Article

Scalable noisy quantum circuits for biased-noise qubits

Check for updates

Marco Fellous-Asiani [1] ✉, Moein Naseri [2], Chandan Datta [3,4], Alexander Streltsov [2] & Michał Oszmaniec [5]

In this work, we consider biased-noise qubits affected only by bit-flip errors, which is motivated by existing systems of stabilized cat qubits. This property allows us to design a class of noisy Hadamard tests involving entangling and certain non-Clifford gates, which can be conducted reliably with only a polynomial overhead in algorithm repetitions. On the flip side, we also found classical algorithms able to efficiently simulate both the noisy and noiseless versions of our specific variants of the Hadamard test. We propose to use these algorithms as a benchmark of the biasness of the noise at the scale of large circuits. The bias being checked on a full computational task makes our benchmark sensitive to crosstalk or time-correlated errors, which are usually invisible from individual gate tomography. For realistic noise models, phase-flip will not be negligible, but in the Pauli-Twirling approximation, we show that our benchmark could check the correctness of circuits containing up to 10^6 gates, several orders of magnitude larger than circuits not exploiting a noise-bias. Our benchmark is applicable for an arbitrary noise-bias, beyond Pauli models.

Quantum computers bring the hope of solving useful problems for society that would be out of reach for classical supercomputers. One can think of problems in optimization[1,2], cryptography[3,4], finance[5,6], quantum chemistry or material sciences[7–11]. The main threats toward the realization of useful quantum computers are noise and decoherence[12], which cause errors and degrade the quality of the computation. In the long term, this problem will likely be addressed by quantum error correction and fault-tolerant quantum computing[13,13–18]. Yet, the very high fidelity and considerable overhead this approach requires make it very challenging to implement.

In this context, the existence of quantum algorithms able to scale up to a large size, with low-fidelity hardware, would be highly desirable. Unfortunately, various studies showed that polynomial-time classical algorithms can efficiently simulate the algorithm outputs of specific circuits in the noisy case, such as random circuits[19,20] or general algorithms under some assumptions on the noise structure[21–23] (see also ref. 24 and refs. 25,26 for analogous results in the optical setting). All these studies indicate that without doing error-correction, for realistic noise models, it is not possible to preserve reliable algorithm outputs in a classically intractable regime (see, however, recent work[27], which showed that in certain oracular scenarios noisy quantum computers can offer an advantage over classical computers). One major difficulty to face is that, for most noise models, the fidelity of the output state drops exponentially with the number of gates involved in the

computer[28], suggesting that a reliable estimation of any expectation value would require to run exponentially many times the algorithm, ruining any hope for an exponential speedup. Error mitigation techniques[29–36] have been proposed, with the hope to solve this issue. However, various no-go results show that, for several noise models, error mitigation techniques are not scalable[33,37]: the number of samples they require can grow exponentially with the algorithm's depth or the number of qubits in the algorithm[34,38,39]. Other approaches can be potentially more scalable, but they assume specific noise models[40], require knowledge of the entanglement spectrum of quantum states[41] or have potentially high algorithmic complexity[42].

In our work, motivated by the limitations of error-correction and error mitigation, we propose instances of the Hadamard test[43,44] that can be robustly implemented in systems of so-called biased-noise qubits[45–47], affected only by local stochastic bit flip (or phase-flip) errors. In technical terms, we show that, for suitably designed circuits tailored to a Pauli biased noise, the outcome of the *ideal* version of the Hadamard test can be estimated reliably by execution of its noisy versions with only a polynomial overhead in the number of repetitions of the algorithm. The model of biased-noise qubits is motivated by existing hardware realizing stabilized cat qubits[46–48]. The key ingredient of our approach is that only specific gates are allowed, in order to preserve the noise bias along the computation[45,49]. This allows us to design circuits in which the measurement will be isolated from

