SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
Special Topic: Quantum Information

August 2023, Vol. 66 180508:1–180508:11
https://doi.org/10.1007/s11432-023-3786-1

Mitigating quantum errors via truncated Neumann series

Kun WANG*, Yu-Ao CHEN & Xin WANG

Institute for Quantum Computing, Baidu Research, Beijing 100193, China

Received 28 February 2023/Revised 6 May 2023/Accepted 26 May 2023/Published online 5 July 2023

Abstract Quantum gates and measurements on quantum hardware are inevitably subject to hardware imperfections that lead to quantum errors. Mitigating such unavoidable errors is crucial to explore the power of quantum hardware better. In this paper, we propose a unified framework that can mitigate quantum gate and measurement errors in computing quantum expectation values utilizing the truncated Neumann series. The essential idea is to cancel the effect of quantum error by approximating its inverse via linearly combining quantum errors of different orders produced by sequential applications of the quantum devices with carefully chosen coefficients. Remarkably, the estimation error decays exponentially in the truncated order, and the incurred error mitigation overhead is independent of the system size, as long as the noise resistance of the quantum device is moderate. We numerically test this framework for different quantum errors and find that the computation accuracy is substantially improved. Our framework possesses several vital advantages: it mitigates quantum gate and measurement errors in a unified manner, it neither assumes any error structure nor requires the tomography procedure to completely characterize the quantum errors, and most importantly, it is scalable. These advantages empower our quantum error mitigation framework to be efficient and practical and extend the ability of near-term quantum devices to deliver quantum applications.

Keywords near-term quantum devices, quantum error mitigation, quantum computing

Citation Wang K, Chen Y-A, Wang X. Mitigating quantum errors via truncated Neumann series. Sci China Inf Sci, 2023, 66(8): 180508, https://doi.org/10.1007/s11432-023-3786-1

1 Introduction

Quantum computers hold great promise for a variety of scientific and industrial applications [1–3]. Nonetheless, the challenges we face are still formidable. In the current noisy intermediate-scale quantum (NISQ) era [4], quantum computers introduce significant quantum errors that must be dealt with before performing any exhilarating tasks. Such errors occur either due to unwanted interactions of qubits with the environment or the physical imperfections of qubit initializations, quantum gates, and measurements [5–8]. The problem can be theoretically resolved with quantum error correction [9–13], which is far beyond the reach of NISQ quantum computers. This motivates the question of alleviating quantum errors and increasing the quantum computation accuracy without quantum error correction.

Quantum error mitigation [14] provides an inspirational solution to this question and has been experimentally implemented [15–18]. Typically, errors in a quantum device are classified into quantum gate and measurement errors. For gate errors, numerous techniques have been designed such as zero-noise extrapolations [14, 19–25], probabilistic error cancellations [14, 20, 26–29], subspace expansions [30–32], purification-based methods [33–39], learning-based methods [40–42], and many others [43–45]. For measurement errors, the focus is not as much as that on gate errors though measurement errors are significantly larger than gate errors on many quantum platforms [15, 46, 47]. A well-known strategy is to regard the measurement error as a classical noise model and handle it via classical post-processing [48–65]. We recommend the interested reader to [66–68] for more detailed surveys of this topic. However, most existing error mitigation techniques require complete characterization of the quantum device and are not scalable

* Corresponding author (email: wangkun28@baidu.com)

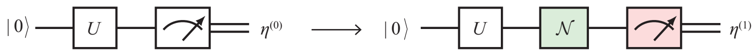

Figure 1 (Color online) Estimation of the expectation value $\mathrm{Tr}[O\rho]$ with the ideal quantum devices (left) and the noisy quantum devices (right).

in general. On the other hand, the techniques targeting one type of error are not directly applicable to the other at large.

We overcome these challenges by proposing a general error mitigation framework that can reduce both quantum gate and measurement errors in computing the expectation values of quantum observables. This framework does not require complete characterization of the quantum devices and is theoretically scalable. The essential idea is to effectively approximate the inverse of quantum error using the truncated Neumann series. Notably, the estimation accuracy of the expectation value is improved in the presence of quantum errors.

2 Computing the expectation value

A common quantum computation task is to estimate the expectation value $\mathrm{Tr}[O\rho]$ within a specified precision ε for a given quantum observable O and an n-qubit quantum state $\rho := U|\mathbf{0}\rangle\langle\mathbf{0}|U^\dagger$ generated by a quantum gate U with initial n-qubit state $|\mathbf{0}\rangle$. Without loss of generality, we may assume that O is diagonal in the computational basis and $\|O\|_2 \leqslant 1$ where $\|\cdot\|_2$ is the matrix 2-norm. This task is the building component of multifarious quantum algorithms, and notable practical examples are the variational quantum eigensolvers [69,70], quantum approximate optimization algorithm [71], and quantum machine learning [72, 73].

Ideally, $\mathrm{Tr}[O\rho]$ can be estimated in the following way. Consider M independent experiments where in each round we prepare the state ρ using U and measure each qubit in the computational basis as shown in Figure 1. Let $\boldsymbol{s}^m \in \{0,1\}^n$ be the outcome in the m-th round. Define the empirical mean value

$$\eta^{(0)} := \frac{1}{M}\sum_{m=1}^{M} O(\boldsymbol{s}^m), \tag{1}$$

where $O(\boldsymbol{x})$ is the $\boldsymbol{x}$-th diagonal element of O. Let $\mathrm{vec}(\rho)$ be the 2^n-dimensional column diagonal vector of ρ. Then [55]

$$E^{(0)} := \mathrm{E}[\eta^{(0)}] = \sum_{\boldsymbol{x}\in\{0,1\}^n} O(\boldsymbol{x})\langle\boldsymbol{x}|\,\mathrm{vec}(\rho) = \mathrm{Tr}[O\rho], \tag{2}$$

where $\mathrm{E}[X]$ is the expectation of the random variable X. Eq. (2) implies that $\eta^{(0)}$ is an unbiased estimator of $\mathrm{Tr}[O\rho]$. Furthermore, the standard deviation $\sigma(\eta^{(0)}) \leqslant 1/\sqrt{M}$. By Hoeffding's inequality [74], $M = 2\log(2/\delta)/\varepsilon^2$ would guarantee that $\eta^{(0)}$ approximates $\mathrm{Tr}[O\rho]$ within ε at probability greater than $1-\delta$, where δ is the confidence and all logarithms are in base 2 throughout this paper.

