

Classical certification of quantum gates under the dimension assumption

Jan Nöller¹, Nikolai Miklin², Martin Kliesch², and Mariami Gachechiladze¹

¹Department of Computer Science, Technical University of Darmstadt, Darmstadt, 64289 Germany

²Institute for Quantum Inspired and Quantum Optimization, Hamburg University of Technology, Germany

The rapid advancement of quantum hardware necessitates the development of reliable methods to certify its correct functioning. However, existing certification tests fall short, as they either suffer from systematic errors or do not guarantee that only a correctly functioning quantum device can pass the test. We introduce a certification method for quantum gates tailored for a practical server-user scenario, where a classical user tests the results of exact quantum computations performed by a quantum server. This method is free from the systematic state preparation and measurement (SPAM) errors. For single-qubit gates, including those that form a universal set for single-qubit quantum computation, we demonstrate that our approach offers soundness guarantees based solely on the dimension assumption. Additionally, for a highly-relevant phase gate – which corresponds experimentally to a $\pi/2$ -pulse – we prove that the method’s sample complexity scales as $O(\varepsilon^{-1})$ relative to the average gate infidelity ε . By combining the SPAM-error-free and sound notion of certification with practical applicability, our approach paves the way for promising research into efficient and reliable certification methods for full-scale quantum computation.

1 Introduction

The problem of certifying the correct functioning of quantum devices is crucial for developing quantum hardware and has naturally evolved into a field of study known as quantum system characterization (see Refs. [1, 2] for reviews). Particularly challenging is assessing the quality of quantum gates due to unavoidable state preparation and measurement (SPAM) errors [2]. They are a limiting factor in standard quantum process tomography [3, 4] and direct certification methods [5–7]. Two broad families of characterization methods have been developed to address this challenge: gate set tomography (GST) [8–10] and random-

ized benchmarking (RB) [11–17] with its many variants (see Refs. [17, 18] for a recent overview). Both types of protocols can be used to estimate gate errors in a SPAM-robust manner by executing sequences of gates with varying lengths. However, neither of them is suitable for certification, as they cannot definitively rule out the implementation of incorrect gates. GST provides a set of compatible descriptions of the underlying experiment, some of which may not be connected to the implemented operations by any physical gauge. RB, in turn, already requires that the implemented gates are close to the target ones in order to interpret the output decay parameters [17].

In general, any certification method should satisfy two key properties: *completeness* and *soundness* [2]. In simple terms, the former means that a certification test should accept correctly implemented target quantum operations, and the latter that *only* the correct implementation should be accepted. While existing SPAM-robust characterization methods satisfy completeness, currently, there is no method in the quantum characterization literature that is both sound and is free from SPAM errors.

An independent line of research, known as *self-testing* [19], offers a framework for sound certification of quantum devices while treating them as black boxes. The black-box nature of certification ensures that all quantum operations, including the state preparation, quantum evolution, and the measurement are certified at the same time. In the context of quantum gate certification, this implies that conclusions that one draws from self-testing are free from SPAM errors, because any such systematic errors are detected.

The self-testing literature primarily focuses on certifying entangled states and local measurements in the Bell test (see e.g., Ref. [20] for a review). However, some works have also considered quantum channels [21] or instruments [22], that can be applied to one of the subsystems in the Bell test, and also entangling interactions [23]. Extending the framework of self-testing to include quantum dynamics, and in particular quantum gates [24, 25], makes it relevant for the problem of testing quantum computers. The general idea is to combine self-testing of states and measurements with protocols such as

Jan Nöller: jan.noeller@tu-darmstadt.de

Nikolai Miklin: nikolai.miklin@tuhh.de, these authors contributed equally to this work.

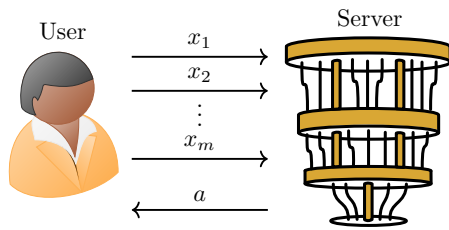


Figure 1: Schematic figure of a server-user interaction. A classical user via a classical channel transmits instructions $x_1, x_2, \dots, x_m$ to a quantum server, which performs a quantum computation and returns the result a .

process tomography, which typically require trusted devices, to obtain device-independent certification of quantum operations. However, all these schemes demand an experimental setup consisting of two well-isolated parts, a requirement that is challenging to achieve within a single quantum processor. A recent self-testing study [26] suggested relaxing this experimental requirement by introducing computational assumptions [27]. For these assumptions to be accepted, however, experimental capabilities beyond the reach of current quantum hardware are required [28].

Finally, a related question of learnability of quantum operations was recently raised in Ref. [29]. There, the authors proposed a framework to investigate the learnability of the intrinsic descriptions of quantum experiments from observational data. However, the general result obtained there relies on more assumptions than desirable for a certification task.

In this work, our objective is to achieve SPAM error-free certification of a quantum computer's correct functioning, without requiring physical access to it, assuming as little about the quantum computer's internal functioning as possible. This investigation is particularly relevant in a practical server-user configuration (see Fig. 1): a classical user via a classical channel transmits gate sequences to a quantum server, which then implements them and returns the measurement outcomes. Since achieving this goal without any assumptions is impossible, we adopt the commonly considered assumption on the quantum system's dimension [30–35].

Since we assume that no part of the quantum apparatus, such as the measurement device, is characterized prior to the test, certification is only possible up to the degrees of freedom inherent to quantum mechanics, i.e., unitary or anti-unitary transformation (a unitary and the complex conjugation) [36]. It is important to note that this is the absolute minimum degree of freedom that cannot be excluded in black-box tests, as it corresponds to the simultaneous change of bases. Our method is based on a very intuitive idea of testing outcomes of exact quantum computation for quantum gate sequences that can be resolved efficiently classically. Examples include gate sequences that compose into the identity gate or simply a zero-

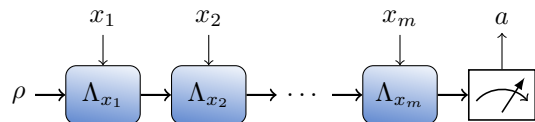


Figure 2: Scenario of the certification test. A system is prepared in state ρ and undergoes a series of transformations, $\Lambda_{x_1}, \Lambda_{x_2}, \dots, \Lambda_{x_m}$, specified by classical instructions $\{x_i\}_{i=1}^m$ from a fixed set, after which it is measured in a fixed basis, producing outcome a . The sequences $\mathbf{x} = x_1 x_2 \dots x_m$ of varied length m are chosen at random from the predetermined set $\mathcal{X}$.

length sequence, for which the system is measured directly after the state preparation. For certification of quantum gates, only the input-output correlations are used and no entanglement with an auxiliary system is required.

Here, we prove that a gate set universal for single qubit quantum operations can be certified within our framework, and analyze in detail the certification of a relevant single-qubit gate, which corresponds to the application a $\pi/2$ -pulse. The latter is routinely implemented on a quantum computer for creating an equal superposition of qubit basis states [37]. We show that the required *sample complexity*, measured in terms of the number of individual runs of the experiment [2], scales as $O(\varepsilon^{-1})$ with respect to the average gate infidelity ε .

This favorable scaling paves the way for a promising research in the certification of quantum systems. In this work, we present proof-of-principle results for a number of important single-qubit gates. At the same time, the proposed protocol has the potential to be extended for application in full-scale quantum computation. While these generalizations are intriguing, they may result in less optimal guarantees. As a result, the scaling demonstrated in our proofs not only minimizes the resources required to certify single-qubit gates but also sets a benchmark for assessing the effectiveness of future generalizations.

The rest of the paper is organized as follows. In Section 2, we explain the experimental setup, outline the assumptions, and describe the protocol. In Section 3, we present our results on certification of single-qubit quantum operations, with the main contributions stated in Theorem 4, Corollary 5, and Theorem 6. Technical details supporting the main claims of the paper are left to the appendix.

2 Setup and protocol

The experimental setup that we consider is common to many established certification methods, such as RB [17] and GST [8]. The setup, or scenario as it would be called in the self-testing literature, is shown schematically in Fig. 2. A quantum system is prepared in some initial state, after which a sequence of

quantum gates is applied to it, and it is finally measured in some fixed basis. For each gate in a given sequence, a label, chosen from some finite set X , is communicated to the quantum computer. The certification protocol relies on a particular finite subset of sequences $\mathcal{X} \subset X^*$, which are determined before the protocol begins. In a single repetition of the protocol, a random string $\mathbf{x} := x_1 x_2 \dots x_m$ is selected from $\mathcal{X}$, and after the quantum computer implements the corresponding computation, the outcome $a \in A$ of the measurement is read out, where A is the set of all possible outcomes. The certification protocol decides to proceed or abort, depending on the deterministic outcome $a_{\mathbf{x}}$, corresponding to the ideal implementation of the target quantum computation. The length m of different sequences $\mathbf{x} \in \mathcal{X}$ can be different, and, in particular, be zero, which we denote by the empty string ϵ , and by which we mean that the system is measured directly after the state preparation.

Next, we give a definition of a quantum model.

Definition 1. A d -dimensional quantum model is a 3-tuple $(\rho, \{\Lambda_x\}_{x \in X}, \{M_a\}_{a \in A})$, consisting of a quantum state ρ , prepared at the beginning of the quantum computation, a finite set of quantum channels (gates) $\{\Lambda_x\}_{x \in X}$, from which a quantum circuit is composed, and a positive operator-valued measure (POVM) $\{M_a\}_{a \in A}$, measured at the end of the computation, all defined over a Hilbert space $\mathcal{H}$ with $\dim(\mathcal{H}) = d$.

For a quantum model involving only unitary quantum channels, we use the corresponding unitary operators in the definition of the model. We treat our setup as a single black box, making only the *minimal* assumptions – such that removing any of these assumptions would render the results of this paper unattainable in the given scenario.

- (i) *The dimension assumption.* We assume that in each experimental round, the implemented state preparation is mathematically described by a density operator, the operations are described by completely positive trace preserving (CPTP) maps, and the measurement is described by a POVM, all defined over a Hilbert space of a specified dimension.
- (ii) *Context independence.* We assume that in each repetition of our protocol, for each label x , there is a single corresponding quantum channel Λ_x that the quantum computer implements. That is, for any input sequence $\mathbf{x} = x_1 x_2 \dots x_m$, we assume that the input state undergoes the corresponding series of transformations $\Lambda_{x_m} \circ \dots \circ \Lambda_{x_2} \circ \Lambda_{x_1}$. In the certification literature, this assumption would be referred to as existence of a single-shot implementation function.

The above assumptions allow us to mathematically describe an experiment in the considered scenario by

a quantum model, as in Definition 1. Other minor assumptions include error-free functioning of the classical part of the quantum computer, such as control circuits, and our ability to randomly select the gate sequences. Notably, it is not necessary to assume the ability to sample the gates precisely according to a fixed distribution.

The dimension assumption is critical, because a single quantum system, with which a classical user is interacting, can be simulated by a classical system of a larger dimension [31, 33]. We note here, that the dimension assumption also means that no side channels, e.g., operated by an adversary, are considered. The assumption of context independence cannot be removed, because for each gate sequence one can always assign a POVM that reproduces the required statistics, and a quantum computer can simply perform this measurement on the preparation state, ignoring the structure of the gate sequence.