¹Centre for Quantum Optical Technologies, Centre of New Technologies, University of Warsaw, Warsaw, Poland. ²Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland. ³Institute for Theoretical Physics III, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ⁴Department of Physics, Indian Institute of Technology Jodhpur, Jodhpur, India. ⁵Center for Theoretical Physics, Polish Academy of Sciences, Warszawa, Poland. ✉e-mail: fellous.asiani.marco@gmail.com
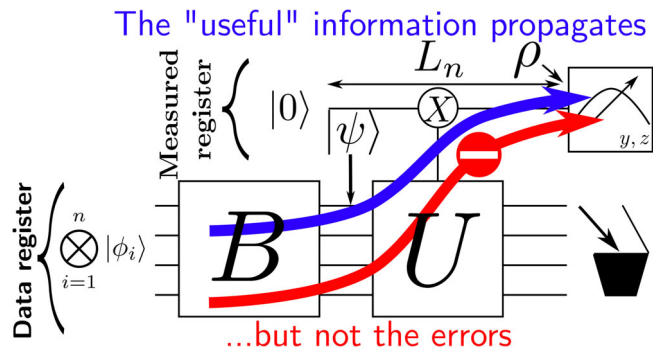
Fig. 1 | The Hadamard test. The Hadamard test represented on this figure allows to estimate $\langle\psi|U|\psi\rangle$ where $|\psi\rangle \equiv B\otimes_{i=1}^{n}|\phi_i\rangle$, ($|\phi_i\rangle$ being a single-qubit state, B a unitary). This estimation is being done by initializing a measured register in $|0\rangle$, and by implementing the unitary $c_X U$ (see the paragraph "Notations and terminology" in the section Results to understand our notations). Finally, measuring the measured register in Pauli Y and Z bases gives access to the imaginary and real part of $\langle\psi|U|\psi\rangle$. We call L_n the number of gates applied on the measured register (including potential noisy identity gates), once $c_X U$ has been decomposed on an experimentally feasible gateset. In this paper, we show that for a noise model only composed of bit-flips, and under some restrictions on the n-qubit unitaries U and B, the measurement will only be sensitive to bit-flips produced on the measured register, and entirely insensitive to the ones produced on the data register. Nonetheless, some `useful information" contained in the entangled state $|\psi\rangle$ will propagate toward the measurement. The intuitive principle we use is to encode the "useful" information in a "Pauli Z (or Y) channel" while making sure that any error is inside a "Pauli X channel". If any Pauli X error (i.e., bit-flip) from the data register cannot propagate toward the measurement while the Pauli Z can, the information will reach the measurements but not the errors. Overall, if $L_n = O(\log(n))$, the noise will only introduce a polynomial overhead in the number of algorithm repetitions to guarantee a reliable outcome: the algorithm will be scalable despite the noise. Note that while $L_n = O(\log(n))$, the algorithm can nonetheless have a polynomial depth.

most of the errors occurring (see Fig. 1). Importantly, the circuits themselves can be quite complicated: the allowed set of gates can generate large amounts of entanglement and contain certain non-Clifford gates, thus avoiding natural classical simulation techniques[50–52]. Nonetheless, the restricted nature of these circuits allowed us to find an efficient classical algorithm that simulates realizations of our family of Hadamard tests, also in the presence of arbitrary local biased noise. We propose to use this algorithm as a simple benchmark of the biasness of the noise at the scale of large and complicated quantum circuits. The interest of our benchmark is that it is scalable and can detect some collective effects of the noise that cannot, by definition, be observed at the level of individual gates. For instance, it can detect some crosstalk and correlated errors. The overall principle of the benchmark, summarized at the end of the results section and detailed in the methods section, is to compare the experiment to the classical simulation, which assumes the noise model of each gate used inside the complete algorithm is identical to the one deduced from individual gate tomography. Hence, if the simulation and experiment differ, it would imply that collective effects are degrading the quality of the hardware when a complete algorithm is implemented, indicating a potential threat to the scalability of the hardware. For pedagogy, in the main text, we mainly focus on the case of Pauli noise, but our benchmark is applicable for the most general model of local biased model as explained in the methods section.