However, both the quantum gate and measurement of the quantum states suffer from Markovian errors inherent in quantum devices. For simplicity, we ignore errors in preparing the initial state, which can be accounted for by regulating noisy quantum gates after qubits initialization. For the quantum gate error, we assume that the overall evolution is modeled as the ideal gate evolution U followed by some quantum noisy channel $\mathcal{N}$ [14], i.e., the actual prepared state is $\mathcal{N}(\rho)$ rather than ρ. The gate error leads to the noisy expectation value $\mathrm{Tr}[O\mathcal{N}(\rho)]$. Experimentally, the noise channel $\mathcal{N}$ can be fully characterized via quantum process or gateset tomography [75]. For the measurement error, it was established that such error can be well understood using classical noise models [48–50]. Specifically, an n-qubit noisy measurement device can be characterized by an error matrix A of size $2^n \times 2^n$. The element in the $\boldsymbol{x}$-th row and $\boldsymbol{y}$-th column, $A_{\boldsymbol{x}\boldsymbol{y}}$, is the probability of obtaining an outcome $\boldsymbol{x}$ provided that the true outcome is $\boldsymbol{y}$. Experimentally, the error matrix can be learned via calibration [49].

Suppose now that we adopt the same procedure for computing $\eta^{(0)}$ and obtain the noisy estimator $\eta^{(1)}$ (cf. the right side of Figure 1), where superscript 1 indicates that the noisy devices are functioning. We prove in Appendix A that

$$E^{(1)} := \mathrm{E}[\eta^{(1)}] = \sum_{\boldsymbol{x}\in\{0,1\}^n} O(\boldsymbol{x})\langle\boldsymbol{x}|A\,\mathrm{vec}(\mathcal{N}(\rho)), \tag{3}$$

indicating that $\eta^{(1)}$ is no longer an estimator of $\mathrm{Tr}[O\rho]$. Comparing (2) and (3), we find that in the ideal case, the sampled distribution approximates $\mathrm{vec}(\rho)$ thanks to the weak law of large numbers, while in the noisy case, the sampled distribution approximates $A\,\mathrm{vec}(\mathcal{N}(\rho))$ due to both gate and measurement errors, leading to a bias in the estimator.

Our main result is a unified error mitigation framework that can alleviate the quantum errors identified by $\mathcal{N}$ and A leading to the noisy estimator $A\,\mathrm{vec}(\mathcal{N}(\rho))$. In the following, we first describe the general framework and then specialize it to deal with quantum gate and measurement errors.

3 Error mitigation via truncated Neumann series

3.1 General theory

Let $\mathbb{R}$ be the real field, $\boldsymbol{M}_d$ be the set of $d \times d$ real square matrices, and $I \in \boldsymbol{M}_d$ be the identity matrix. Let $f : \boldsymbol{M}_d \to \mathbb{R}$ be a linear function and $A \in \boldsymbol{M}_d$ be a matrix. We are interested in estimating $f(I)$ given access to $f(A)$. This abstract task encapsulates many important computational tasks in both classical and quantum computing including the expectation value estimation task described above. We show in the following proposition that, under certain conditions, the target $f(I)$ can be efficiently approximated via a linear combination of accessible terms $f(A^k)$ of different orders with carefully chosen coefficients, employing the Neumann series expansion [76, Theorem 4.20]. Since $I = AA^{-1}$, the essential idea is to simulate the effect of the inverse A^{-1} via a suitable truncated Neumann series. The proof is given in Appendix B.

Proposition 1. If $\|I - A\| < 1$ in some consistent norm, then

$$\left| f(I) - \sum_{k=1}^{K+1} c_K(k-1) f\left(A^k\right) \right| = \left| f\left((I-A)^{K+1}\right) \right|, \tag{4}$$

where the coefficient function is defined as

$$c_K(k) := (-1)^k \binom{K+1}{k+1} \tag{5}$$

and $\binom{n}{k}$ is the binomial coefficient.

Intuitively, Eq. (4) indicates that one may approximate the target value $f(I)$ using the first K truncated Neumann terms $f(A^k)$, if (1) the precondition $\|I - A\| < 1$ is satisfied, (2) the k-th order value $f(A^k)$ can be obtained in a similar way as that of $f(A)$, and (3) the remaining term $f((I - A)^{K+1})$ can be upper bounded theoretically. And, even better, if the upper bound decays exponentially with K, the approximation quickly converges. We show that with appropriate representations all these three conditions can be satisfied in the quantum error mitigation tasks. This idea has previously been applied for linear data detection in massive multiuser multiple-input multiple-output wireless systems [77].

3.2 Gate error mitigation (GEM)

We illustrate how to use the Neumann series framework to mitigate gate errors. For this aim, we first recall the Pauli transfer matrix (PTM) representation in Appendix C. In PTM, quantum states $|\rho\rangle\rangle$ and observables $\langle\langle O|$ are represented by vectors and quantum channels $[\mathcal{N}]$ are represented by real matrices. For n qubits, vectors and matrices are 4^n-dimensional. The expected value of the observable O in the state ρ going through the noisy quantum channel $\mathcal{N}$ reads as follows:

$$\mathrm{Tr}[O\mathcal{N}(\rho)] = \langle\langle O|[\mathcal{N}]|\rho\rangle\rangle. \tag{6}$$

Setting $f \equiv \mathrm{Tr}$ and $A \equiv [\mathcal{N}]$, the above task fits into the truncated Neumann series framework. Define the noise resistance of the quantum channel $\mathcal{N}$ as

$$\xi_g(\mathcal{N}) := \|I - [\mathcal{N}]\|_\infty, \tag{7}$$

where $\|\cdot\|_\infty$ is the matrix ∞-norm. We assume that $\xi_g(\mathcal{N}) < 1$, which is a sufficient condition so that Proposition 1 holds. For Pauli noise $\mathcal{N}$, $\xi_g < 1$ corresponds to the case that all Pauli eigenvalues must

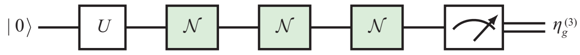

Figure 2 (Color online) Experimental setup for estimating $E_g^{(3)}$, where the noisy gate device (box in green) is executed 3 times sequentially.

be strictly positive [78, 79], which can be satisfied whenever the noise is weak [80]. We show in the following theorem that the approximation error (right hand side (RHS.) of (4)) can be exponentially upper bounded in terms of ξ_g. The proof is given in Appendix D.