Finally, when we estimate the sample complexity of our protocol, we need to use an assumption (iii) of *independence of repetitions*. As we argue below, (iii) is a part of (i), but we separate it here for clarity. We need the independence of repetitions to treat events in different repetitions of the protocol as statistically independent. However, the dimension assumption already implies that there is a tensor product between quantum models implemented in different rounds of the protocol, which leads to the independence of the observed outcomes. Indeed, if one does not assume this tensor structure, there is always a possibility of a global unitary gauge applied to several copies of state, measurements, and quantum gates in a way that makes these objects entangled across the protocol repetitions. This would not change the observed statistics, but would make it impossible to say anything about the models implemented in a single experimental run. Note, however, that for the certification, we *do not* need the assumption that quantum models in different repetitions of the protocol are identical.

We are ready to present our Protocol 1 for classical certification of quantum gates. We give a general formulation for a given set of gates X , with an important property that among all possible sequences X^* , there are such $\mathbf{x}$, for which we can predict the deterministic outcome $a_{\mathbf{x}}$, which a noiseless quantum computer should output.

The basic idea is that if Protocol 1 accepts a model for some large N , and therefore obtained the correct outcome $a_{\mathbf{x}}$ for different sequences $\mathbf{x}$ in all these tests, we can obtain a certain level of confidence, typically denoted by $1 - \delta$, that our quantum computer implements the quantum model correctly. We define more precisely below what we mean by the latter, building on similar definitions in the self-testing literature [33].

Protocol 1 Classical certification of quantum gates

Set N – the number of repetitions, $\mathcal{X}$ – the set of gate sequences $\mathbf{x}$, each with the corresponding deterministic outcome $a_{\mathbf{x}}$, and μ – a probability mass function over $\mathcal{X}$.

for $i \in [N]$ **do**

 sample $\mathbf{x}$ from $\mathcal{X}$, according to μ ;

 run the quantum circuit consisting of the state preparation, the sequence $\mathbf{x}$ of gates, and the measurement, record the outcome a ;

if $a \neq a_{\mathbf{x}}$ **then**

 output "reject" and end the protocol;

 output "accept".

Since we do not assume any part of the quantum computer to be characterized, and rely only on the classical data in our certification, any two quantum models which are equivalent up to the symmetries in quantum mechanics [36] will produce the same statistics, and we will not be able to distinguish between them. At the same time, we would like to exclude any other quantum model, which is formalized by the following definition.

Definition 2. For a target d -dimensional quantum model $(\rho, \{U_x\}_{x \in X}, \{M_a\}_{a \in A})$ with unitary channels, we say that another d -dimensional quantum model $(\tilde{\rho}, \{\tilde{U}_x\}_{x \in X}, \{\tilde{M}_a\}_{a \in A})$ is its correct implementation if there exists a unitary operator U (with a possible complex conjugation $(*)$), such that

$$\begin{aligned}\tilde{\rho} &= U\rho^{(*)}U^\dagger, \\ \tilde{U}_x(\cdot) &= UU_x^{(*)}U^\dagger(\cdot)UU_x^{\dagger(*)}U^\dagger, \quad \forall x \in X, \\ \tilde{M}_a &= UM_a^{(*)}U^\dagger, \quad \forall a \in A.\end{aligned}\quad (1)$$

The negation of Definition 2 provides a definition of an incorrect implementation. Following the terminology of the self-testing literature [20], we say that the target outcomes $a_{\mathbf{x}}$ *self-test* a quantum model, if from the fact that the observed outcomes correspond to $a_{\mathbf{x}}$, we can infer that the implemented quantum model is a correct implementation of the target model, in the sense of Definition 2. Moreover, we say that the self-test is *robust*, if for small deviations in the outcomes, the target and the implemented models are close in some distance. For quantum states and quantum measurements, we use the infidelity and the spectral distance, respectively, and for quantum gates, we use the average gate infidelity. Here, we use the following expression for the average gate fidelity between a qubit channel $\tilde{\Lambda}$ and qubit unitary channel Λ (see e.g., [2]),

$$F_{\text{avg}}(\tilde{\Lambda}, \Lambda) = \frac{2}{3} \text{Tr}[J(\tilde{\Lambda})J(\Lambda)] + \frac{1}{3}, \quad (2)$$

where $J(\cdot)$ is the Choi-Jamiołkowski [38, 39] map,

defined for a qubit channel Λ as

$$J(\Lambda) := \frac{1}{2} \sum_{i,j \in \{0,1\}} \Lambda(|i\rangle\langle j|) \otimes |i\rangle\langle j|. \quad (3)$$

Definition 2 is motivated by the natural symmetries in quantum mechanics [36]: two quantum models connected by a (anti-) unitary transformation as in Eq. (1) will always lead to the same observed statistics. For this reason, whenever a collection of deterministic outcomes produced by a quantum computer self-test a target quantum model, we also say that the quantum computer implemented this model *correctly*, even if the target and the implemented models are not exactly equal, but only equivalent up to a (anti-) unitary gauge. For the same reason, we do not treat the (anti-) unitary gauge in the implemented model as *noise*.

Following the terminology of the certification literature [2], we say that Protocol 1 is a certification test for a target quantum model with respect to appropriately chosen distances, if the protocol is *complete* and *sound*.

Definition 3. Given a null hypothesis H_0 and an alternative hypothesis H_1 , a test is complete, if

$$\mathbb{P}[\text{"accept"}|H_1] \geq 1 - \delta', \quad (4)$$

for $\delta' < \frac{1}{2}$ and sound, if

$$\mathbb{P}[\text{"reject"}|H_0] \geq 1 - \delta, \quad (5)$$

for $\delta < \frac{1}{2}$.

It is common to take $\delta' = 0$ in Eq. (4), which is also what we do in this work. In our certification test, H_1 is the hypothesis that the implemented model is a correct implementation of a target model, as given by Definition 2. For our certification results for a gate set universal for single-qubit quantum computation, we take H_0 to be the negation of H_1 . One can refer to this case as “non-robust soundness”, as this would require the test to run infinitely. In the self-testing literature, this is often referred to as the *ideal* case [20]. For our results for a phase gate, we relax the hypothesis H_0 , and only exclude models for which no unitary can bring them ε -close in the chosen distance to the target model. This allows us to set an upper-bound on the required number of repetitions of Protocol 1, i.e., the sample complexity.

3 Certification of single-qubit quantum models

In this section, first we prove that Protocol 1 is an ε -certification test for a quantum model $(|+\rangle\langle+|, \{S, S^\dagger\}, \{|+\rangle\langle+|, |-\rangle\langle-|\})$, where $S := |0\rangle\langle 0| + i|1\rangle\langle 1|$. The soundness of this test follows from a more general robust self-testing-type

result, which we prove first. Here, we use S gate to model a $\pi/2$ -pulse with respect to an orthonormal basis (ONB) $\{|+\rangle, |-\rangle\}$ [37]. However, all the following results for the quantum model $(|+\rangle\langle+|, \{S, S^\dagger\}, \{|+\rangle\langle+|, |-\rangle\langle-|\})$ also apply for other unitarily equivalent models such as e.g., $(|0\rangle\langle 0|, \{\sqrt{X}, \sqrt{X}^\dagger\}, \{|0\rangle\langle 0|, |1\rangle\langle 1|\})$, where $\sqrt{X} := |+\rangle\langle+| + i|-\rangle\langle-|$. Then, we prove an ideal self-testing result for a quantum model with an additional H gate and the square root of the S gate, i.e., the model $(|+\rangle\langle+|, \{S, S^\dagger, H, T\}, \{|+\rangle\langle+|, |-\rangle\langle-|\})$, which shows that a single-qubit universal gate set can be certified using Protocol 1.

3.1 Certification of a phase gate

We explain in detail certification of the S gate, or more precisely, the quantum model

$$(|+\rangle\langle+|, \{S, S^\dagger\}, \{|+\rangle\langle+|, |-\rangle\langle-|\}). \quad (6)$$

We set $X = \{s, s^{-1}\}$ for the classical instructions given to a quantum computer which specify whether it should implement the S gate or its inverse, respectively.

Surprisingly, it is sufficient to consider the following set of strings in Protocol 1

$$\mathcal{X} = \{\epsilon, ss, ss^{-1}, s^{-1}s, s^{-1}s^{-1}\}, \quad (7)$$

where ϵ denotes the empty string. Next, we set $A = \{+, -\}$, and for the sequences in Eq. (7), determine that the deterministic outcomes corresponding to the target model are the following

$$a_{\mathbf{x}} = \begin{cases} + & \text{for } \mathbf{x} \in \{\epsilon, ss^{-1}, s^{-1}s\}, \\ - & \text{for } \mathbf{x} \in \{ss, s^{-1}s^{-1}\}. \end{cases} \quad (8)$$

It can be easily seen that in case of noiseless implementation of the state preparation $|+\rangle\langle+|$, the gates $S, S^\dagger$, and the measurement $\{|+\rangle\langle+|, |-\rangle\langle-|\}$, Protocol 1 always accepts. In other words, Protocol 1 is *complete* for certification of the quantum model $(|+\rangle\langle+|, \{S, S^\dagger\}, \{|+\rangle\langle+|, |-\rangle\langle-|\})$, with $\mathcal{X}$ and $\{a_{\mathbf{x}}\}_{\mathbf{x} \in \mathcal{X}}$ specified by Eq. (7) and Eq. (8), respectively.

Proving the *soundness* of the protocol is much less straightforward. To achieve this, we first state and prove the following self-testing result.

Theorem 4. *Let $(\tilde{\rho}, \{\tilde{\Lambda}_s, \tilde{\Lambda}_{s^{-1}}\}, \{\tilde{M}_+, \tilde{M}_-\})$ be a single-qubit quantum model that passes with probability at least $1 - \varepsilon$ a single repetition of Protocol 1 with uniform sampling from the gate sequences (7) and for deterministic measurement outcomes (8). Then the quantum model is $O(\varepsilon)$ -close to the target model (6),*

i.e., there is a unitary $U \in U(2)$ such that

$$\begin{aligned} F_{\text{avg}}(\tilde{\Lambda}_s, USU^\dagger) &\geq 1 - O(\varepsilon), \\ F_{\text{avg}}(\tilde{\Lambda}_{s^{-1}}, US^\dagger U^\dagger) &\geq 1 - O(\varepsilon), \\ \text{Tr}[\tilde{\rho}U|+\rangle\langle+|U^\dagger] &\geq 1 - \frac{15}{2}\varepsilon, \\ \|\tilde{M}_+ - U|+\rangle\langle+|U^\dagger\|_\infty &\leq \frac{5}{2}\varepsilon. \end{aligned} \quad (9)$$

In the case of unitary channels, we use the respective operators as the argument of the fidelity function for simplicity of notation. Below, we give a sketch of the proof, and the full proof can be found in Appendix A.