Results

Notations and terminology

Let $(\sigma_0, \sigma_1, \sigma_2, \sigma_3) \equiv (I, X, Y, Z)$ denote single-qubit Pauli matrices. Let H denote the Hadamard gate. We call $\mathbb{P}_n^X$ the set of X-Pauli operators acting on n-qubits, $\mathbb{P}_n^X \equiv \{\otimes_{k=1}^{n}\sigma_{i_k}, |\forall k, i_k \in \{0,1\}\}$. We say that $f_n \in \text{poly}(n)$ if there exist two reals $C, a > 0$ such that $\lim_{n\to\infty} f_n/(Cn^a) = 1$. Moreover, when $\text{poly}(n)$ appears in an equation, it means that the equation remains true by replacing $\text{poly}(n)$ by any function $f_n \in \text{poly}(n)$. Lastly, we say that

$f_n = O(g_n)$ if there exists $C > 0$ such that $\lim_{n\to\infty} |f_n/g_n| \leq C$. For any unitary U, we define its coherently controlled operation in the X-basis as $c_X U \equiv |+\rangle\langle+| \otimes I + |-\rangle\langle-| \otimes U$. Let G be a single-qubit unitary. Unless stated otherwise, G_i indicates that G is applied on the i'th qubit in the tensor product (and I is applied elsewhere). Calling $\mathcal{G}$ the unitary map implementing a unitary quantum gate G, and $\mathcal{E}_\mathcal{G}$ the CPTP (Completely Positive Trace Preserving) operation describing the noisy implementation of the gate, we define the "noise map of G" (or the noise map associated with $\mathcal{G}$) the CPTP $\mathcal{N}_\mathcal{G}$ defined as $\mathcal{N}_\mathcal{G} \equiv \mathcal{E}_\mathcal{G} \circ \mathcal{G}^\dagger$. The noise map of a state preparation is the CPTP map that is applied after a noiseless state preparation, so that the overall process (noiseless preparation followed by CPTP map) describes the noisy state preparation. The noise map of a measurement is the CPTP map that must be applied before the noiseless measurement so that the overall process (CPTP map followed by noiseless measurement) describes the noisy measurement (noise maps for state preparation and noisy measurements are not uniquely defined). We call $\text{supp}(\mathcal{G})$ the set of qubits over which the CPTP $\mathcal{G}$ acts non-trivially. The notation $\sum_{\alpha \subset \text{supp}(\mathcal{G})}$ means that the summation is over every possible subset of $\text{supp}(\mathcal{G})$. For instance, if $\mathcal{G}$ is a two-qubit gate acting on qubits 1 and 2, we have $\text{supp}(\mathcal{G}) = \{1, 2\}$, and α will take every one the 4 values in the following set $\{\{\emptyset\}, \{1\}, \{2\}, \{1, 2\}\}$. These 4 values correspond to the 4 possible subsets of $\text{supp}(\mathcal{G})$. Finally, $I_{\text{supp}(\mathcal{G})}$ is the identity operator acting on the qubits where the CPTP map $\mathcal{G}$ acts non-trivially.

Introduction to the Hadamard test

The Hadamard test is the task over which all our examples are built. It allows for the estimation of the expectation value $\langle\psi|U|\psi\rangle$ of a unitary U on some prepared n-qubit state $|\psi\rangle = B\otimes_{i=1}^{n}|\phi_i\rangle$ ($|\phi_i\rangle$ being a single-qubit state, B a unitary operation), leading to several applications[44,53,54]. A way to implement it is represented in Fig. 1. It consists of a "measured" register initialized in $|0\rangle$ combination with a "data" register, where the state $|\psi\rangle$ has been prepared. The reduced state of the measured register right before measurement takes the following form

$$\rho = \frac{1}{2}\left(I + \alpha_n(yY + zZ)\right), \tag{1}$$

where $y = -\text{Im}(\langle\psi|U|\psi\rangle)$, $z = \text{Re}(\langle\psi|U|\psi\rangle)$ and $\alpha_n = 1$ from now on. Hence, measuring the first register in the Y (resp. Z) basis allows to estimate the imaginary (resp. real) part of $\langle\psi|U|\psi\rangle$. Using Hoeffding's inequality[33] we get that $N = 2\log(2/\delta)/\epsilon^2$ experimental repetitions are sufficient to estimate y or z up to ϵ-precision with a probability $1-\delta$. It is well-known that estimating $\langle\psi|U|\psi\rangle$, for a general polynomial circuit U, to additive precision is a BQP-complete problem[53] and therefore, in general, we do not expect an efficient classical algorithm that would realize this task.