Theorem 1. Assume that $\xi_g(\mathcal{N}) < 1$. For arbitrary positive integer K, it holds that

$$\left|\mathrm{Tr}[O\rho] - \sum_{k=1}^{K+1} c_K(k-1)E_g^{(k)}\right| \leqslant \|\langle\langle O|\|_\infty \, \xi_g^{K+1}, \tag{8}$$

where $E_g^{(k)} := \langle\langle O|[\mathcal{N}]^k|\rho\rangle\rangle$.

As evident from Theorem 1, the noise resistance ξ_g of $\mathcal{N}$ uniquely determines the number of terms required in the truncated Neumann series to approximate $\mathrm{Tr}[O\rho]$ to the desired precision. What is more, since $\xi_g < 1$, the approximation error decays exponentially in terms of K, indicating that small K suffices to reach high estimating accuracy. Thanks to the multiplicativity property of PTM, which states that the PTM of $\mathcal{N}^{\circ k}$ is exactly $[\mathcal{N}]^k$, each $E_g^{(k)}$ can be viewed as the noisy expectation value generated by the noisy gate device executed k times sequentially. Since measurement errors can be handled independently, we do not concern such errors in GEM. Let $\overline{E}_g := \sum_{k=1}^{K+1} c_K(k-1)E_g^{(k)}$. Theorem 1 inspires a systematic way to estimate the expectation value $\mathrm{Tr}[O\rho]$. Firstly, we choose K so that the RHS. of (8) evaluates to the desired precision ε, yielding the optimal gate truncated number

$$K_g = \left\lceil \frac{\log\varepsilon - \log\|\langle\langle O|\|_\infty}{\log \xi_g} - 1 \right\rceil. \tag{9}$$

Secondly, we compute $\overline{E}_g$ by estimating each $E_g^{(k)}$ and linearly combining them with coefficients c_K. Since $\overline{E}_g$ itself is only an ε-estimate of $\mathrm{Tr}[O\rho]$, it suffices to approximate $\overline{E}$ within an error ε. Motivated by the relation between $\eta^{(1)}$ and $E^{(1)}$ in (3), we propose the following procedure to estimate $E_g^{(k)}$ for arbitrary $1 \leqslant k \leqslant K+1$:

(1) Generate a quantum state ρ.

(2) Execute the channel $\mathcal{N}$ sequentially k times, yielding the final state $\mathcal{N}^{\circ k}(\rho)$. Measure the final state and collect the measurement outcome.

(3) Repeat the above two steps M rounds.

(4) Output the average $\eta_g^{(k)}$ as an estimate of $E_g^{(k)}$.

We claim that the average $\overline{\eta}_g := \sum_{k=1}^{K+1} c_K(k-1)\eta_g^{(k)}$ approximates $\mathrm{Tr}[O\rho]$ within error 2ε with high probability, i.e.,

$$\Pr\left\{|\mathrm{Tr}[O\rho] - \overline{\eta}_g| \leqslant 2\varepsilon\right\} \geqslant 1-\delta. \tag{10}$$

The proof is given in Appendix E. For illustrative purposes, we demonstrate in Figure 2 the experimental setup for estimating the noisy expectation value $E_g^{(3)}$, where the noisy device is repeated sequentially three times in each round. It is worth noting that the GEM method does not necessarily require the complete characterization of $\mathcal{N}$, as the only relevant quantity is the noise resistance ξ_g. In principle, this means that ξ_g can be estimated efficiently without resorting to time-consuming tomography procedures.

Comparison with existing results. The proposed GEM method shares similarities with the Richardson extrapolation-based error mitigation method [14, 24] in that both increase the effective noise level and then infer the noiseless case through classical reconstruction. As demonstrated in [24], Richardson extrapolation produces the same fitting coefficients $c_K(k)$ as defined in (5) when the noise scale factors are evenly spaced. However, our GEM method fundamentally differs from the Richardson extrapolation method as it assumes that the expectation value $E^{(1)}$ defined in (3) is influenced by a scalar noise factor and corresponds to a Taylor series approximation. This results in the inability to analytically determine the number of extrapolation terms. Conversely, our GEM method evaluates the impact of noise on the expectation value in its most general form and employs the truncated Neumann series to mitigate the noise effect at the matrix level. As a result, our method provides rigorous convergence rate bounds, allowing the optimal determination of truncated terms. In some respects, our GEM method is a multi-dimensional Richardson extrapolation based error mitigation method with rigorous convergence rate bounds. Nonetheless, our method has shortcomings when compared to the Richardson extrapolation method since it requires evenly spaced noise scale factors, thus preventing us from employing the standard identity insertion technique to increase the effective noise level effectively. Consequently, our method faces challenges in experimental implementation, which we will discuss shortly. In their recent work, Takagi et al. [27] established fundamental bounds on the reduction of computation error by error mitigation methods in terms of sampling overhead, which place universal performance limits on our proposed GEM method. Additionally, Cao et al. [81] introduced a scalable extrapolation approach to mitigate algorithmic errors in quantum optimization algorithms. We believe that our proposed GEM method, or the truncated Neumann series framework more generally, could also be extended to mitigate algorithmic errors, given the similarity between our method and the extrapolation-based error mitigation method.