Proof sketch. The conclusions of the theorem follow from the condition $\mathbb{P}[\text{"pass"}] \geq 1 - \varepsilon$ and the dimension assumption. As a first step, we show that for small ε , the measurement effects $\tilde{M}_+$ and $\tilde{M}_-$ are close to being rank-1 projectors, which we denote as $|\psi\rangle\langle\psi|$ and $|\psi^\perp\rangle\langle\psi^\perp|$. Next, we show that the POVMs which one obtains by applying the adjoint maps $\tilde{\Lambda}_s^\dagger$ and $\tilde{\Lambda}_{s^{-1}}^\dagger$ to $|\psi\rangle\langle\psi|$ and $|\psi^\perp\rangle\langle\psi^\perp|$ are also close to being projective for small ε . We denote the corresponding projectors by $|\phi\rangle\langle\phi|$ and $|\phi^\perp\rangle\langle\phi^\perp|$. Importantly, we find that $\tilde{\Lambda}_s^\dagger(|\psi\rangle\langle\psi|) \approx |\phi\rangle\langle\phi|$ and $\tilde{\Lambda}_{s^{-1}}^\dagger(|\psi\rangle\langle\psi|) \approx |\phi^\perp\rangle\langle\phi^\perp|$, and since the adjoint maps of channels are unital, also $\tilde{\Lambda}_s^\dagger(|\psi^\perp\rangle\langle\psi^\perp|) \approx |\phi^\perp\rangle\langle\phi^\perp|$ and $\tilde{\Lambda}_{s^{-1}}^\dagger(|\psi^\perp\rangle\langle\psi^\perp|) \approx |\phi\rangle\langle\phi|$. Next, we obtain a partial characterization of the Choi-Jamiołkowski state of the channel $\tilde{\Lambda}_s$ in the bases of $|\psi\rangle, |\psi^\perp\rangle$ and $|\phi\rangle, |\phi^\perp\rangle$, with the leading terms corresponding to the subspace spanned by $|\psi\rangle|\phi\rangle$ and $|\psi^\perp\rangle|\phi^\perp\rangle$. Then, we show that $\tilde{\Lambda}_s$ is close to being a unitary channel, for which we use the conditions $\tilde{\rho} \approx |\psi\rangle\langle\psi|$ and $\tilde{\Lambda}_s(\tilde{\rho}) \approx |\phi^\perp\rangle\langle\phi^\perp|$. Finally, by fixing the global phases of the vectors $|\psi\rangle, |\psi^\perp\rangle$ appropriately, we construct a suitable gauge unitary,

$$U = |\psi\rangle\langle+| - i|\psi^\perp\rangle\langle-|, \quad (10)$$

for which the condition for $\tilde{\Lambda}_s$ in Eq. (9) is satisfied. Because we obtain characterization of $\tilde{\Lambda}_s$ and $\tilde{\Lambda}_{s^{-1}}$ in the same basis, the proof also easily extends to the channel $\tilde{\Lambda}_{s^{-1}}$. The bounds for the state and the measurement in Eq. (9) with the chosen unitary are also immediate. $\square$

Interestingly, in the ideal case of $\varepsilon = 0$, the self-testing argument can also be made for the set of gate sequences without $s^{-1}s^{-1}$, or without ss . Moreover, the effect that the sampling distribution μ over $\mathcal{X}$ plays in Theorem 4 is purely in determining the constants in front of ε , with the only requirement that each sequence in $\mathcal{X}$ is chosen with some nonzero probability.

We can use the known relation that connects the average gate fidelity and the diamond distance for an arbitrary qubit channel $\tilde{\Lambda}$ and a unitary channel Λ [2],

$$\|\tilde{\Lambda} - \Lambda\|_\diamond \leq 2\sqrt{6}\sqrt{1 - F_{\text{avg}}(\tilde{\Lambda}, \Lambda)}, \quad (11)$$

to reformulate Theorem 4 for the diamond distance with an upper bound of $O(\sqrt{\varepsilon})$.

Next, using Theorem 4, we show that Protocol 1 is sound for certification of the model $(|+\rangle\langle+|, \{S, S^\dagger\}, \{|+\rangle\langle+|, |- \rangle\langle-|\})$, as stated by the following corollary. This corollary is the main practical result of the paper.

Corollary 5. *Protocol 1 with uniform sampling from the gate sequences (7) and for deterministic measurement outcomes (8) is an ε -certification test for the S gate and its inverse with respect to the average gate infidelity, as well as initial state $|+\rangle\langle+|$ and measurement $\{|+\rangle\langle+|, |- \rangle\langle-|\}$ with respect to infidelity and spectral norm from N independent samples for $N \geq N_0$ with*

$$N_0 = O(\varepsilon^{-1}) \ln(\delta^{-1}) \quad (12)$$

with confidence at least $1 - \delta$. Moreover, Protocol 1 accepts the target model $(|+\rangle\langle+|, \{S, S^\dagger\}, \{|+\rangle\langle+|, |- \rangle\langle-|\})$ with probability 1.

The lower bound of N_0 in Corollary 5 should be interpreted as a sufficient number of repetitions of Protocol 1 to reach the target confidence level $1 - \delta$.

Proof. We can invert the statement of Theorem 4, and obtain that for a noisy quantum model, for which there does not exist a unitary $U \in U(2)$, satisfying Eq. (9), the probability of passing a single repetition of Protocol 1 is $\mathbb{P}[\text{"pass"}] \leq 1 - \varepsilon$. If we take N independent copies of such noisy $\tilde{\Lambda}_s, \tilde{\Lambda}_{s^{-1}}$, as well as $\tilde{\rho}$ and $\{\tilde{M}_+, \tilde{M}_-\}$, the probability of Protocol 1 accepting them is

$$\mathbb{P}[\text{"accept"}] \leq (1 - \varepsilon)^N. \quad (13)$$

For the target confidence level $1 - \delta$, the acceptance probability in Eq. (13) should be upper-bounded by δ , which leads to a lower bound on N ,

$$N \geq \frac{\ln(\delta^{-1})}{\ln \frac{1}{1-\varepsilon}}. \quad (14)$$

Approximating the logarithm function for $\frac{1}{1-\varepsilon}$, and rescaling ε such that the lower bounds on the average gate fidelity in Eq. (9) are exactly $1 - \varepsilon$, leads to the sample complexity stated in Eq. (12). $\square$

As Corollary 5 demonstrates, our method for certification of quantum gates is as efficient as the direct certification of quantum processes [5], which requires trust in state preparations and measurements (and hence is not free from SPAM errors) and also requires an auxiliary system to prepare the Choi-Jamiołkowski state of the process. The only price to pay is a possibly larger constant factor, which we numerically estimate for the uniform μ in the next section.

The commonly used SPAM-robust characterization methods, RB and GST, are sample efficient, however, as mentioned earlier, if used for certification, they do not come with the soundness guarantees. The proposal of Ref. [24] for self-testing of quantum gates in the Bell test is sound, but can tolerate very little amount of noise (see discussion in Ref. [21]). The proposal of Ref. [21] reports higher noise tolerance, but as noted in Ref. [40], self-tests based on violations of a Bell inequality that do not reach the algebraic maximum suffer from a quadratically worse scaling of sample complexity with respect to ε .

Finally, since Protocol 1 uses the same experimental setup as GST and RB, it can be seamlessly integrated into these protocols. In particular, if S and $S^\dagger$ are included in the gate set of a GST experiment, the statistics gathered there can be used to obtain a lower bound on the average gate fidelity from Theorem 4 (under an additional i.i.d. assumption). Moreover, the gauge freedom in the GST output can be reduced to unitary for the gates S and $S^\dagger$. Similarly, if the acceptance probability of Protocol 1 is estimated before conducting an RB experiment, it can be used in the guarantees of the RB protocol's output.

3.2 Numerical investigations

In this subsection, we supplement our theoretical results of Theorem 4 and Corollary 5 by numerical investigations. This also allows us to estimate the coefficients of the linear scaling in Theorem 4, which then translates to an explicit formula for the sample complexity in Corollary 5. The results of our numerical investigations are shown in Fig. 3.

For each randomly generated quantum model $(\tilde{\rho}, \{\tilde{\Lambda}_s, \tilde{\Lambda}_{s^{-1}}\}, \{\tilde{M}_+, \tilde{M}_-\})$, we calculate the probability of it failing a single repetition of the protocol, which corresponds to ε in the statement of Theorem 4. We then apply a unitary to the target model, which explicitly depends on the noisy random model, as described in the proof of Theorem 4, and calculate the distance between the noisy model and the target model. As the latter, we take the maximum of the two average gate infidelities for the S and $S^\dagger$ gates.

We consider four different noise models: unitary noise, two combinations of unitary noise with depolarizing noise, and the depolarizing noise. To generate random noisy quantum models, we apply independent unitaries to the target state, channels, and the measurement, i.e., we take $\tilde{\rho} = U_1|+\rangle\langle+|U_1^\dagger$, $\tilde{\Lambda}_s(\cdot) = U_2SU_3(\cdot)U_3^\dagger S^\dagger U_2^\dagger$, $\tilde{\Lambda}_{s^{-1}}(\cdot) = U_4S^\dagger U_5(\cdot)U_5^\dagger SU_4^\dagger$, $\tilde{M}_+ = U_6|+\rangle\langle+|U_6^\dagger$, and $\tilde{M}_- = \mathbb{1} - \tilde{M}_+$, for randomly sampled $U_1 \dots U_6$. Sampling Haar-random unitaries $U_1 \dots U_6$ would result in a quantum model far from the target one, which would not be useful for our numerical investigation. Therefore, we generate each U_i , $i \in \{1, \dots, 6\}$, by randomly sampling $u_i \in \mathfrak{su}(2)$ with $\|u_i\|_\infty = 1$ and then setting

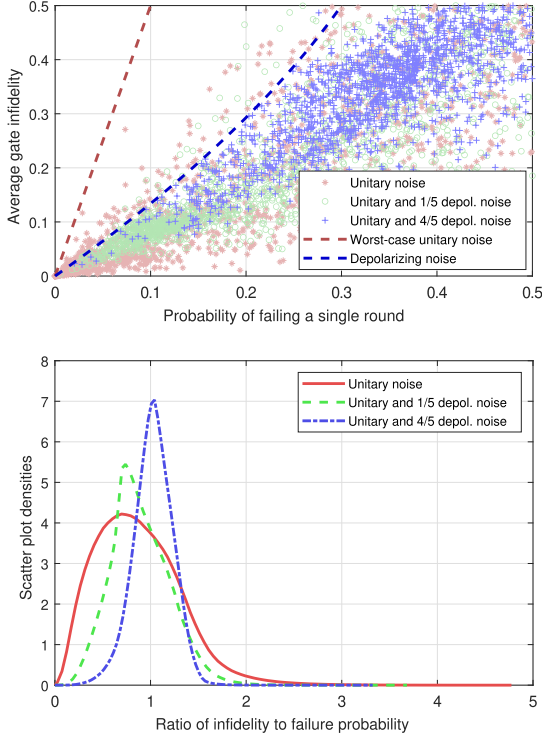


Figure 3: Results of our numerical investigations. Top: dependence of the average gate infidelity on the probability of failing a single round of Protocol 1 for noisy models, as in the statement of Theorem 4. Each of the approximately 10^3 depicted points represents a single noisy model, with different noise models considered. The straight dotted line corresponds to the worst-case scaling among the randomly sampled noisy models, and the dotted curve corresponds to the depolarizing noise. Bottom: a collection of histograms representing distributions of the ratios of the average gate infidelity to the failure probability for the three considered cases of noise models. Each histogram is based on 10^7 random samples.

$U_i = e^{\alpha_i u_i}$ for uniformly sampled $\alpha_i \in [0, 1]$.

In addition to the unitary noise, we consider adding various amounts of depolarizing noise to the channels $\tilde{\Lambda}_s$ and $\tilde{\Lambda}_{s^{-1}}$. Finally, we consider a noise model in which the state preparation, measurement, and the S-gate are implemented ideally, while the $S^\dagger$ -gate suffers from the depolarizing noise, $\tilde{\Lambda}_{s^{-1}}(\rho) = (1 - \eta)S^\dagger(\rho)S + \eta \text{Tr}[\rho] \frac{1}{2}$, with the depolarizing parameter $\eta \in [0, 1]$. A straightforward calculation shows that for this model $1 - \mathbb{P}[\text{"pass"}] = \frac{(4-\eta)\eta}{10}$, and $1 - F_{\text{avg}}(\tilde{\Lambda}_{s^{-1}}, S^\dagger) = \frac{\eta}{2}$.

In Fig. 3 (top) we plot approximately 10^3 randomly generated noisy quantum models, as well as a linear worst-case upper bound on the distance given the failing probability, which we estimated from sampling 10^7 random models. In addition to the unitary noise, we consider adding depolarizing noise with the depolarizing parameters $\frac{1}{5}$ and $\frac{4}{5}$, as well as the purely depolarizing noise added to the $S^\dagger$, as described above. In Fig. 3 (bottom) we present histograms for the consid-

ered noise models, apart from the purely depolarizing noise. As one can see from the histograms, the unitary noise results in better scaling on average, but it is also the one which produces the worst-case behavior.