So far, we discussed what happens when the measured qubit is noiseless. Let us now assume that a bit-flip channel right before the measurement occurs in such a way that $0 < \alpha_n < 1$ in Eq. (1). Then, repeating the algorithm $N_n = 2\log(2/\delta)/(\alpha_n\epsilon)^2$ times would be sufficient to estimate y and z to ϵ precision with high probability (the scaling in $1/(\alpha_n\epsilon)^2$ is also optimal[55]). If α_n decreases exponentially with n, the total number of algorithm calls will necessarily grow exponentially with n and the algorithm wouldn't be scalable. However, for $\alpha_n = 1/\text{poly}(n)$, only $\text{poly}(n)$ overhead in experiment repetitions would be sufficient to reliably estimate $\langle\psi|U|\psi\rangle$. In this work, we will show that for biased-noise qubits, it is possible to design a class of non-trivial Hadamard tests for which $\alpha_n = 1/\text{poly}(n)$ in the presence of non-vanishing local biased noise.

Noise model

In general, both B, and $c_X U$ have to be decomposed on a gateset implementable at the experimental level. Let $\mathcal{G}$ be a unitary channel describing a gate belonging to the accessible gateset, and $\mathcal{E}_\mathcal{G}$ be its noisy Completely Positive Trace Preserving (CPTP) implementation in the laboratory. We will assume a local biased noise model: $\mathcal{E}_\mathcal{G} = \mathcal{N}_\mathcal{G} \circ \mathcal{G}$, where the CPTP "noise map" $\mathcal{N}_\mathcal{G}$ will only introduce (possibly correlated) bit-flip errors on the set of

qubits over which $\mathcal{G}$ acts non-trivially. We call this set $\mathrm{supp}(\mathcal{G})$. We have:

$$\mathcal{N}_{\mathcal{G}}(\rho) = \sum_{\alpha \subset \mathrm{supp}(\mathcal{G})} p_\alpha^{\mathcal{G}} X_\alpha \rho X_\alpha, \tag{2}$$

where the summation is over every possible subsets α of $\mathrm{supp}(\mathcal{G})$ (we clarify what it means with an example in a few lines). We also have $X_\alpha \equiv \prod_{i \in \alpha} X_i$ for $\alpha \neq \{\emptyset\}$ and $X_{\{\emptyset\}} \equiv I_{\mathrm{supp}(\mathcal{G})}$. Here, $I_{\mathrm{supp}(\mathcal{G})}$ is the identity operator acting on the qubits where $\mathcal{G}$ acts non-trivially. Therefore, α denotes the set of qubits on which a bit-flip operator is applied in the summation in (2). Finally, $\{p_\alpha^{\mathcal{G}}\}$ is a probability distribution supported on the subsets of $\mathrm{supp}(\mathcal{G})$. For instance, if $\mathcal{G}$ is a two-qubit gate acting on qubits 1 and 2, we have $\mathrm{supp}(\mathcal{G}) = \{\{1\}, \{2\}\}$, and α will take each of the 4 values in the following set $\{\{\emptyset\}, \{1\}, \{2\}, \{1,2\}\}$. These 4 values correspond to the 4 possible subsets of $\mathrm{supp}(\mathcal{G})$. It means that $\mathcal{N}_{\mathcal{G}}$ will be a bit-flip noise model having the following 4 Kraus operators: $\sqrt{p_{\{\emptyset\}}} I_1 \otimes I_2$, $\sqrt{p_{\{1\}}} X_1 \otimes I_2$, $\sqrt{p_{\{2\}}} I_1 \otimes X_2$ and $\sqrt{p_{\{1,2\}}} X_1 \otimes X_2$. Noisy measurements are modeled by a perfect measurement followed by a probability p_{meas} to flip the outcome. Lastly, we assume that single-qubit noisy state preparation consists of a perfect state preparation followed by the application of a Pauli X-error with probability p_{prep}. Our noise model is based on an idealization of cat qubits that are able to exponentially suppress other noise channels than bit-flip, at the cost of a linear increase in bit-flip rate[46–48] (or the other way around). It is an idealization as (i) the Kraus operators are a linear combination of Pauli X and I operators only: this is what we call a perfect bias, (ii) we also neglect coherent errors, meaning that our noise model is a Pauli noise. The assumption (ii) is mainly used for pedagogical purposes in the main text. Indeed, our main result (the benchmarking protocol) is applicable for biased qubits which are designed in such a manner that their noise model is well approximated by the assumption (i) only (see Theorem 4 and Definition 4). We now provide the definition of "an error".