Remarks on the experimental implementation. In practical quantum computing platforms, two-qubit gates typically suffer from error rates that are orders of magnitude higher than that of single-qubit gates [15, 46, 47]. Therefore, we may focus on mitigating errors induced by two-qubit gates such as the CNOT gate as in [21]. Specifically, to implement the power $\mathcal{N}^m$ physically, we first learn the noise model $\mathcal{N}$ of the CNOT gate to high accuracy and then carefully design a virtual identity gate whose noise model approximates $\mathcal{N}$. In this way, the power $\mathcal{N}^m$ can be experimentally accomplished by inserting the virtual identity gates after the noisy CNOT gate. This approach is particularly suitable for near-term quantum computers where two-qubit gate fidelities are limited and can be characterized.

3.3 Measurement error mitigation (MEM)

Recall that the error of a measurement device is well understood using classical noise models and is characterized by an error matrix A. It is straightforward to classically reverse the noise effects by multiplying the sampled distribution by the inversion A^{-1}. Nonetheless, there are several limitations of the direct inverse approach: (i) Completely characterizing A requires 2^n calibration setups and is not scalable; (ii) A may be singular which prevents direct inversion; and (iii) A^{-1} is hard to compute in general and might not be column stochastic, resulting unphysical estimates.

We manifest how to use the Neumann series framework to mitigate measurement errors while avoiding the limitations. Similar to the GEM case, the essential idea of MEM is to effectively simulate the inverse of the error matrix A, utilizing the truncated Neumann series. First of all, we define the noise resistance of the error matrix A as

$$\xi_m(A) := 2\left(1 - \min_{\boldsymbol{x}\in\{0,1\}^n} \langle\boldsymbol{x}|\, A\,|\boldsymbol{x}\rangle\right). \tag{11}$$

By definition, $1 - \xi_m/2$ is the minimal diagonal element of A. Intuitively, $\xi_m/2$ characterizes the measurement device's worst-case behavior since it is the maximal probability for which the true and actual outcomes mismatch. In the following, we assume $\xi_m < 1$, which is equivalent to the condition that the minimal diagonal element of A is larger than 0.5. This assumption is reasonable since otherwise the measurement device is too noisy to be applicable from the practical perspective. It is also a sufficient condition under which Proposition 1 holds. We deploy the truncated Neumann series framework to alleviate the measurement errors and obtain the following, in a similar favor of Theorem 1. The proof is given in Appendix F.

Theorem 2. Assume that $\xi_m < 1$. For arbitrary positive integer K, it holds that

$$\left|\mathrm{Tr}[O\rho] - \sum_{k=1}^{K+1} c_K(k-1)E_m^{(k)}\right| \leqslant \xi_m^{K+1}, \tag{12}$$

where

$$E_m^{(k)} := \sum_{\boldsymbol{x}\in\{0,1\}^n} O(\boldsymbol{x})\langle\boldsymbol{x}|A^k\,\mathrm{vec}(\rho). \tag{13}$$

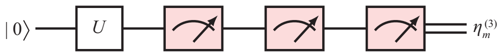

Figure 3 (Color online) Experimental setup for estimating $E_m^{(3)}$, in which the noisy measurement device (box in red) is executed 3 times sequentially.

Upper bounding the RHS. of (12) with the desired precision ε yields the optimal measurement truncated number

$$K_m = \left\lceil \frac{\log \varepsilon}{\log \xi_m} - 1 \right\rceil. \tag{14}$$

We can evaluate the noisy values $E_m^{(k)}$ up to the optimal order K_m in almost the same manner as we estimated $E_g^{(k)}$ in GEM. The different step is that we replace the single measurement with sequential measurements so that A^k appeared in (13) can be recovered:

(1) Generate a quantum state ρ.

(2) Using ρ as input, execute the noisy measurement device k times sequentially and collect the outcome produced by the final k-th measurement device.

(3) Repeat the above two steps M rounds.

(4) Output the average $\eta_m^{(k)}$ as an estimate of $E_m^{(k)}$.

Likewise, the average $\overline{\eta}_m := \sum_{k=1}^{K+1} c_K(k-1)\eta_m^{(k)}$ approximates $\mathrm{Tr}[O\rho]$ within error 2ε with a probability larger than $1-\delta$. The proof is the same as that of (10).

A crucial concept we introduce in the above MEM method is the sequential measurement. Roughly speaking, it means that we use the output of one measurement device as the input of the other. In Appendix G, we elaborate thoroughly on this concept and show that the classical noise model describing the sequential measurement repeating k times is indeed characterized by the error matrix A^k. For illustrative purpose, we demonstrate in Figure 3 the experimental setup for estimating $E_m^{(3)}$, where the measurement device is executed three times sequentially. Indeed, one can think of the rightmost $k-1$ measurements as implementing the calibration subroutine since they always have the computational basis states as inputs. In some sense, this is a dynamic calibration where we do not statically enumerate all computational bases as input states but dynamically prepare the input states based on the output information of the target state from the first measurement device. We note that our MEM method has been implemented in the quantum error processing toolkit developed on the Baidu Quantum Platform [82].

3.4 Resource analysis

When applying the truncated Neumann series framework to mitigate quantum errors, we repeat the quantum devices sequentially in different numbers of times, compute the noisy values, and linearly combine them to approximate the target. We use the number of quantum states consumed as the resource metric and analyze the complexity of the proposed methods GEM and MEM. Fundamentally, the resource costs of both methods are dominated by the optimal truncated number K – K_g (9) in GEM and K_m (14) in MEM – that determines the maximal number of truncated terms and the number of quantum states prepared to evaluate each truncated term. The detailed analysis has been given in Appendix E. Set $\Delta := \binom{2K+2}{K+1} - 1$. For each $1 \leqslant k \leqslant K$, we need $M = 2(K+1)\Delta\log(2/\delta)/\varepsilon^2$ copies of quantum states to achieve the desired accuracy ε and confidence δ. As so, the total number of quantum states consumed is roughly given by