The performed numerical investigation follows the theoretical predictions of the linear scaling of the distances as a function of ε in Theorem 4, and allows us to estimate the coefficient of this linear function to be approximately 5. This results in the sample complexity of $\frac{5}{\varepsilon} \ln(\delta^{-1})$, which amounts to approximately 2000 repetitions for $\varepsilon = \delta = 0.01$, or approximately 300 repetitions for $\varepsilon = \delta = 0.05$.

3.3 Certification of a gate set universal for single-qubit quantum computation

Next, we show how to employ Protocol 1 to certify a universal gate set for single-qubit quantum computation. For this, we rely on already proven Theorem 4 for self-testing of the S gate, and incorporate the Hadamard gate H and the T gate to the sequences considered in the protocol. There are, however, important differences from the case of the S gate certification. First, in order to include the Hadamard we also need to account for a possible complex conjugation, which is still in accordance with Definition 2. To certify the T gate, we would either need to modify Protocol 1 to include estimation of outcome probabilities, or as we do it here, change the goal of the certification. In particular, we show that using Protocol 1, we can certify implementation of the square root of the S gate, which can be either $T = |0\rangle\langle 0| + e^{i\frac{\pi}{4}}|1\rangle\langle 1|$, or $ZT = |0\rangle\langle 0| - e^{i\frac{\pi}{4}}|1\rangle\langle 1|$. Simultaneously, either of the two gates, T or ZT, in conjunction with the S gate and the Hadamard gate, constitute a universal gate set for single-qubit quantum computation.

We use the following gate sequences for certification of the quantum model $(|+\rangle\langle +|, \{S, S^\dagger, H, T\}, \{|+\rangle\langle +|, |-\rangle\langle -|\})$,

$$\mathcal{X} = \{\epsilon, ss^{-1}, s^{-1}s, ss, s^{-1}s^{-1}, shs, s^{-1}hs, hh, hsh, hth, sshth, tts\}, \quad (15)$$

where the labels $X = \{s, s^{-1}, h, t\}$ correspond to the gates in the self-explanatory way. Recall, that we read the sequences from left to right, e.g., in the sequence $s^{-1}hs$, the gate corresponding to s^{-1} is applied first. We again take $A = \{+, -\}$, and set the deterministic outcomes expected by Protocol 1 to

$$a_{\mathbf{x}} = \begin{cases} + & \text{for } \mathbf{x} \in \{\epsilon, ss^{-1}, s^{-1}s, shs, hh, hsh, hth\}, \\ - & \text{for } \mathbf{x} \in \{ss, s^{-1}s^{-1}, s^{-1}hs, sshth, tts\}. \end{cases} \quad (16)$$

We formalize our findings in this direction in the following theorem, which is an ideal self-testing type result.

Theorem 6. *If a single-qubit quantum model $(\tilde{\rho}, \{\tilde{\Lambda}_x\}_{x \in X}, \{\tilde{M}_+, \tilde{M}_-\})$, with $X = \{s, s^{-1}, h, t\}$*

passes a single repetition of Protocol 1 with probability 1, for $\mathcal{X}$ and $a_{\mathbf{x}}$ specified in Eq. (15) and Eq. (16), respectively, and for any sampling distribution such that $\mu(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathcal{X}$, then each quantum channel $\tilde{\Lambda}_x$ in the model is unitary, i.e., there exist $\tilde{U}_x \in \text{U}(2)$, such that

$$\tilde{\Lambda}_x(\cdot) = \tilde{U}_x(\cdot)\tilde{U}_x^\dagger, \quad \forall x \in X, \quad (17)$$

and there exists a unitary $U \in \text{U}(2)$ (with possible complex conjugation $(*)$), such that

$$\begin{aligned} \tilde{U}_s &= US^{(*)}U^\dagger, \\ \tilde{U}_{s^{-1}} &= US^{\dagger(*)}U^\dagger, \\ \tilde{U}_h &= UHU^\dagger, \end{aligned} \quad (18)$$

and either $\tilde{U}_t = UT^{(*)}U^\dagger$ or $\tilde{U}_t = UZT^{(*)}U^\dagger$. Moreover, for the same unitary U it holds that,

$$\tilde{\rho} = U|+\rangle\langle+|U^\dagger, \quad \tilde{M}_+ = U|+\rangle\langle+|U^\dagger. \quad (19)$$

In the statement of Theorem 6, we define the complex conjugation with respect to the computational basis.

Proof. We start with a brief overview of the proof steps. The starting point is to consider the sequences $\{\epsilon, ss^{-1}, s^{-1}s, ss, s^{-1}s^{-1}\}$, which achieve the self-testing result for the S, and $S^\dagger$ gates. By additionally considering all the sequences composed of $\{h, s, s^{-1}\}$, we can self-test the Hadamard gate. Lastly, we consider sequences involving t and prove that either the T gate or ZT is implemented when we pass this instruction label to the quantum computer.

The first step follows immediately from the proof of Theorem 4 for the special case $\varepsilon = 0$, which provides us with a unitary U such that the first two equations in Eq. (18) as well as Eq. (19) are satisfied. Note, that at this point we do not need additional complex conjugation. Let $|\psi\rangle\langle\psi| = \tilde{\rho}$ be the initial state and $|\phi\rangle\langle\phi| = \tilde{\Lambda}_{s^{-1}}(|\psi\rangle\langle\psi|)$. Following Theorem 4, we know that $\tilde{\Lambda}_s$ and $\tilde{\Lambda}_{s^{-1}}$ implement a unitary $\tilde{U}_s$ and its inverse, respectively, where

$$\tilde{U}_s = |\psi\rangle\langle\phi| + |\psi^\perp\rangle\langle\phi^\perp|. \quad (20)$$

Moreover, following the proof of Theorem 4, we also obtain that the bases $\{|\psi\rangle, |\psi^\perp\rangle\}$ and $\{|\phi\rangle, |\phi^\perp\rangle\}$ are mutually unbiased, and we can choose all the inner products of $\{|\psi\rangle, |\psi^\perp\rangle, |\phi\rangle, |\phi^\perp\rangle\}$ to be real (see Eq. (37)).

We continue the proof with self-testing of the Hadamard gate. The observed correlations for the strings $\mathbf{x} = shs$, $\mathbf{x} = s^{-1}hs$ imply that $\tilde{\Lambda}_h(|\phi\rangle\langle\phi|) = |\phi^\perp\rangle\langle\phi^\perp|$ and $\tilde{\Lambda}_h(|\phi^\perp\rangle\langle\phi^\perp|) = |\phi\rangle\langle\phi|$, i.e., $\tilde{\Lambda}_h$ maps an ONB to an ONB. From the input string $\mathbf{x} = hh$, we also determine that $\tilde{\Lambda}_h \circ \tilde{\Lambda}_h(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|$, which implies that $\tilde{\Lambda}_h(|\psi\rangle\langle\psi|)$ is a pure state (see Lemma 8). Combining these, we then conclude that

$\tilde{\Lambda}_h$ is a unitary channel (see Lemma 9). Moreover, the corresponding unitary operator $\tilde{U}_h$ must be of the form

$$\tilde{U}_h = |\phi\rangle\langle\phi^\perp| + e^{i\theta}|\phi^\perp\rangle\langle\phi| \quad (21)$$

for some $\theta \in \mathbb{R}$.

We can identify the phase θ by considering the sequence input string $\mathbf{x} = hsh$. Specifically, the observed deterministic behavior implies that $|\langle\psi|\tilde{U}_h\tilde{U}_s\tilde{U}_h|\psi\rangle| = 1$, with $\tilde{U}_s$ as in Eq. (20). This results in the condition

$$|1 + 2e^{i\theta} - e^{2i\theta}| = 2\sqrt{2}, \quad (22)$$

which is satisfied if and only if $e^{i\theta} = \pm i$. By applying the gauge unitary U , this leaves either of the two possibilities $U^\dagger\tilde{U}_hU \in \{H, XHX\}$ (up to a global phase). In the latter case, we absorb the X gate into the gauge unitary $U \mapsto UX$, at the cost of interchanging S and $S^\dagger$ (and an additional global phase), which effectively amounts to applying a complex conjugation to S and $S^\dagger$. Note, that adding X to the gauge unitary does not change the results for the state and the measurement in Eq. (19) since $|+\rangle, |-\rangle$ are eigenstates of X.

Finally, we also consider the sequences involving the 't' input. The observed correlations for the input strings $\mathbf{x} = hth$ and $\mathbf{x} = sshth$ imply that the channel $\tilde{\Lambda}_t$ maps an ONB $\{\tilde{U}_h|\psi\rangle, \tilde{U}_h|\psi^\perp\rangle\}$ to an ONB (in fact, to itself). Moreover, from the sequence $\mathbf{x} = tts$ we can deduce that $\tilde{\Lambda}_t(\psi)$ is a pure state (see Lemma 8). Moreover, since $|\langle\psi|\tilde{U}_h|\psi\rangle| = 1/\sqrt{2}$, we can invoke Lemma 9 to conclude that $\tilde{\Lambda}_t$ is a unitary channel. From $|\langle\psi|\tilde{U}_h\tilde{U}_t\tilde{U}_h|\psi\rangle| = 1$ we deduce, that after applying the gauge unitary (and a possible complex conjugation), we have

$$U^\dagger\tilde{U}_tU = |0\rangle\langle 0| + e^{i\varphi}|1\rangle\langle 1|, \quad (23)$$

for a suitable phase $\varphi \in \mathbb{R}$. This phase is constrained by the input string $\mathbf{x} = tts$, since $|\langle\psi^\perp|\tilde{U}_s\tilde{U}_t\tilde{U}_t|\psi\rangle| = 1$ is equivalent to $e^{2i\varphi} = i$, which leaves the two possibilities $\varphi \in \{\frac{\pi}{4}, \pi + \frac{\pi}{4}\}$, which we cannot distinguish further with Protocol 1, but which lets us to conclude that $\tilde{U}_t \in \{UT^{(*)}U^\dagger, UZT^{(*)}U^\dagger\}$ (up to the global phase). This finishes the proof. $\square$

It is possible to modify Protocol 1 for self-testing the T gate in the sense of Definition 2, by including sequences such as, e.g., $\mathbf{x} = t$. However, this means that the target statistics will stop being deterministic, and we will need to estimate the corresponding outcome probabilities up to some precision. At the same time, it does not mean that the overall sampling complexity should change drastically, because, at least in the ideal case, we will only need to distinguish between the two cases $|\langle+|T|+\rangle|^2 = \frac{1}{2} + \frac{1}{2\sqrt{2}}$, and $|\langle+|ZT|+\rangle|^2 = \frac{1}{2} - \frac{1}{2\sqrt{2}}$.

4 Conclusions and outlook

In this paper, we propose a novel method for certifying quantum gates within a practical scenario where a classical user interacts with a quantum server, considered as a black box. The method certifies the target gates together with the state preparation and measurement, thus making it free from SPAM errors. Here, we focus on single-qubit gates and prove soundness of certification for a gate set that is universal for single-qubit quantum computation, based on a few minimal assumptions. Moreover, for a relevant single-qubit phase gate, which corresponds experimentally to a $\pi/2$ -pulse, we show that the sample complexity of our method scales like $O(\varepsilon^{-1})$ with respect to the average gate infidelity ε .