Definition 1. (Error). Let $|\Psi\rangle$ be the state the qubits should be in at some timestep of the algorithm if all the gates were perfect. Because we consider a probabilistic noise model, the actual n-qubit quantum state will take the form $\rho = \sum_i p_i E_i |\Psi\rangle\langle\Psi| E_i^\dagger$, where E_i is a unitary operator and p_i some probability ($p_i \geq 0$, $\sum_i p_i = 1$). The operator E_i is what we call the error that affected $|\Psi\rangle$.

Designing noise-resilient Hadamard test
The core idea behind our work is to exploit the fact that only bit-flips are produced, in order to design circuits guaranteeing that most of these errors will never reach the measurement in the Hadamard test. In order to explain our results we need to introduce first a number of auxiliary technical definitions.

Definition 2. (X-type unitary operators and errors). We call X-type unitary operators the set of unitaries that can be written as a linear combination of Pauli X matrices. For n-qubits, we formally define it as:

$$\mathbb{U}_n^X \equiv \left\{ U = \sum_i c_i P_i, |P_i \in \mathbb{P}_n^X, c_i \in \mathbb{C}, U^\dagger = U^{-1} \right\}. \tag{3}$$

Alternatively, $\mathbb{U}_n^X$ can be understood in terms of unitaries that are diagonal in the product basis $|\mathbf{s}\rangle = |s_1\rangle|s_2\rangle \ldots |s_n\rangle$, where $s_i = \pm$ and $|\pm\rangle$ are eigenstates of Pauli X matrix. We will call "X-error", an error belonging to $\mathbb{U}_n^X$ (if the error is additionally Pauli, we will call it "Pauli X error", or simply "bit-flip").

What we need to do is to guarantee that any error occurring at any step of the computation is an X-error. It is possible with the use of "bias-preserving" gates: such gates map any initial X-error to another X-error. It motivates the following definition and property (all our results are derived in the Supplemental material).

Definition 3. (Bias-preserving gates). Let G be an n-qubit unitary operator. We say that G preserves the X-errors (or X-bias), if it satisfies the following property

$$\forall P \in \mathbb{P}_n^X, \exists A \in \mathbb{U}_n^X \text{ such that } GP = AG. \tag{4}$$

We denote $\mathbb{B}_n$ the set of such gates.

Property 1. (Preservation of the bias). If a quantum circuit is only composed of gates in $\mathbb{B}_n$, each subject to a local biased noise model from Eq. (2) (and the paragraph that follows for measurement and preparation), then any error affecting the state of the computation is an X-error.

As examples of bias-preserving gates, there are all the unitaries in $\mathbb{U}_n^X$, the cNOT, and what we call Toffoli' $\equiv H_1 H_2 H_3 \times \text{Toffoli} \times (H_1 H_2 H_3)^\dagger$ (We assume Toffoli' can be implemented "natively" without having to actually apply the Hadamards (otherwise, because Hadamard would also be noisy, the full sequence wouldn't be bias-preserving)). An example of a gate that does not preserve the bias is the Hadamard gate. The existence of gate preserving the bias also *during* the gate implementation is not straightforward, but cat-qubits are able to overcome this complication, at least for cNOT, Toffoli' and gates in $\mathbb{U}_n^X$: see the section I C in the Supplemental Material. Finally, bias preserving gates have a nice interpretation: they correspond to permutations (up to a phase) in the Pauli X-eigenstates basis.