$$M(K+1) = 2(K+1)^2\Delta\log(2/\delta)/\varepsilon^2 \approx 4^K\log(2/\delta)/\varepsilon^2, \tag{15}$$

where the approximation follows from Stirling's approximation. In other words, the number of quantum states consumed by GEM or MEM is much more than that of the ideal case by a factor of 4^K. We shall call the factor 4^K the error mitigation overhead with truncated Neumann series, characterizing the overall increased number of samples necessary to ensure a certain accuracy. At first glance, the exponential factor 4^K renders the error mitigation methods utilizing truncated Neumann series infeasible when K becomes large. However, we argue in the following that for near-term quantum devices with moderate noise resistance, K is quite small and thus is acceptable experimentally. In Figure 4, we plot the optimal truncated number $K = \lceil\log\varepsilon/\log\xi - 1\rceil$ as a function of the noise resistance ξ, where the error tolerance parameter is fixed as $\varepsilon = 0.01$. Notice that ξ can be either ξ_g (7) in GEM (ignoring

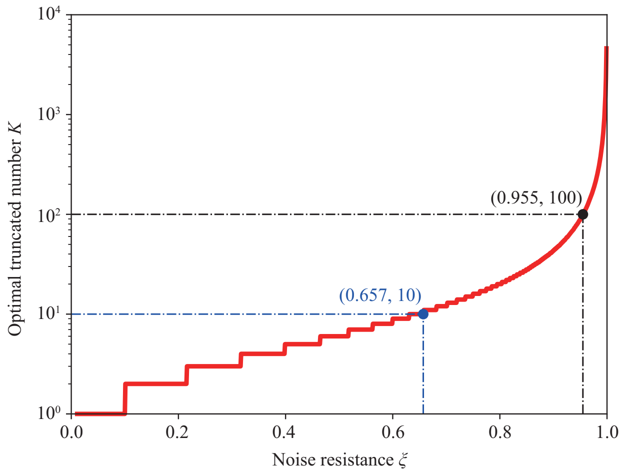

Figure 4 (Color online) The (simplified) truncated number $K = \lceil \log \varepsilon / \log \xi - 1 \rceil$ as a function of the noise resistance ξ, where $\varepsilon = 0.01$.

negligible extra terms) or ξ_m (11) in MEM. One can check from the figure that $K \leqslant 10$ whenever the noise resistance satisfies $\xi \leqslant 0.657$. In GEM with Pauli noise, this corresponds to that the minimal Pauli eigenvalue is larger than 0.343, while in MEM, this means that the minimal diagonal element of A is greater than 0.67. These conditions are easily met by many publicly available quantum devices, as shown in [15, 49]. Furthermore, in Appendix H, we illustrate that prevalent noise channels, including depolarizing, dephasing, and amplitude damping channels, exhibit noise resistances that are linearly proportional to their noise parameters. We also provide a more straightforward representation of the relationship between the optimal truncated number K and the noise parameters of these channels, which could serve as a useful guide for the practical implementation of our error mitigation method. In general, the overhead incurred by error mitigation, 4^K, can be independent of the system size (the number of qubits), as long as the noise resistance ξ of the quantum device is moderate and below a certain threshold, such as 0.657.

3.5 Numerical results

We use the qubit depolarizing channel as an example to verify the GEM method, and more examples can be found in Appendix H. This channel is defined as $\Omega_p(\rho) := \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z)$, where $p \in [0, 1]$ and X, Y, Z are the Pauli operators. Consider the task where the ideal state is $\rho = |0\rangle\langle 0|$ and $O = Z$. The ideal expectation value is $\mathrm{Tr}[Z\rho] = 1$ while the noisy expectation value suffering from depolarizing noise is $\mathrm{Tr}[Z\Omega_p(\rho)] = 1 - p$. To apply GEM, we first compute the optimal truncated number. Since $\xi_g(\Omega_p) = p$ and $\|\langle\langle Z|\|_\infty = \sqrt{2}$, we get $K_g = \lceil (\log \varepsilon - \log \sqrt{2}) / \log p - 1 \rceil$. Figure 5 shows the noisy and mitigated expectation values obtained via exact simulation for the range of noise parameters $p \in [0, 0.3]$, where $\varepsilon = 0.01$ is fixed. To account for the effect of finite measurement shots, we numerically estimate the noisy and mitigated expectation values for six noise parameters $p = 0.025, 0.075, \ldots, 0.275$, forming an arithmetic sequence. Each estimate is obtained from 10^4 measurement shots, and we repeat the mitigation procedure 10^3 times to report the mean and standard deviation of each estimate. The resulting data is also shown in Figure 5. We observe that the variance of the mitigated value increases as the noise parameter p becomes large, because a large noise parameter leads to an increased number of truncated terms $\eta_g^{(k)}$. The sample variance of $\eta_g^{(k)}$ depends on the number of measurement shots, and $\bar{\eta}_g$ is a linear combination of $\eta_g^{(k)}$. Thus, the variance of $\bar{\eta}_g$ must increase when the noise parameter becomes large while the number of measurement shots remains unchanged. From our numerical results, we can see that GEM works well and substantially improves computation accuracy, despite an increase in estimation variance, which is a common trade-off in error mitigation methods. We note that the kink-like behavior observed in Figure 5 is due to using the same truncation number K for a certain range of p. As p increases within this range, more terms of the Neumann series are needed to achieve a good approximation, and the approximation quality degrades. Therefore, using the same truncation number K across a wide range of p values can result in insufficient approximation accuracy for some values of p.