Within the range of quantum system characterization techniques, the proposed method occupies a unique position and cannot be substituted by any existing tools. While direct gate certification and tomography are affected by SPAM errors, SPAM-robust characterization methods are not supported by the soundness guarantees if applied for certification. Self-testing either requires two isolated parts of an experimental apparatus or computational assumptions, both posing challenges in a way of it being applied in practice at the current level of technological development of quantum hardware. This work presents a fresh perspective that can inspire the development of practical and reliable certification techniques for testing quantum computers.

In this work, we focus on certification of single-qubit gates. Nevertheless, some of the introduced concepts have the potential for extension. Theorem 4,

can be used to construct a *fidelity witness* for quantum gates, if Protocol 1 is modified to estimate the acceptance probability. This will be particularly relevant for an experimental demonstration of the proposed method. The soundness proof of Theorem 6 can be extended to multi-qubit quantum gates, as our preliminary analysis suggests. However, this extension requires new techniques for the case of entangling gates, and, thus, deserves a separate study.

It also seems possible to translate some of the ideas from Ref. [41] to the framework of the dimension assumption, removing the requirement on the ideal preparation of the computational basis states, assumed therein. Finally, we find the connection between the classical simulability of quantum computation and the types of quantum gates which can be efficiently certified with deterministic measurement outcomes intriguing, which also deserves a separate investigation.

Acknowledgments

We thank Michał Oszmaniec, Costantino Budroni, and Ingo Roth for inspiring discussions. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project numbers 441423094, 236615297 - SFB 1119) and the Fujitsu Services GmbH as part of the endowed professorship “Quantum Inspired and Quantum Optimization”. Figure 1 is drawn using tikzpeople package, developed by Nils Fleischhacker. Figure 3 is generated in MATLAB.

Appendix

A Proof of Theorem 4

We repeat the statement of the theorem for convenience. We omit “tilde” over the implemented state, channels, and the measurement to keep the presentation simple.

Theorem 4. *Let $(\rho, \{\Lambda_s, \Lambda_{s^{-1}}\}, \{M_+, M_-\})$ be a single-qubit quantum model that passes with probability at least $1 - \varepsilon$ a single repetition of Protocol 1 with uniform sampling from the gate sequences (7) and for deterministic measurement outcomes (8). Then the quantum model is $O(\varepsilon)$ -close to the target model (6), i.e., there is a unitary $U \in U(2)$ such that*

$$\begin{aligned} F_{\text{avg}}(\Lambda_s, USU^\dagger) &\geq 1 - O(\varepsilon), \\ F_{\text{avg}}(\Lambda_{s^{-1}}, US^\dagger U^\dagger) &\geq 1 - O(\varepsilon), \\ \text{Tr}[\rho U|+\rangle\langle +|U^\dagger] &\geq 1 - \frac{15}{2}\varepsilon, \\ \|M_+ - U|+\rangle\langle +|U^\dagger\|_\infty &\leq \frac{5}{2}\varepsilon. \end{aligned} \tag{24}$$

Proof. Because the following proof is lengthy and technical in parts, we start by giving a general outline. The conclusions of the theorem follow from the condition $\mathbb{P}[\text{"pass"}] \geq 1 - \varepsilon$ and the dimension assumption, that is $\rho, M_+, M_- \in \mathcal{L}(\mathbb{C}^2)$, and $\Lambda_s : \mathcal{L}(\mathbb{C}^2) \rightarrow \mathcal{L}(\mathbb{C}^2)$, $\Lambda_{s^{-1}} : \mathcal{L}(\mathbb{C}^2) \rightarrow \mathcal{L}(\mathbb{C}^2)$. As a first step, we show that for small ε , the measurement effects M_+ and M_- are close to being rank-1 projectors, which we denote as ψ and $\psi^\perp$.

Next, we show that POVMs which one obtains by applying the adjoint maps $\Lambda_s^\dagger$ and $\Lambda_{s-1}^\dagger$ to ψ and $\psi^\perp$ are also close to be projective for small ε . We denote the corresponding projectors by ϕ and $\phi^\perp$. Importantly, we find that $\Lambda_s^\dagger(\psi) \approx \phi$ and $\Lambda_{s-1}^\dagger(\psi) \approx \phi^\perp$, and since the adjoint maps of channels are unital, also $\Lambda_s^\dagger(\psi^\perp) \approx \phi^\perp$ and $\Lambda_{s-1}^\dagger(\psi^\perp) \approx \phi$. Next, we obtain a partial characterization of the Choi state of the channel Λ_s in the basis of ψ and ϕ , with the leading terms which we denote as a_1, a_2, a_3, a_3^* corresponding to the subspace spanned by $|\psi\rangle|\phi\rangle$ and $|\psi^\perp\rangle|\phi^\perp\rangle$. Here, a_3 and a_3^* correspond to the off-diagonal terms of the matrix representation of $J(\Lambda_s)$, which at this point can only be upper-bounded by $\frac{1}{2}$. The case $|a_3| \approx \frac{1}{2}$ corresponds to Λ_s being a unitary channel. In order to show that actually $a_3 \approx \frac{1}{2}$, we use the conditions $\rho \approx \psi$ and $\Lambda_s(\rho) \approx \phi^\perp$. Finally, we find a suitable gauge unitary $U \in U(2)$ for which the condition for Λ_s in Eq. (24) follows. Because we obtain characterization of Λ_s and Λ_{s-1} in the same basis, the proof also easily extends to the channel Λ_{s-1} . The bounds for the state and the measurement in Eq. (24) for the chosen unitary are also immediate. Showing each step of the above sketch is, in principle, not too technical, but a lot of involving calculations in the proof are there to ensure the linear scaling of the bounds in Eq. (24) with respect to ε .

We start the proof by writing the probability of a quantum model, given by ρ , Λ_s , Λ_{s-1} , and $\{M_+, M_-\}$ passing a single repetition of the protocol.

$$\begin{aligned} \mathbb{P}[\text{"pass"}] = & \frac{1}{5} \left(\text{Tr}[M_+\rho] + \text{Tr}[M_+\Lambda_s \circ \Lambda_{s-1}(\rho)] + \text{Tr}[M_+\Lambda_{s-1} \circ \Lambda_s(\rho)] \right. \\ & \left. + \text{Tr}[M_-\Lambda_s \circ \Lambda_s(\rho)] + \text{Tr}[M_-\Lambda_{s-1} \circ \Lambda_{s-1}(\rho)] \right). \end{aligned} \quad (25)$$

For simplicity, let us take ε such that $\mathbb{P}[\text{"pass"}] \geq 1 - \frac{\varepsilon}{5}$, and rescale ε at the end of the proof. We separate the condition in Eq. (25) into $\text{Tr}[M_+\rho] \geq 1 - \varepsilon$ and

$$\text{Tr}[M_+\Lambda_s \circ \Lambda_{s-1}(\rho)] + \text{Tr}[M_+\Lambda_{s-1} \circ \Lambda_s(\rho)] - \text{Tr}[M_+\Lambda_s \circ \Lambda_s(\rho)] - \text{Tr}[M_+\Lambda_{s-1} \circ \Lambda_{s-1}(\rho)] \geq 2 - \varepsilon. \quad (26)$$

Let the eigendecomposition of M_+ be $M_+ = (1 - \lambda_+)\psi + \lambda_-\psi^\perp$, where $\psi := |\psi\rangle\langle\psi|$, $\psi^\perp := |\psi^\perp\rangle\langle\psi^\perp|$, $\langle\psi|\psi^\perp\rangle = 0$, and $\lambda_+ + \lambda_- \leq 1$. We can then substitute M_+ in Eq. (26) with $(1 - \lambda_+ - \lambda_-)\psi + \lambda_-\mathbb{1}$, and due to the normalization of states and $1 \geq (1 - \lambda_+ - \lambda_-)$, we arrive at the same condition as Eq. (26), but with ψ instead of M_+ . We also obtain that $\lambda_+ + \lambda_- \leq \frac{\varepsilon}{2}$, because we can upper-bound the expression, which is multiplied by $(1 - \lambda_+ - \lambda_-)$ on the left-hand side of Eq. (26) by 2. Next, for each trace, we move the second channel in the sequence to the measurement side, and denote the adjoint maps as $\Lambda_s^\dagger$ and $\Lambda_{s-1}^\dagger$. Grouping the terms together, we obtain

$$\text{Tr} \left[\left(\Lambda_s^\dagger(\psi) - \Lambda_{s-1}^\dagger(\psi) \right) \left(\Lambda_{s-1}(\rho) - \Lambda_s(\rho) \right) \right] \geq 2 - \varepsilon. \quad (27)$$

Let $\Lambda_s^\dagger(\psi) - \Lambda_{s-1}^\dagger(\psi) = \eta_+\phi - \eta_-\phi^\perp$, where $\phi := |\phi\rangle\langle\phi|$, $\phi^\perp := |\phi^\perp\rangle\langle\phi^\perp|$, $\langle\phi|\phi^\perp\rangle = 0$, and $\eta_+, \eta_- \in [-1, 1]$ due to the fact that POVM effects are positive semidefinite (PSD) and bounded. Inserting this eigendecomposition into Eq. (27), leads to

$$(\eta_+ + \eta_-) \text{Tr} \left[\phi \left(\Lambda_{s-1}(\rho) - \Lambda_s(\rho) \right) \right] \geq 2 - \varepsilon. \quad (28)$$

Since the trace in Eq. (28) can be at most 1, and each of η_- and η_+ are upper-bounded by 1, we conclude that $\eta_- \geq 1 - \varepsilon$ and $\eta_+ \geq 1 - \varepsilon$. From this conclusion, we arrive at a first set of important conditions that characterize the channels Λ_s and Λ_{s-1} , namely

$$\text{Tr}[\Lambda_s^\dagger(\psi)\phi] \geq 1 - \varepsilon, \quad \text{Tr}[\Lambda_s^\dagger(\psi^\perp)\phi^\perp] \geq 1 - \varepsilon, \quad \text{Tr}[\Lambda_{s-1}^\dagger(\psi)\phi^\perp] \geq 1 - \varepsilon, \quad \text{Tr}[\Lambda_{s-1}^\dagger(\psi^\perp)\phi] \geq 1 - \varepsilon. \quad (29)$$

Next, we focus on channel Λ_s and derive a partial characterization of its Choi state. We define the Choi-Jamiołkowski state [38, 39], or the Choi state for short, of a qubit channel Λ and the inverse Choi map with respect to the canonical product basis $(|i\rangle|j\rangle)_{i,j \in \{0,1\}}$ in $\mathbb{C}^4$ as

$$J(\Lambda) := \frac{1}{2} \sum_{i,j \in \{0,1\}} \Lambda(|i\rangle\langle j|) \otimes |i\rangle\langle j|, \quad \Lambda^\dagger(\cdot) = 2(\text{Tr}_1[(\cdot) \otimes \mathbb{1} J(\Lambda)])^\top, \quad (30)$$

where $\text{Tr}_1[\cdot]$ denotes the partial trace with respect to the first subsystem. Let us specify the matrix representation of $J(\Lambda_s)$ in the basis $\text{ONB}_1 := (|\psi\rangle|\phi\rangle^*, |\psi^\perp\rangle|\phi^\perp\rangle^*, |\psi\rangle|\phi^\perp\rangle^*, |\psi^\perp\rangle|\phi\rangle^*)$ as follows

$$[J(\Lambda_s)]_{\text{ONB}_1} = \begin{bmatrix} A & B \\ B^\dagger & C \end{bmatrix} := \begin{bmatrix} a_1 & a_3 & b_1 & b_2 \\ a_3^* & a_2 & b_3^* & -b_1^* \\ b_1^\dagger & b_3 & c_2 & c_3 \\ b_2^* & -b_1 & c_3^* & c_1 \end{bmatrix}, \quad (31)$$

where $A, B, C \in \mathbb{C}^{2 \times 2}$ represent the 2×2 blocks of $[J(\Lambda_s)]_{\text{ONB}_1}$, and $a_1, a_2, c_1, c_2 \in \mathbb{R}$, and $a_3, b_1, b_2, b_3, c_3 \in \mathbb{C}$ represent the entries. From the derived condition in Eq. (29), we have that $a_1 \geq \frac{1}{2} - \frac{\varepsilon}{2}$ and $a_2 \geq \frac{1}{2} - \frac{\varepsilon}{2}$. From the normalization condition $\text{Tr}_1[J(\Lambda_s)] = \frac{1}{2}$, we have that $c_1 = \frac{1}{2} - a_1$ and $c_2 = \frac{1}{2} - a_2$, and, therefore, $c_1 \leq \frac{\varepsilon}{2}$ and $c_2 \leq \frac{\varepsilon}{2}$. From the PSD condition $[J(\Lambda_s)]_{\text{ONB}_1} \geq 0$, we obtain that

$$|a_3| \leq \frac{1}{2}, \quad |c_3| \leq \frac{\varepsilon}{2}, \quad |b_1| \leq \frac{\sqrt{\varepsilon}}{2}, \quad |b_2| \leq \frac{\sqrt{\varepsilon}}{2}, \quad |b_3| \leq \frac{\sqrt{\varepsilon}}{2}. \quad (32)$$

We can use the above estimates to upper-bound the unwanted terms, i.e., all except for the ones in submatrix A , in channel Λ_s . However, they are not sufficient for obtaining the linear scaling in ε of the bounds in Eq. (24). We will also need a tighter upper-bound on $|b_2 + b_3|$.

