$$S_{ij} = \mathrm{Tr}\left[E_i \rho E_j^\dagger\right], \tag{49}$$

symmetry verification is equivalent to QSE if instead of the $\{E_i\}$ the set $\{\mathbb{1}, S\}$ is chosen, but the correct symmetry subspace is chosen automatically.
Symmetry verification can also be expressed in the stabilizer formalism with the stabilizer group $\mathcal{S} \subset P_n$ with generator S. The logical states are the $+1$ eigenstates of the stabilizer group. Another formulation is that the code space is the degenerate groundstate of

$$H_c = -\sum_{M_i \in \mathcal{M}} M_i \tag{50}$$

with M_i being the check operators $\in \mathcal{S}$ used to deduce the error and anticommuting with the error.

$$S \subseteq \mathcal{M} \subseteq \mathcal{S} \tag{51}$$

Traditionally M_i are measured and the information is used to correct the error. But with NISQ devices no on-the-fly measurements. Thus use the M_i to define projections

into the +1 subspace

$$P_i = \frac{1}{2}(\mathbb{1} + M_i) \tag{52}$$

Thus with $S_iS_j = S_k$ and therefore $P_iP_jP_k = P_iP_j$

$$\overline{P} = \prod_i^m P_i = \prod_{S_i \in S}^m \frac{1}{2}(\mathbb{1} + S_i) = \frac{1}{2^m} \sum_{M_i \in \mathcal{S}} M_i \tag{53}$$

with $m = n - k$ being the number of generators of the stabilizer group.
Then the expectation value of a observable is given by

$$\langle \Gamma \rangle = \frac{1}{c} \operatorname{Tr}\left[\overline{P}\rho\overline{P}^\dagger\Gamma\right] \tag{54}$$

$$= \frac{1}{c2^m} \sum_{i,j} \operatorname{Tr}\left[\rho M_i^\dagger \Gamma M_j\right] \tag{55}$$

$$= \frac{1}{c2^m} \sum_i \operatorname{Tr}\left[\rho\Gamma M_i\right] \tag{56}$$

$$c = \operatorname{Tr}\left[\overline{P}\rho\right] \tag{57}$$

It is advised to use stochastic sampling of the terms for better scaling as described in [9] .
Another idea is to project onto the error subspace corresponding to error $\mathcal{E}_i$ by

$$\overline{P}_{\mathcal{E}_i} = \prod_j \frac{1}{2}\left(\mathbb{1} + (-1)^{s_j^i} S_j\right) \tag{58}$$

with $s_j^i \in \{0, 1\}$ being the error syndrome associated with error $\mathcal{E}_i$ and generator S_j. After projection one can use the recovery operation R_i to map the state back to the code space. Considering every correctable error (including the identity) we get

$$\langle \Gamma \rangle \frac{1}{c} \sum_i \operatorname{Tr}\left[R_i \overline{P}_{\mathcal{E}_i} \rho \overline{P}_{\mathcal{E}_i}^\dagger \Gamma\right] \tag{59}$$

The advantage of this approach is that the variance is smaller, however less Pauli errors can be corrected ($\frac{d-1}{2}$ to $d-1$).

$$\overline{P}_c = \sum_i c_i M_i \tag{60}$$

not $\overline{P} \propto \prod_i (\mathbb{1} + S_i)$ Find c_i

$$\min_{c_i} \mathrm{Tr}\left[\overline{P}_c \rho \overline{P}_c^\dagger H_c\right] \tag{61}$$

with $\mathrm{Tr}\left[\overline{P}_c \rho \overline{P}_c^\dagger\right] = 1$ Solution is given by generalized eigenvalue problem

$$HC = SCE \tag{62}$$

$$H_{ij} = \mathrm{Tr}\left[M_i^\dagger H_c M_j \rho\right] \tag{63}$$

$$S_{ij} = \mathrm{Tr}\left[M_i^\dagger M_j \rho\right] \tag{64}$$

If M_i are build from generators, the solution coincides with the previous result. Using this together with the encoded problem Hamiltonian by expanding to a set of operators and using the symmetries, logical errors can be mitigated as well. In addition to error mitigation this approach can be used to study error correction codes.

5.4 Strengths

- Overlap between exact state and noisy state is bigger
- resulting state has correct symmetries

5.5 Weaknesses

- does not trivially work for dephasing (one needs to be careful with rotating)

5.6 Important papers

Idea: [1]
Experimental realization: [13]
Error mitigated digital quantum simulation: [8]
QSE and Symmetry verification: [9]

References

[1] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien. Low-cost error mitigation by symmetry verification. *Phys. Rev. A*, 98:062339, 12 2018.

[2] Sergey Bravyi, Jay M. Gambetta, Antonio Mezzacapo, and Kristan Temme. Tapering off qubits to simulate fermionic hamiltonians. *arXiv:1701.08213*, 2017.

[3] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X*, 8:011021, 02 2018.

[4] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications. *Phys. Rev. X*, 8:031027, 07 2018.

[5] Suguru Endo, Qi Zhao, Ying Li, Simon Benjamin, and Xiao Yuan. Mitigating algorithmic errors in a hamiltonian simulation. *Phys. Rev. A*, 99:012334, 01 2019.

[6] Abhinav Kandala, Kristan Temme, Antonio D. Corcoles, Antonio Mezzacapo, and Jay M. Chow, Jerry M.and Gambetta. Error mitigation extends the computational reach of a noisy quantum processor. *Nature*, 567(7749):491–495, 2019.

[7] Ying Li and Simon C. Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X*, 7:021050, 06 2017.

[8] Sam McArdle, Xiao Yuan, and Simon Benjamin. Error mitigated digital quantum simulation. *arXiv:1807.02467*, 2018.

[9] Jarrod R. McClean, Zhang Jiang, Nicholas C. Rubin, Ryan Babbush, and Hartmut Neven. Decoding quantum errors with subspace expansions. *arXiv:1903.05786*, 2019.

[10] Jarrod R. McClean, Mollie E. Kimchi-Schwartz, Jonathan Carter, and Wibe A. de Jong. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A*, 95:042308, 04 2017.

[11] Matthew Otten and Stephen K. Gray. Recovering noise-free quantum observables. *Phys. Rev. A*, 99:012338, 01 2019.

[12] Hakop Pashayan, Joel J. Wallman, and Stephen D. Bartlett. Estimating outcome probabilities of quantum circuits using quasiprobabilities. *Phys. Rev. Lett.*, 115:070501, 08 2015.

[13] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo. Error mitigation by symmetry verification on a variational quantum eigensolver. *arXiv:1902.11258*, 2019.

[14] Avram Sidi. *Practical Extrapolation Methods: Theory and Applications*. Cambridge University Press, New York, NY, USA, 1st edition, 2002.

[15] Chao Song, Jing Cui, H. Wang, J. Hao, H. Feng, and Ying Li. Quantum computation with universal error mitigation on superconducting quantum processor. *arXiv:1812.10903*, 2018.

[16] Tyler Takeshita, Nicholas C. Rubin, Zhang Jiang, Eunseok Lee, Ryan Babbush, and Jarrod R. McClean. Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources. *arXiv:1902.10679*, 2019.

[17] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.*, 119:180509, 11 2017.