We consider another example to verify the MEM method. Consider the input state $\rho = |\Phi\rangle\langle\Phi|$, where

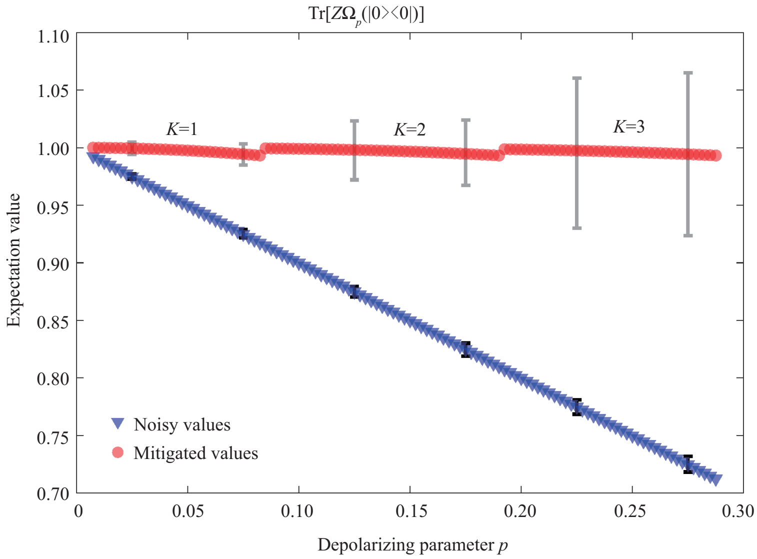

Figure 5 (Color online) Quantum GEM via truncated Neumann series for the qubit depolarizing channel, where $\varepsilon = 0.01$. The values are obtained via exact simulation. For the chosen six noise parameters $p = 0.025, 0.075, \ldots, 0.275$, we repeat the mitigation procedure 10^3 times to report the mean and standard deviation of each estimate, while each data point is obtained from 10^4 measurement shots.

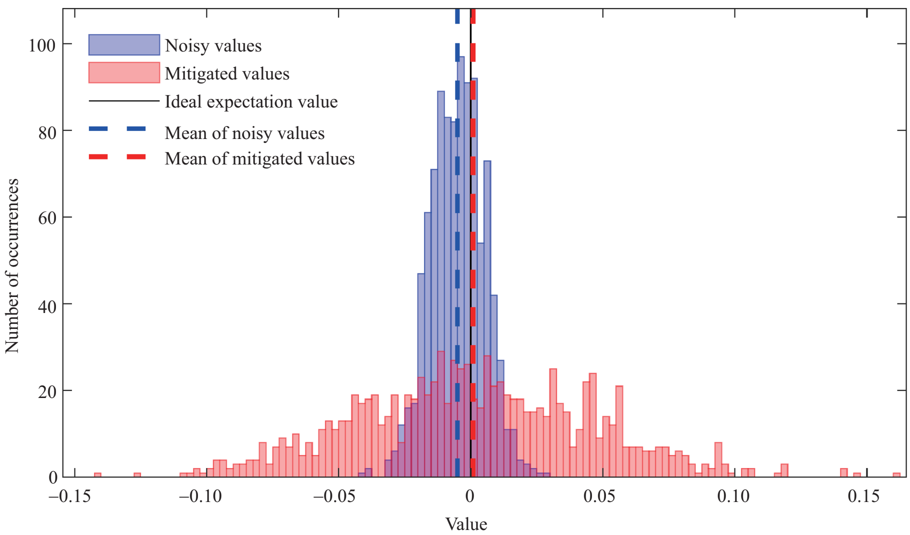

Figure 6 (Color online) 1000 noisy estimates $\eta^{(1)}$ (blue) and mitigated estimates $\overline{\eta}_m$ (red) via truncated Neumann series for the ideal expectation value 0. Here, the number of qubits is 8 and $\varepsilon = 0.01$.

Φ is the maximal superposition state $|\Phi\rangle := \sum_{i=0}^{2^n-1} |i\rangle / \sqrt{2^n}$. The observable O is a tensor product of Pauli Z operators, i.e., $O = Z^{\otimes n}$. The ideal expectation value is $\mathrm{Tr}[O\rho] = 0$. We choose $n = 8$ and randomly generate an error matrix A whose noise resistance satisfies $\xi(A) \approx 0.2$ so that the measurement error falls into the moderate regime. To account for the effect of finite measurement shots, we numerically estimate the noisy and mitigated expectation values. Each estimate is based on 10^4 measurement shots, and we repeat the mitigation procedure 10^3 times to obtain the mean and standard deviation of each estimate. Note that all these experiments use the same error matrix A with fixed parameters $\varepsilon = 0.01$. The resulting data is visualized in a bar graph in Figure 6, where the blue and red histograms indicate the noisy and mitigated values, respectively, and the blue and red dotted lines indicate their mean. The black line indicates the ideal expectation value of 0. We observe that the noisy measurement device, characterized by the error matrix A, incurs a bias of approximately -0.005 to the estimated expectation value. On the other hand, the mean of the mitigated expectation values approximates the ideal value well, despite an increase in estimation variance, which is a common trade-off in error mitigation methods.

4 Conclusion

We introduced a general framework to mitigate quantum gate and measurement errors in computing expectation values of quantum observables, an essential building block of numerous quantum algorithms. The idea behind this method is to approximate the inverse of the quantum error characterizing the noisy behavior of the underlying quantum device using a small number of truncated Neumann series terms. Remarkably, the estimation error decays exponentially in the truncated order, and the incurred error mitigation overhead is independent of the system size, as long as the noise resistance of the quantum device is moderate. The proposed error mitigation framework theoretically works for any quantum error and does not require the tomography procedure to completely characterize the quantum errors. This property is beneficial and will be more and more important as the quantum circuit sizes increase. We numerically tested this method for both gate and measurement errors and found that the computation accuracy is substantially improved. We believe that this framework will be helpful for quantum error mitigation in NISQ quantum devices. We emphasize that quantum error mitigation is still an active research area. For example, there is emerging interest in syncretizing error mitigation and correction techniques [83, 84]. It would be interesting to explore how the proposed error mitigation framework can be enhanced via error correction.

Acknowledgements We thank Runyao Duan for his helpful suggestions.