In order to derive a tighter upper-bound on $|b_2 + b_3|$, we use the following constraint on the blocks A, B, C that form a PSD matrix,

$$|\langle v|B|w\rangle|^2 \leq \langle v|A|v\rangle \langle w|C|w\rangle, \quad \forall |v\rangle, |w\rangle \in \mathbb{C}^2. \quad (33)$$

This result can be found in Ref. [42] (Theorem 7.7.7), and we also provide a proof of Eq. (33) in Appendix B for completeness. Let us first take $|v\rangle = -\frac{a_3^*}{|a_3|}|\phi^\perp\rangle^* + |\phi\rangle^*$ and $|w\rangle = |\phi^\perp\rangle^*$. The condition in Eq. (33) then implies

$$\left| -\frac{a_3}{|a_3|}b_1^* + b_3 \right|^2 \leq (a_1 + a_2 - 2|a_3|)c_2 \leq (1 - 2|a_3|)\frac{\varepsilon}{2}. \quad (34)$$

Next, take $|v\rangle = |\phi^\perp\rangle^* - \frac{a_3}{|a_3|}|\phi\rangle^*$ and $|w\rangle = |\phi\rangle^*$, which results in a similar condition,

$$\left| b_2 + \frac{a_3}{|a_3|}b_1^* \right|^2 \leq (a_1 + a_2 - 2|a_3|)c_1 \leq (1 - 2|a_3|)\frac{\varepsilon}{2}. \quad (35)$$

Using the triangular inequality, we then obtain a condition

$$|b_2 + b_3| \leq \sqrt{2\varepsilon} \sqrt{1 - 2|a_3|}, \quad (36)$$

which we use later in the proof.

We continue the proof by returning to Eq. (28) and using the condition $\eta_+ + \eta_- \leq 2$, obtain that $\text{Tr}[\phi\Lambda_{s^{-1}}(\rho)] \geq 1 - \frac{\varepsilon}{2}$ and $\text{Tr}[\phi\Lambda_s(\rho)] \leq \frac{\varepsilon}{2}$. Again, we focus on channel Λ_s first, and rewrite the aforementioned condition for it as

$$\text{Tr}[\Lambda_s^\dagger(\phi^\perp)\rho] \geq 1 - \frac{\varepsilon}{2}. \quad (37)$$

This is the second important condition alongside Eq. (29) that allows us to characterize channel Λ_s .

It is useful at this point of the proof to fix the relative phases between the vectors $|\phi\rangle, |\phi^\perp\rangle, |\psi\rangle$, and $|\psi^\perp\rangle$. Without loss of generality, we set

$$\langle \psi|\phi\rangle = \langle \psi^\perp|\phi^\perp\rangle = |\langle \psi|\phi\rangle|, \quad -\langle \psi|\phi^\perp\rangle = \langle \psi^\perp|\phi\rangle = |\langle \psi^\perp|\phi\rangle|. \quad (38)$$

Let us first express ρ in the basis $\{|\psi\rangle, |\psi^\perp\rangle\}$ as

$$\rho = d_1\psi + d_2\psi^\perp + d_3|\psi\rangle\langle\psi^\perp| + d_3^*|\psi^\perp\rangle\langle\psi|. \quad (39)$$

From the condition $\text{Tr}[M_+\rho] \geq 1 - \varepsilon$, which we obtained directly from Eq. (25), and from the condition on the eigenvalues of M_+ , namely, $\lambda_+ + \lambda_- \leq \frac{\varepsilon}{2}$, we obtain that $d_1 = \text{Tr}[\psi\rho] \geq 1 - \frac{3}{2}\varepsilon$, and, consequently, $d_2 \leq \frac{3}{2}\varepsilon$. From $\rho \geq 0$, we obtain additionally that $|d_3| \leq \mathcal{O}(\sqrt{\varepsilon})$.

From now on, we express the bounds using the Big-O notation, because we are interested in the scaling w.r.t. ε , and we estimate the constants in our numerical studies in Section 3. Using the expansion in Eq. (39), we can reduce the condition in Eq. (37) to

$$\text{Tr}[\Lambda_s^\dagger(\phi^\perp)\psi] + 2\text{Re}[d_3\text{Tr}[\Lambda_s^\dagger(\phi^\perp)|\psi\rangle\langle\psi^\perp|]] \geq 1 - \mathcal{O}(\varepsilon). \quad (40)$$

We do not simply use the upper bound of $\mathcal{O}(\sqrt{\varepsilon})$ on the second term in Eq. (40), but instead carefully analyze both terms. We use the expansion of $\phi^\perp$ in the basis of $(|\psi\rangle, |\psi^\perp\rangle)$ to write the POVM effect $\Lambda_s^\dagger(\phi^\perp)$ as

$$\Lambda_s^\dagger(\phi^\perp) = |\langle \psi^\perp|\phi\rangle|^2\Lambda_s^\dagger(\psi) + |\langle \psi|\phi\rangle|^2\Lambda_s^\dagger(\psi^\perp) - |\langle \psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \left(\Lambda_s^\dagger(|\psi\rangle\langle\psi^\perp|) + \Lambda_s^\dagger(|\psi^\perp\rangle\langle\psi|) \right). \quad (41)$$

We can use Eq. (41) and the partial characterization of Λ_s in Eq. (31) to express $\Lambda_s^\dagger(\phi^\perp)$ in the basis $\text{ONB}_2 := (|\phi\rangle, |\phi^\perp\rangle)$,

$$\begin{aligned} [\Lambda_s^\dagger(\phi^\perp)]_{\text{ONB}_2} = & 2 \begin{bmatrix} c_1 |\langle\psi|\phi\rangle|^2 & -c_3 |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \\ -c_3^* |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| & c_2 |\langle\psi^\perp|\phi\rangle|^2 \end{bmatrix} \\ & + 2 \begin{bmatrix} a_1 |\langle\psi^\perp|\phi\rangle|^2 - 2 \text{Re}[b_2] |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| & b_1^* (|\langle\psi^\perp|\phi\rangle|^2 - |\langle\psi|\phi\rangle|^2) - a_3^* |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \\ b_1 (|\langle\psi^\perp|\phi\rangle|^2 - |\langle\psi|\phi\rangle|^2) - a_3 |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| & a_2 |\langle\psi|\phi\rangle|^2 - 2 \text{Re}[b_3] |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \end{bmatrix}. \end{aligned} \quad (42)$$

The first summand in the above expression can be safely ignored, because its contribution is of the order of $O(\varepsilon)$, due to the upper-bounds on its entries. On the other hand, the matrix representations of ψ and $|\psi\rangle\langle\psi^\perp|$ in the basis ONB_2 , are

$$\begin{aligned} [\psi]_{\text{ONB}_2} &= \begin{bmatrix} |\langle\psi|\phi\rangle|^2 & -|\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \\ -|\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| & |\langle\psi^\perp|\phi\rangle|^2 \end{bmatrix}, \\ [|\psi\rangle\langle\psi^\perp|]_{\text{ONB}_2} &= \begin{bmatrix} |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| & |\langle\psi|\phi\rangle|^2 \\ -|\langle\psi^\perp|\phi\rangle|^2 & -|\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \end{bmatrix}. \end{aligned} \quad (43)$$

Using Eq. (42) and Eq. (43), we can upper-bound the first term on the left-hand side of Eq. (40) as

$$\begin{aligned} \text{Tr}[\Lambda_s^\dagger(\phi^\perp)\psi] &\leq O(\varepsilon) + 2(a_1 + a_2 + 2 \text{Re}[a_3]) |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle|^2 \\ &\quad - 4 |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \left(|\langle\psi|\phi\rangle|^2 \text{Re}[b_2] + |\langle\psi^\perp|\phi\rangle|^2 \text{Re}[b_3] + (|\langle\psi^\perp|\phi\rangle|^2 - |\langle\psi|\phi\rangle|^2) \text{Re}[b_1] \right). \end{aligned} \quad (44)$$

Similarly, the second term on the left-hand side of Eq. (40) can be upper-bounded as

$$\begin{aligned} \text{Re} [d_3 \text{Tr}[\Lambda_s^\dagger(\phi^\perp)|\psi\rangle\langle\psi^\perp|]] &\leq O(\varepsilon) + 2 \text{Re} \left[d_3 |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \left((a_1 + a_3^*) |\langle\psi^\perp|\phi\rangle|^2 - (a_2 + a_3) |\langle\psi|\phi\rangle|^2 \right. \right. \\ &\quad \left. \left. + 2 |\langle\psi|\phi\rangle\langle\psi^\perp|\phi\rangle| \text{Re}[b_3 - b_2] \right) + d_3 \left(|\langle\psi^\perp|\phi\rangle|^2 - |\langle\psi|\phi\rangle|^2 \right) \left(b_1 |\langle\psi|\phi\rangle|^2 - b_1^* |\langle\psi^\perp|\phi\rangle|^2 \right) \right]. \end{aligned} \quad (45)$$

To simplify the estimates of the quantities in Eq. (44) and Eq. (45), we introduce the last bit of notation, namely two functions $f : [0, 1] \rightarrow [-1, 1]$ and $g : [0, 1] \rightarrow [0, 2]$, such that

$$|\langle\psi|\phi\rangle|^2 = \frac{1-f(\varepsilon)}{2}, \quad |\langle\psi^\perp|\phi\rangle|^2 = \frac{1+f(\varepsilon)}{2}, \quad \text{Re}[a_3] = \frac{1-g(\varepsilon)}{2}. \quad (46)$$

Note, that even though we use ε as the argument for functions f and g , there is no loss of generality in making the above assignments. In particular, we can take $g(\varepsilon) \geq 0$, because from Eq. (32), we know that $|a_3| \leq \frac{1}{2}$. Due to the same reason, we can upper-bound the absolute value of the imaginary part of a_3 as $|\text{Im}[a_3]| \leq \sqrt{\frac{g(\varepsilon)}{2}}$.