Property 2. (Characterization of bias-preserving gates). $V \in \mathbb{B}_n$ if and only if for any $\mathbf{s} \in \{+, -\}^n$, there exists a real phase $\varphi_{\mathbf{s},V}$ such that: $V|\mathbf{s}\rangle = e^{i\varphi_{\mathbf{s},V}} |\sigma_V(\mathbf{s})\rangle$ where σ_V is some permutation acting on $\{+, -\}^n$.

We now sketch the sufficient ingredients guaranteeing the existence of noise-resilient Hadamard tests. First, (i) assume that individual gate errors, as well as individual measurement and state-preparation errors, occur with a probability smaller than $p < 1/2$. Furthermore, assume (ii) that only X-errors occur in the algorithm (it can be satisfied with the assumptions of Property 1), (iii) these errors cannot propagate from the data to the measured register, (iv) the number of interactions of the measured register with the data register satisfies $L_n = O(\log(n))$ (which implies the measured register will only be impacted by X-errors introduced at $O(\log(n))$ locations). Subject to conditions (i–iv) the reduced state ρ will satisfy Eq. (1) with α_n efficiently computable classically, and satisfying $\alpha_n \geq 1/\mathrm{poly}(n)$. As explained earlier, this would guarantee a scalable algorithm to estimate $\langle \psi | U | \psi \rangle$. In particular, the following Theorem holds.

Theorem 1. (Hadamard test resilient to biased noise). Let:

$$|\psi\rangle = B \bigotimes_{i=1}^{N_B} |\phi_i\rangle, \quad U = W \cdot V,$$
$$W \equiv \prod_{i=1}^{N_W} W_i, \quad V \equiv \prod_{i=1}^{N_V} V_i, \tag{5}$$

where B is a product of local bias preserving gates, gates V_i and W_i are local gates and belong to $\mathbb{U}_n^X$. In the context of this theorem, the index i on V_i and W_i refers to the i'th gate in the product (it does not necessarily mean the gate is applied on the i'th qubit). Additionally, the gates W_i are assumed to be Hermitian. We assume the circuit is implemented as indicated on Fig. 2. There, W is implemented thanks to the "parallelization register", while V is implemented by making the measured and data register directly interact.

Furthermore, we assume the local bias noise model introduced in Eq. (2), and that state preparation, measurements, and each non-trivial gate applied on the measurement register have a probability at most $p < 1/2$ to introduce a bit-flip on the measured register.

Under these conditions, there exists a quantum circuit realizing a Hadamard test such that, in the presence of noise, the reduced state ρ satisfies Eq. (1) with $\alpha_n \geq (1 - 2p)^{O(N_V)}$.

Additionally, α_n is efficiently computable classically. Hence, if $N_V = O(\log(n))$, it is possible to implement the Hadamard test in such a
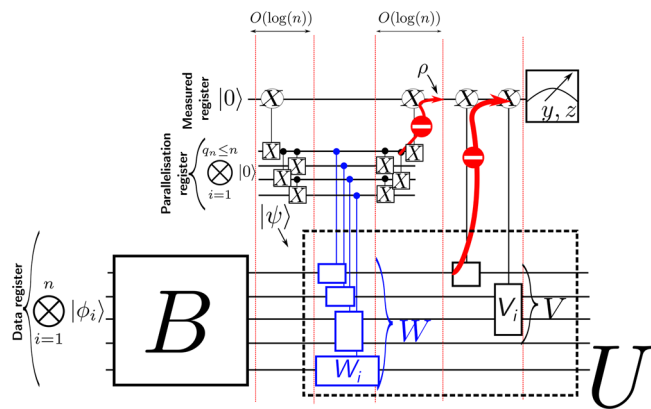
Fig. 2 | Concrete implementation of the noise-resilient Hadamard test. In this figure, we illustrate how $U = W \times V = \otimes_{i=1}^{N_W} W_i \times \otimes_{i=1}^{N_V} V_i$ mentioned in Theorem 1 can be implemented in a noise-resilient manner. The coherently controlled unitaries with a blue contour implement the controlled Hermitian unitaries W_i, which overall implement the controlled W. The coherently controlled unitaries with a black contour implement the controlled V_i, which overall implement the controlled V. The central element in this construction is the parallelization register[68], which allows the implementation of the unitary W in such a way that it acts upon all the qubits in the data register while preserving the noise-resilience. Hence, it allows the implementation of a Hadamard test with U acting on all the qubits of the data register. Such unitaries are useful for our benchmarking protocol: see the third paragraph in the discussion section. The key point behind the parallelization register is that, while it introduces additional components that can introduce bit-flip errors, these bit-flips cannot, by construction, propagate to the measured register (in practice, they commute with the last $c_X X$ gate drawn, (see the paragraph "Notations and terminology" in the section Results to understand our notations). It is an example where trading space (using more qubits) to gain time (guaranteeing that the measured register interacts with $O(\log(n))$ gates and not $\text{poly}(n)$ gates) is worth doing. While the parallelization register does not propagate bit-flips toward the measured register, it will propagate phase-flip errors. We make use of this to easily detect an excessive production of phase-flip errors in the third paragraph of the discussion section.