Supporting information Appendixes A–H. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 McArdle S, Endo S, Aspuru-Guzik A, et al. Quantum computational chemistry. Rev Mod Phys, 2020, 92: 015003
2 Cerezo M, Poremba A, Cincio L, et al. Variational quantum fidelity estimation. Quantum, 2020, 4: 248
3 Bharti K, Cervera-Lierta A, Kyaw T H, et al. Noisy intermediate-scale quantum algorithms. Rev Mod Phys, 2022, 94: 015004
4 Preskill J. Quantum computing in the NISQ era and beyond. Quantum, 2018, 2: 79
5 Kandala A, Mezzacapo A, Temme K, et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature, 2017, 549: 242–246
6 Arute F, Arya K, Babbush R, et al. Quantum supremacy using a programmable superconducting processor. Nature, 2019, 574: 505–510
7 Arute F, Arya K, Babbush R, et al. Hartree-fock on a superconducting qubit quantum computer. Science, 2020, 369: 1084–1089
8 Chen Z, Satzinger K J, Atalaya J, et al. Exponential suppression of bit or phase errors with cyclic error correction. Nature, 2021, 595: 383–387
9 Shor P W. Scheme for reducing decoherence in quantum computer memory. Phys Rev A, 1995, 52: R2493–R2496
10 Steane A M. Error correcting codes in quantum theory. Phys Rev Lett, 1996, 77: 793–797
11 Calderbank A R, Shor P W. Good quantum error-correcting codes exist. Phys Rev A, 1996, 54: 1098–1105
12 Aharonov D, Ben-Or M. Fault-tolerant quantum computation with constant error rate. SIAM J Comput, 2008, 38: 1207–1282
13 Nielsen M A, Chuang I L. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge: Cambridge University Press, 2010
14 Temme K, Bravyi S, Gambetta J M. Error mitigation for short-depth quantum circuits. Phys Rev Lett, 2017, 119: 180509
15 Kandala A, Temme K, Córcoles A D, et al. Error mitigation extends the computational reach of a noisy quantum processor. Nature, 2019, 567: 491–495
16 Song C, Cui J, Wang H, et al. Quantum computation with universal error mitigation on a superconducting quantum processor. Sci Adv, 2019, 5: eaaw5686
17 Zhang S, Lu Y, Zhang K, et al. Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system. Nat Commun, 2020, 11: 587
18 Kim Y, Wood C J, Yoder T J, et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. Nat Phys, 2023, 19: 752–759
19 Li Y, Benjamin S C. Efficient variational quantum simulator incorporating active error minimization. Phys Rev X, 2017, 7: 021050
20 Endo S, Benjamin S C, Li Y. Practical quantum error mitigation for near-future applications. Phys Rev X, 2018, 8: 031027
21 Dumitrescu E F, McCaskey A J, Hagen G, et al. Cloud quantum computing of an atomic nucleus. Phys Rev Lett, 2018, 120: 210501
22 Otten M, Gray S K. Recovering noise-free quantum observables. Phys Rev A, 2019, 99: 012338
23 He A, Nachman B, de Jong W A, et al. Zero-noise extrapolation for quantum-gate error mitigation with identity insertions. Phys Rev A, 2020, 102: 012426
24 Giurgica-Tiron T, Hindy Y, LaRose R, et al. Digital zero noise extrapolation for quantum error mitigation. In: Proceedings of IEEE International Conference on Quantum Computing and Engineering (QCE), 2020. 306–316
25 Schultz K, LaRose R, Mari A, et al. Impact of time-correlated noise on zero-noise extrapolation. Phys Rev A, 2022, 106: 052406