Using the new notations in Eq. (46), as well as the upper bounds in Eq. (32) and Eq. (36), we can simplify the bound in Eq. (44) as

$$\text{Tr}[\Lambda_s^\dagger(\phi^\perp)\psi] \leq \frac{1}{2}(1-f(\varepsilon)^2)(2-g(\varepsilon)) + \sqrt{1-f(\varepsilon)^2} \sqrt{g(\varepsilon)} \sqrt{2\varepsilon} + 2|f(\varepsilon)|\sqrt{\varepsilon} + O(\varepsilon). \quad (47)$$

Similarly, the bound in Eq. (45) can be simplified as

$$\text{Re} [d_3 \text{Tr}[\Lambda_s^\dagger(\phi^\perp)|\psi\rangle\langle\psi^\perp|]] \leq O(\sqrt{\varepsilon}) \sqrt{1-f(\varepsilon)^2} (|f(\varepsilon)| + \sqrt{g(\varepsilon)}) + O(\varepsilon). \quad (48)$$

Combining these two bounds together and inserting them back to the condition in Eq. (40), we finally arrive at

$$(|f(\varepsilon)| - O(\sqrt{\varepsilon}))^2 + \frac{1}{2} \left(\sqrt{1-f(\varepsilon)^2} \sqrt{g(\varepsilon)} - O(\sqrt{\varepsilon}) \right)^2 \leq O(\varepsilon). \quad (49)$$

This allows us to deduce that $|f(\varepsilon)| \leq O(\sqrt{\varepsilon})$ and $g(\varepsilon) \leq O(\varepsilon)$.

As the final part of the proof, we choose the gauge unitary U to be

$$U = |\psi\rangle\langle+| - i|\psi^\perp\rangle\langle-|. \quad (50)$$

Up to this gauge, the ideal gate S takes the form

$$USU^\dagger = \frac{e^{i\frac{\pi}{4}}}{\sqrt{2}} (\psi + \psi^\perp + |\psi\rangle\langle\psi^\perp| - |\psi^\perp\rangle\langle\psi|). \quad (51)$$

Consequently, we find the Choi state vector $|J(USU^\dagger)\rangle$, which we define as $|J(USU^\dagger)\rangle\langle J(USU^\dagger)| := J(USU^\dagger)$,

$$\begin{aligned} |J(USU^\dagger)\rangle &= \frac{1}{\sqrt{2}} USU^\dagger \otimes \mathbb{1} (|0\rangle|0\rangle + |1\rangle|1\rangle) \\ &= \frac{e^{i\frac{\pi}{4}}}{2} \left((|\langle\psi|\phi\rangle| + |\langle\psi^\perp|\phi\rangle|)(|\psi\rangle|\phi\rangle^* + |\psi^\perp\rangle|\phi^\perp\rangle^*) + (|\langle\psi|\phi\rangle| - |\langle\psi^\perp|\phi\rangle|)(|\psi\rangle|\phi^\perp\rangle^* - |\psi^\perp\rangle|\phi\rangle^*) \right). \end{aligned} \quad (52)$$

Its matrix representation in the ONB_1 is

$$[|J(USU^\dagger)\rangle]_{\text{ONB}_1} = \frac{e^{i\frac{\pi}{4}}}{2} \begin{bmatrix} |\langle\psi|\phi\rangle| + |\langle\psi^\perp|\phi\rangle| \\ |\langle\psi|\phi\rangle| + |\langle\psi^\perp|\phi\rangle| \\ |\langle\psi|\phi\rangle| - |\langle\psi^\perp|\phi\rangle| \\ -(|\langle\psi|\phi\rangle| - |\langle\psi^\perp|\phi\rangle|) \end{bmatrix}. \quad (53)$$

Having the explicit forms of the Choi states in Eq. (31) and Eq. (53) allows us to estimate their inner product,

$$\text{Tr} [J(\Lambda_s) J(USU^\dagger)] = \frac{1}{4} \left((|\langle\psi|\phi\rangle| + |\langle\psi^\perp|\phi\rangle|)^2 (a_1 + a_2 + 2 \text{Re}[a_3]) \right. \quad (54)$$

$$\left. + (|\langle\psi|\phi\rangle|^2 - |\langle\psi^\perp|\phi\rangle|^2) (2 \text{Re}[2b_1 - b_2 + b_3]) \right) + O(\varepsilon) \quad (55)$$

$$\geq \frac{1}{4} \left((1 + \sqrt{1 - f(\varepsilon)^2}) (1 - \varepsilon + 1 - g(\varepsilon)) - |f(\varepsilon)| O(\sqrt{\varepsilon}) \right) + O(\varepsilon). \quad (56)$$

Inserting the bounds $|f(\varepsilon)| \leq O(\sqrt{\varepsilon})$ and $g(\varepsilon) \leq O(\varepsilon)$ leads to the lower bound of $1 - O(\varepsilon)$ on the inner product of the Choi states of Λ_s and the target unitary channel with the unitary operator $USU^\dagger$. This inner product is sometimes referred to as the entanglement fidelity [2], which is related to the average gate fidelity through a known relation [2],

$$F_{\text{avg}}(\Lambda_s, USU^\dagger) = \frac{2}{3} \text{Tr} [J(\Lambda_s) J(USU^\dagger)] + \frac{1}{3}. \quad (57)$$

Equation (57) leads directly to the first claim of the theorem in Eq. (24).

The proof for channel $\Lambda_{s^{-1}}$ follows exactly the same steps as for channel Λ_s . It is important, however, that the lower bound on $\text{Tr} [J(\Lambda_{s^{-1}}) J(US^\dagger U^\dagger)]$ is shown to hold for the same gauge unitary U in Eq. (50). The main difference from the case of Λ_s , is that roles of states ϕ and $\phi^\perp$ are swapped, and in ONB_1 , the matrix representation of $J(\Lambda_{s^{-1}})$ has the leading terms in the block C rather than the block A , if we look at Eq. (31). We can notice that the matrix representation of the Choi state vector $|J(US^\dagger U^\dagger)\rangle$ in the same basis is

$$[|J(US^\dagger U^\dagger)\rangle]_{\text{ONB}_1} = \frac{e^{-i\frac{\pi}{4}}}{2} \begin{bmatrix} |\langle\psi|\phi\rangle| - |\langle\psi^\perp|\phi\rangle| \\ |\langle\psi|\phi\rangle| - |\langle\psi^\perp|\phi\rangle| \\ -(|\langle\psi|\phi\rangle| + |\langle\psi^\perp|\phi\rangle|) \\ |\langle\psi|\phi\rangle| + |\langle\psi^\perp|\phi\rangle| \end{bmatrix}, \quad (58)$$

again with the leading terms in the lower half of the vector. Apart from that, the reasoning is exactly the same, and the second claim in Eq. (24) follows.

As for the third and fourth claims in Eq. (24), we notice that $U|+\rangle\langle+|U^\dagger = \psi$, and since we already showed that $\text{Tr}[\rho\psi] \geq 1 - \frac{3}{2}\varepsilon$ when characterizing ρ in Eq. (39), we directly conclude that $\text{Tr}[\rho U|+\rangle\langle+|U^\dagger] \geq 1 - \frac{3}{2}\varepsilon$. Since $M_+ = (1 - \lambda_+)\psi + \lambda_-\psi^\perp$, we also immediately conclude that $\|M_+ - U|+\rangle\langle+|U^\dagger\|_\infty = \max\{\lambda_+, \lambda_-\} \leq \frac{\varepsilon}{2}$. Note, that at the beginning of the proof we rescaled ε by the factor of 5. This finishes the proof. $\square$

B Supporting Lemmata

In this section of Appendix, we list the supporting lemmata.

Lemma 7. For a PSD matrix $\Gamma = \begin{bmatrix} A & B \\ B^\dagger & C \end{bmatrix}$, with $A, C, B \in \mathbb{C}^{n \times n}$ it holds that

$$|\langle v|B|w\rangle|^2 \leq \langle v|A|v\rangle \langle w|C|w\rangle, \quad \forall |v\rangle, |w\rangle \in \mathbb{C}^n. \quad (59)$$

Proof. Let $K := |0\rangle|v\rangle\langle 0| + |1\rangle|w\rangle\langle 1| \in \mathbb{C}^{2n \times 2}$. We can write $\Gamma = |0\rangle\langle 0| \otimes A + |0\rangle\langle 1| \otimes B + |1\rangle\langle 0| \otimes B^\dagger + |1\rangle\langle 1| \otimes C$, and express $K^\dagger \Gamma K$,

$$\begin{aligned} K^\dagger \Gamma K &= (|0\rangle\langle 0| \langle v| + |1\rangle\langle 1| \langle w|) (|0\rangle\langle 0| \otimes A + |0\rangle\langle 1| \otimes B + |1\rangle\langle 0| \otimes B^\dagger + |1\rangle\langle 1| \otimes C) (|0\rangle|v\rangle\langle 0| + |1\rangle|w\rangle\langle 1|) \\ &= |0\rangle\langle 0| \langle v|A|v\rangle + |0\rangle\langle 1| \langle v|B|w\rangle + |1\rangle\langle 0| \langle w|B^\dagger|v\rangle + |1\rangle\langle 1| \langle w|C|w\rangle. \end{aligned} \quad (60)$$

The matrix representation of $K^\dagger \Gamma K$ in the computational basis is therefore $K^\dagger \Gamma K = \begin{bmatrix} \langle v|A|v\rangle & \langle v|B|w\rangle \\ \langle w|B^\dagger|v\rangle & \langle w|C|w\rangle \end{bmatrix}$. Since $\Gamma \geq 0$, then also $K^\dagger \Gamma K \geq 0$, since $K^\dagger(\cdot)K$ is completely positive (CP). The claim of the lemma then follows from non-negativity of the determinant of $K^\dagger \Gamma K$. $\square$

In Ref. [42], the above lemma is stated as part of a theorem (Theorem 7.7.7), which holds for $A > 0$ and $C > 0$. Therefore, we preset the proof above for completeness, to account for the cases of non-invertible A and C .

Lemma 8. *Given a qubit channel $\Lambda : \mathcal{L}(\mathbb{C}^2) \rightarrow \mathcal{L}(\mathbb{C}^2)$, if $\Lambda \circ \Lambda(|\psi\rangle\langle\psi|) = |\phi\rangle\langle\phi|$, for any two $|\psi\rangle, |\phi\rangle \in \mathbb{C}^2$, then $\Lambda(|\psi\rangle\langle\psi|)$ is a pure state.*

Proof. Assume the opposite, that is $\Lambda(|\psi\rangle\langle\psi|) = \lambda|\theta\rangle\langle\theta| + (1-\lambda)|\theta^\perp\rangle\langle\theta^\perp|$ for some $|\theta\rangle \in \mathbb{C}^2$, and $\lambda \in (0, 1)$. Then, from linearity it must hold that $\lambda\Lambda(|\theta\rangle\langle\theta|) + (1-\lambda)\Lambda(|\theta^\perp\rangle\langle\theta^\perp|) = |\phi\rangle\langle\phi|$, which is only possible if $\Lambda(|\theta\rangle\langle\theta|) = \Lambda(|\theta^\perp\rangle\langle\theta^\perp|)$, and hence $\Lambda(\mathbb{1}) = 2|\phi\rangle\langle\phi|$, which means that Λ is a measure-and-prepare channel, and, in particular, $\Lambda(|\psi\rangle\langle\psi|) = |\phi\rangle\langle\phi|$. Indeed, if $\Lambda(|\psi\rangle\langle\psi|) \neq |\phi\rangle\langle\phi|$, then due to $\Lambda(\mathbb{1}) = 2|\phi\rangle\langle\phi|$, we would have that $\Lambda(|\psi^\perp\rangle\langle\psi^\perp|) = 2|\phi\rangle\langle\phi| - \Lambda(|\psi\rangle\langle\psi|)$ is not PSD. We reach the contradiction, because we assumed that $\Lambda(|\psi\rangle\langle\psi|)$ is not pure. $\square$

Lemma 9. *Let $\Lambda : \mathcal{L}(\mathbb{C}^2) \rightarrow \mathcal{L}(\mathbb{C}^2)$ be a qubit channel which maps an ONB $\{|\psi\rangle, |\psi^\perp\rangle\}$ to an ONB in $\mathbb{C}^2$. Let further $\Lambda(|\varphi\rangle\langle\varphi|)$ be a pure state for some other state vector $|\varphi\rangle$, such that $0 < |\langle\varphi|\psi\rangle| < 1$. Then the channel Λ is unitary.*

Proof. Let $\{|\phi\rangle, |\phi^\perp\rangle\}$ be an ONB such that $|\phi\rangle\langle\phi| = \Lambda(|\psi\rangle\langle\psi|)$ and $|\phi^\perp\rangle\langle\phi^\perp| = \Lambda(|\psi^\perp\rangle\langle\psi^\perp|)$. From the CPTP condition, we conclude that $\Lambda(|\psi\rangle\langle\psi^\perp|) = z|\phi\rangle\langle\phi^\perp|$, and $\Lambda(|\psi^\perp\rangle\langle\psi|) = z^*|\phi^\perp\rangle\langle\phi|$, for some $z \in \mathbb{C}$, with $|z| \leq 1$.