way that running the algorithm $\text{poly}(n)$ times is sufficient to estimate the real and imaginary parts of $\langle\psi|U|\psi\rangle$ to ϵ precision with high probability.

The Hadamard test is implemented with the circuit shown in Fig. 2. It uses a parallelization register. It is this parallelization register that allows for the implementation of a unitary acting on all the qubits of the data register (the unitary W mentioned in the theorem), while keeping a logarithmic number of interactions with the measured register (this is required to preserve noise-resilience). While Theorem 1 assumes the trivial identity gate to be noiseless, the results can be easily extended if they are noisy (in this case, W should have a depth in $O(\log(n))$, as discussed in the section III C of the Supplemental Material). Our approach is also scalable in the presence of noisy measurements (measuring the data register in the X basis to infer $\langle\psi|U|\psi\rangle$ would not be scalable in general in this case: see ref. 56 and the section IV B of the Supplemental Material).

Efficient simulation of noise-resilient Hadamard test
We just saw that restricted forms of Hadamard tests are resilient to bit-flip errors occurring throughout the circuit. The following result proves that the task realized by such a restricted Hadamard test can be efficiently simulated on a classical computer with polynomial effort.

Theorem 2. (Efficient classical simulation of restricted Hadamard test). Let $B \in \mathbb{B}_n$, $U \in \mathbb{U}_n^X$ be n qubit unitaries specified by R_B and R_U local qubit gates (belonging to respective classes $\mathbb{B}_n$ and $\mathbb{U}_n^X$). Let $|\psi_0\rangle = |\phi_1\rangle|\phi_2\rangle\cdots|\phi_n\rangle$ be an initial product state. Then, there exists a randomized classical algorithm, taking as input classical specifications of circuits defining B, U, and the initial state $|\psi_0\rangle$, that efficiently and with a high probability computes an additive approximation to $\langle\psi_0|B^\dagger UB|\psi_0\rangle$. We call $\mathcal{C}$ this

approximation. Specifically, we have

$$\Pr\big(|\langle\psi_0|B^\dagger UB|\psi_0\rangle - \mathcal{C}| \leq \epsilon\big) \geq 1 - \delta, \qquad (6)$$

while the running time is $T = O(\frac{R_B + R_U + n}{\epsilon^2}\log(1/\delta))$.

The classical simulability of the restricted Hadamard test can be regarded as the limitation of our approach to constructing quantum circuits that are robust to biased noise. At the same time, it allows us to introduce an efficient and scalable benchmarking protocol that is tailored to validating the assumption of bias noise on the level of the whole (possibly complicated) circuit. It allows us to validate this assumption in a manner that individual gate tomography couldn't, as we now explain.

Sketch of the benchmarking protocol
For pedagogy, in this section, we give an intuitive sketch of how the benchmarking would work for the Pauli bit-flip noise model considered so far. In the method section, we give the details behind the benchmarking protocol, which applies to the most general case of a biased noise model for qubits, beyond bit-flip (formally, a noise model satisfying Definition 4). In the methods section, we also provide the classical algorithms able to efficiently simulate the *noisy* implementation of the quantum circuit.