26 Takagi R. Optimal resource cost for error mitigation. Phys Rev Res, 2021, 3: 033178
27 Takagi R, Endo S, Minagawa S, et al. Fundamental limits of quantum error mitigation. npj Quantum Inf, 2022, 8: 114
28 Jiang J, Wang K, Wang X. Physical implementability of linear maps and its application in error mitigation. Quantum, 2021, 5: 600
29 Sun J, Yuan X, Tsunoda T, et al. Mitigating realistic noise in practical noisy intermediate-scale quantum devices. Phys Rev Appl, 2021, 15: 034026
30 McClean J R, Kimchi-Schwartz M E, Carter J, et al. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. Phys Rev A, 2017, 95: 042308
31 McArdle S, Yuan X, Benjamin S. Error-mitigated digital quantum simulation. Phys Rev Lett, 2019, 122: 180501
32 McClean J R, Jiang Z, Rubin N C, et al. Decoding quantum errors with subspace expansions. Nat Commun, 2020, 11: 636
33 Koczor B. Exponential error suppression for near-term quantum devices. Phys Rev X, 2021, 11: 031057
34 Koczor B. The dominant eigenvector of a noisy quantum state. New J Phys, 2021, 23: 123047
35 Huggins W J, McArdle S, O'Brien T E, et al. Virtual distillation for quantum error mitigation. Phys Rev X, 2021, 11: 041036
36 Xiong Y, Ng S X, Hanzo L. Quantum error mitigation relying on permutation filtering. IEEE Trans Commun, 2021, 70: 1927–1942
37 Cai Z Y. Resource-efficient purification-based quantum error mitigation. 2021. ArXiv:2107.07279
38 Huo M, Li Y. Dual-state purification for practical quantum error mitigation. Phys Rev A, 2022, 105: 022427
39 Yoshioka N, Hakoshima H, Matsuzaki Y, et al. Generalized quantum subspace expansion. Phys Rev Lett, 2022, 129: 020502
40 Czarnik P, Arrasmith A, Coles P J, et al. Error mitigation with Clifford quantum-circuit data. Quantum, 2021, 5: 592
41 Lowe A, Gordon M H, Czarnik P, et al. Unified approach to data-driven quantum error mitigation. Phys Rev Res, 2021, 3: 033098
42 Strikis A, Qin D, Chen Y, et al. Learning-based quantum error mitigation. PRX Quantum, 2021, 2: 040330
43 Bonet-Monroig X, Sagastizabal R, Singh M, et al. Low-cost error mitigation by symmetry verification. Phys Rev A, 2018, 98: 062339
44 O'Brien T E, Polla S, Rubin N C, et al. Error mitigation via verified phase estimation. PRX Quantum, 2021, 2: 020317
45 Cai Z Y. A practical framework for quantum error mitigation. 2021. ArXiv:2110.05389
46 Zhao Y, Ye Y, Huang H L, et al. Realization of an error-correcting surface code with superconducting qubits. Phys Rev Lett, 2022, 129: 030501
47 Acharya R, Aleiner I, Allen R, et al. Suppressing quantum errors by scaling a surface code logical qubit. Nature, 2023, 614: 676–681
48 Chow J M, Gambetta J M, Córcoles A D, et al. Universal quantum gate set approaching fault-tolerant thresholds with superconducting qubits. Phys Rev Lett, 2012, 109: 060501
49 Chen Y, Farahzad M, Yoo S, et al. Detector tomography on IBM quantum computers and mitigation of an imperfect measurement. Phys Rev A, 2019, 100: 052315
50 Geller M R. Rigorous measurement error correction. Quantum Sci Technol, 2020, 5: 03LT01
51 Maciejewski F B, Zimborás Z, Oszmaniec M. Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography. Quantum, 2020, 4: 257
52 Tannu S S, Qureshi M K. Mitigating measurement errors in quantum computers by exploiting state-dependent bias. In: Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture, 2019. 279–290
53 Nachman B, Urbanek M, de Jong W A, et al. Unfolding quantum computer readout noise. npj Quantum Inf, 2020, 6: 84
54 Hicks R, Bauer C W, Nachman B. Readout rebalancing for near-term quantum computers. Phys Rev A, 2021, 103: 022407
55 Bravyi S, Sheldon S, Kandala A, et al. Mitigating measurement errors in multiqubit experiments. Phys Rev A, 2021, 103: 042605
56 Geller M R, Sun M Y. Efficient correction of multiqubit measurement errors. 2020. ArXiv:2001.09980
57 Murali P, McKay D C, Martonosi M, et al. Software mitigation of crosstalk on noisy intermediate-scale quantum computers. In: Proceedings of the 25th International Conference on Architectural Support for Programming Languages and Operating Systems, 2020. 1001–1016
58 Kwon H, Bae J. A hybrid quantum-classical approach to mitigating measurement errors in quantum algorithms. IEEE Trans Comput, 2021, 70: 1401–1411
59 Funcke L, Hartung T, Jansen K, et al. Measurement error mitigation in quantum computers through classical bit-flip correction. Phys Rev A, 2022, 105: 062404
60 Zheng M, Li A, Terlaky T, et al. A Bayesian approach for characterizing and mitigating gate and measurement errors. ACM Trans Quantum Comput, 2023, 4: 1–21
61 Maciejewski F B, Baccari F, Zimborás Z, et al. Modeling and mitigation of realistic readout noise with applications to the quantum approximate optimization algorithm. 2021. ArXiv:2101.02331
62 Barron G S, Wood C J. Measurement error mitigation for variational quantum algorithms. 2020. ArXiv:2010.08520
63 van den Berg E, Minev Z K, Temme K. Model-free readout-error mitigation for quantum expectation values. Phys Rev A, 2022, 105: 032620
64 Geller M R. Conditionally rigorous mitigation of multiqubit measurement errors. Phys Rev Lett, 2021, 127: 090502
65 Wang K, Chen Y A, Wang X. Measurement error mitigation via truncated neumann series. In: Proceedings of the 16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021), 2021
66 Endo S, Cai Z, Benjamin S C, et al. Hybrid quantum-classical algorithms and quantum error mitigation. J Phys Soc Jpn, 2021, 90: 032001
67 Cai Z Y, Babbush R, Benjamin S C, et al. Quantum error mitigation. 2022. ArXiv:2210.00921
68 Huang H L, Xu X Y, Guo C, et al. Near-term quantum computing techniques: variational quantum algorithms, error

mitigation, circuit compilation, benchmarking and classical simulation. Sci China-Phys Mech Astron, 2023, 66: 250302
69 Peruzzo A, McClean J, Shadbolt P, et al. A variational eigenvalue solver on a photonic quantum processor. Nat Commun, 2014, 5: 4213
70 McClean J R, Romero J, Babbush R, et al. The theory of variational hybrid quantum-classical algorithms. New J Phys, 2016, 18: 023023
71 Farhi E, Goldstone J, Gutmann S. A quantum approximate optimization algorithm. 2014. ArXiv:1411.4028
72 Biamonte J, Wittek P, Pancotti N, et al. Quantum machine learning. Nature, 2017, 549: 195–202
73 Havlíček V, Córcoles A D, Temme K, et al. Supervised learning with quantum-enhanced feature spaces. Nature, 2019, 567: 209–212
74 Hoeffding W. Probability inequalities for sums of bounded random variables. J Am Stat Assoc, 1963, 58: 13–30
75 Greenbaum D. Introduction to quantum gate set tomography. 2015. ArXiv:1509.02921
76 Stewart G W. Matrix Algorithms: Volume 1: Basic Decompositions. Philadelphia: SIAM, 1998
77 Wu M, Yin B, Vosoughi A, et al. Approximate matrix inversion for high-throughput data detection in the large-scale mimo uplink. In: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS), 2013. 2155–2158
78 Harper R, Flammia S T, Wallman J J. Efficient learning of quantum noise. Nat Phys, 2020, 16: 1184–1188
79 Chen S, Zhou S, Seif A, et al. Quantum advantages for Pauli channel estimation. Phys Rev A, 2022, 105: 032435
80 Flammia S T, Wallman J J. Efficient estimation of Pauli channels. ACM Trans Quantum Computing, 2020, 1: 1–32
81 Cao C, Yu Y, Wu Z, et al. Mitigating algorithmic errors in quantum optimization through energy extrapolation. Quantum Sci Technol, 2023, 8: 015004
82 Baidu Quantum. Quantum error processing toolkit (QEP), 2022. https://quantum-hub.baidu.com/qep/
83 Piveteau C, Sutter D, Bravyi S, et al. Error mitigation for universal gates on encoded qubits. Phys Rev Lett, 2021, 127: 200505
84 Lostaglio M, Ciani A. Error mitigation and quantum-assisted simulation in the error corrected regime. Phys Rev Lett, 2021, 127: 200506