Let $|\varphi\rangle = \sqrt{a}|\psi\rangle + \sqrt{1-a}|\psi^\perp\rangle$ for some $a \in \mathbb{R}$ (which we can always achieve by fixing the global phases of $|\psi\rangle$ and $|\psi^\perp\rangle$), and write

$$\Lambda(|\varphi\rangle\langle\varphi|) = a|\phi\rangle\langle\phi| + (1-a)|\phi^\perp\rangle\langle\phi^\perp| + z\sqrt{a(1-a)}|\phi\rangle\langle\phi^\perp| + z^*\sqrt{a(1-a)}|\phi^\perp\rangle\langle\phi|. \quad (61)$$

The purity of $\Lambda(|\varphi\rangle\langle\varphi|)$ in Eq. (61) leads to the condition

$$1 = \text{Tr}[(\Lambda(|\varphi\rangle\langle\varphi|))^2] = a^2 + (1-a)^2 + 2a(1-a)|z|^2. \quad (62)$$

From the assumptions on $|\varphi\rangle$, we have $0 < a < 1$, and hence $|z| = 1$. From here it is straightforward to see that $\Lambda(\cdot) = U(\cdot)U^\dagger$ for the unitary $U = |\phi\rangle\langle\psi| + z^*|\phi^\perp\rangle\langle\psi^\perp|$. $\square$

References

- [1] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, *Quantum certification and benchmarking*, *Nat. Rev. Phys.* **2**, 382 (2020), [arXiv:1910.06343 \[quant-ph\]](#).
- [2] M. Kliesch and I. Roth, *Theory of quantum system certification*, *PRX Quantum* **2**, 010201 (2021), tutorial, [arXiv:2010.05925 \[quant-ph\]](#).
- [3] I. L. Chuang and M. A. Nielsen, *Prescription for experimental determination of the dynamics of a quantum black box*, *Journal of Modern Optics* **44**, 2455 (1997), [arXiv:quant-ph/9610001](#).
- [4] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, *Quantum-process tomography: Resource analysis of different strategies*, *Phys. Rev. A* **77**, 032322 (2008), [arXiv:quant-ph/0702131](#).
- [5] Y.-C. Liu, J. Shang, X.-D. Yu, and X. Zhang, *Efficient verification of quantum processes*, *Phys. Rev. A* **101**, 042315 (2020), [arXiv:1910.13730 \[quant-ph\]](#).
- [6] H. Zhu and H. Zhang, *Efficient verification of quantum gates with local operations*, *Phys. Rev.*
- [A](#) **101**, 042316 (2020), [arXiv:1910.14032 \[quant-ph\]](#).
- [7] P. Zeng, Y. Zhou, and Z. Liu, *Quantum gate verification and its application in property testing*, *Physical Review Research* **2**, 023306 (2020), [arXiv:1911.06855 \[quant-ph\]](#).
- [8] S. T. Merkel, J. M. Gambetta, J. A. Smolin, S. Poletto, A. D. Córcoles, B. R. Johnson, C. A. Ryan, and M. Steffen, *Self-consistent quantum process tomography*, *Phys. Rev. A* **87**, 062119 (2013), [arXiv:1211.0322 \[quant-ph\]](#).
- [9] R. Blume-Kohout, J. King Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, *Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit*, [arXiv:1310.4492 \[quant-ph\]](#).
- [10] R. Brieger, I. Roth, and M. Kliesch, *Compressive gate set tomography*, *PRX Quantum* **4**, 010325 (2023), [arXiv:2112.05176 \[quant-ph\]](#).
- [11] J. Emerson, R. Alicki, and K. Życzkowski, *Scalable noise estimation with random unitary operators*, *J. Opt. B* **7**, S347 (2005), [arXiv:quant-ph/0503243](#).
- [12] B. Lévi, C. C. López, J. Emerson, and D. G.

- Cory, *Efficient error characterization in quantum information processing*, *Phys. Rev. A* **75**, 022314 (2007), [arXiv:quant-ph/0608246](#).
- [13] C. Dankert, R. Cleve, J. Emerson, and E. Livine, *Exact and approximate unitary 2-designs and their application to fidelity estimation*, *Phys. Rev. A* **80**, 012304 (2009), [arXiv:quant-ph/0606161](#).
- [14] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, *Symmetrized characterization of noisy quantum processes*, *Science* **317**, 1893 (2007), [arXiv:0707.0685 \[quant-ph\]](#).
- [15] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, *Randomized benchmarking of quantum gates*, *Phys. Rev. A* **77**, 012307 (2008), [arXiv:0707.0963 \[quant-ph\]](#).
- [16] E. Magesan, J. M. Gambetta, and J. Emerson, *Characterizing quantum gates via randomized benchmarking*, *Phys. Rev. A* **85**, 042311 (2012), [arXiv:1109.6887 \[quant-ph\]](#).
- [17] J. Helsen, I. Roth, E. Onorati, A. H. Werner, and J. Eisert, *A general framework for randomized benchmarking*, *PRX Quantum* **3**, 020357 (2022), [arXiv:2010.07974 \[quant-ph\]](#).
- [18] M. Heinrich, M. Kliesch, and I. Roth, *Randomized benchmarking with random quantum circuits*, [arXiv:2212.06181 \[quant-ph\]](#) (2022).
- [19] D. Mayers and A. Yao, *Self testing quantum apparatus*, *Quantum Info. Comput.* **4**, 273–286 (2004), [arXiv:quant-ph/0307205](#).
- [20] I. Šupić and J. Bowles, *Self-testing of quantum systems: a review*, *Quantum* **4**, 337 (2020), [arXiv:1904.10042 \[quant-ph\]](#).
- [21] P. Sekatski, J.-D. Bancal, S. Wagner, and N. Sangouard, *Certifying the building blocks of quantum computers from Bell’s theorem*, *Phys. Rev. Lett.* **121**, 180505 (2018), [arXiv:1802.02170 \[quant-ph\]](#).
- [22] S. Wagner, J.-D. Bancal, N. Sangouard, and P. Sekatski, *Device-independent characterization of quantum instruments*, *Quantum* **4**, 243 (2020), [arXiv:1812.02628 \[quant-ph\]](#).
- [23] S. Sarkar, *Model-independent inference of quantum interaction from statistics*, *Phys. Rev. A* **110**, L020402 (2024), [arXiv:2402.08003 \[quant-ph\]](#).
- [24] F. Magniez, D. Mayers, M. Mosca, and H. Olivier, *Self-testing of quantum circuits*, in *Automata, Languages and Programming*, edited by M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006) pp. 72–83, [arXiv:quant-ph/0512111](#).
- [25] B. Reichardt, F. Unger, and U. Vazirani, *Classical command of quantum systems*, *Nature* **496**, 456 (2013), [arXiv:1209.0449 \[quant-ph\]](#).
- [26] T. Metger and T. Vidick, *Self-testing of a single quantum device under computational assumptions*, *Quantum* **5**, 544 (2021), [arXiv:2001.09161 \[quant-ph\]](#).
- [27] U. Mahadev, *Classical verification of quantum computations*, in *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)* (2018) pp. 259–267, [arXiv:1804.01082 \[quant-ph\]](#).
- [28] R. Stricker, J. Carrasco, M. Ringbauer, L. Postler, M. Meth, C. Edmunds, P. Schindler, R. Blatt, P. Zoller, B. Kraus, and T. Monz, *Towards experimental classical verification of quantum computation*, *Quantum Science and Technology* **9**, 02LT01 (2024), [arXiv:2203.07395 \[quant-ph\]](#).
- [29] H.-Y. R. Huang, S. T. Flammia, and J. Preskill, *Foundations for learning from noisy quantum experiments* (2022), presented at QIP 2022, Paderborn, California, [arXiv:2204.13691 \[quant-ph\]](#).
- [30] K. Mohan, A. Tavakoli, and N. Brunner, *Sequential random access codes and self-testing of quantum measurement instruments*, *New Journal of Physics* **21**, 083034 (2019), [arXiv:1905.06726 \[quant-ph\]](#).
- [31] N. Miklin, J. J. Borkala, and M. Pawłowski, *Semi-device-independent self-testing of unsharp measurements*, *Phys. Rev. Res.* **2**, 033014 (2020), [arXiv:1903.12533 \[quant-ph\]](#).
- [32] A. Tavakoli, M. Smania, T. Vértesi, N. Brunner, and M. Bourennane, *Self-testing nonprojective quantum measurements in prepare-and-measure experiments*, *Science Advances* **6**, eaaw6664 (2020), [arXiv:1811.12712 \[quant-ph\]](#).
- [33] N. Miklin and M. Oszmaniec, *A universal scheme for robust self-testing in the prepare-and-measure scenario*, *Quantum* **5**, 424 (2021), [arXiv:2003.01032 \[quant-ph\]](#).
- [34] H. Anwer, S. Muhammad, W. Cherifi, N. Miklin, A. Tavakoli, and M. Bourennane, *Experimental characterization of unsharp qubit observables and sequential measurement incompatibility via quantum random access codes*, *Phys. Rev. Lett.* **125**, 080403 (2020), [arXiv:2001.04768 \[quant-ph\]](#).
- [35] M. Navascués, K. F. Pál, T. Vértesi, and M. Araújo, *Self-testing in prepare-and-measure scenarios and a robust version of Wigner’s theorem*, *Phys. Rev. Lett.* **131**, 250802 (2023), [arXiv:2306.00730 \[quant-ph\]](#).
- [36] E. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Vieweg+Teubner Verlag, 1931).
- [37] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, *Quantum dynamics of single trapped ions*, *Rev. Mod. Phys.* **75**, 281 (2003).
- [38] M.-D. Choi, *Completely positive linear maps on complex matrices*, *Lin. Alg. App.* **10**, 285 (1975).
- [39] A. Jamiolkowski, *Linear transformations which preserve trace and positive semidefiniteness of operators*, *Rep. Math. Phys.* **3**, 275 (1972).

- [40] A. Gočanin, I. Šupić, and B. Dakić, *Sample-efficient device-independent quantum state verification and certification*, [PRX Quantum](#) **3**, 010317 (2022), [arXiv:2105.05832 \[quant-ph\]](#).
- [41] W. van Dam, F. Magniez, M. Mosca, and M. Santha, *Self-testing of universal and fault-tolerant sets of quantum gates*, in *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, STOC00 (ACM, 2000) [arXiv:quant-ph/9904108](#).
- [42] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, 1985).