The issue with an individual benchmark of quantum gates is that it is blind to collective effects that can only build up when a full circuit, composed of many gates in sequence and in parallel, is implemented. One example is the presence of correlated errors, or more generally, non-local effects in the noise, that individual gate tomography cannot usually detect. Another example of collective effects is the presence of "scale-dependent noise", i.e., a noise for which the intensity depends on the number of qubits or surrounding gates used in the algorithm[57–59]. Note that scale-dependent noise can sometimes be implied by correlated noise models[60]. Scale-dependent noise can be particularly damaging for biased qubits, as it could rule out the intrinsic interest of using such qubits in the case that the nearly suppressed error rate happens to grow when the circuits are scaled up. All these behaviors for the noise represent a major threat both in the near term, and to reach the large-scale[57,61,62]. Our benchmarking protocol, which we now sketch, allows us to detect some of these effects. The intuitive idea is that we are able to classically predict the algorithm's output under the assumption that the noise deduced from individual gate tomography is still realized for each gate, but now inside a whole algorithm. A mismatch between the classical algorithm and the quantum experiment would necessarily indicate a violation of the assumption behind the noise.

We emphasize here that while there exist scalable protocols able to benchmark the noise strength at the scale of entire circuits (for instance, randomized benchmarking[63]), to our knowledge, there doesn't exist any scalable protocol able to see whether or not the noise model is *biased*, at the scale of large circuits. Our benchmark is precisely able to do this, which is something crucial for cat qubit platforms. This is the main reason why our protocol differs from the existing approaches. Additionally, we are not aware of other classical simulation methods than the one we developed in Theorems 2, 3, for the class of circuits affected by biased noise considered by us. Specifically, none of the existing classical simulation techniques can efficiently simulate geometrically nonlocal circuits covered by our work, while such simulations are at the core of our benchmark.

For simplicity of the explanations, in all that follows we propose to implement our benchmark in the exact circuits we analyzed so far (i.e. the ones of Fig. 2). However, as mentioned at the end of the third paragraph of the discussion section, depending on which violation in the noise we aim to be sensitive to, simpler circuits could potentially be considered.

For a Pauli bit-flip noise, we first need to perform tomography on the gates that need to be applied on the measured register, in order to find the probability that each such gate introduces a bit-flip on the measured register. Other noise parameters p_{meas} and p_{prep} would also need to be estimated. Then, a Hadamard test satisfying the constraints of Theorem 1 is chosen. Subsequently, one needs to classically compute α_n, and estimate y, z up to ϵ-precision, with high probability (using Theorem 2). The estimates of the

noiselessly, z are compared to values of $Tr(\rho Y)/\alpha_n$, $Tr(\rho Z)/\alpha_n$ estimated experimentally by running the quantum circuit polynomially many times. The experimental and theoretical predictions are then compared. If they don't match (up to error resulting from a finite number of experiments), it would necessarily indicate a violation of the assumptions behind the assumed noise model (bit-flip) for the circuit participating in the Hadamard test. The precise implementation of the benchmarking protocol is provided in the methods section.

Discussion

In supplemental V B, we quantify the total number of quantum gates allowed in the circuit for which our benchmark could be practically useful. Based on available data in the literature, we show that, under the Pauli-Twirling approximation, our benchmark could be used for circuits containing 10^6 gates. As it represents circuits three to four orders of magnitude bigger than in current experiments[64], it is a strong indication of the practical usefulness of our benchmark for the NISQ regime, and beyond, allowing us to check the hardware reliability for large circuits. Loosely speaking, our benchmark shall be used in the regime where the suppressed source of errors (phase-flip in our convention) is expected to be negligible. This is because it relies on the classical simulation, Theorem 3, which assumes a perfect bias, hence that no phase-flip are produced in the algorithm. This is precisely the regime where having a benchmark is interesting as there are propositions to use biased-qubits in large-scale algorithms where only the dominant source of errors is corrected, because the suppressed errors would have a negligible impact at the level of the algorithm (see supplemental E2 of ref. 65 and references therein). Hence, our protocol is well-suited to analyze this regime of interest.

Our protocol can first detect some correlated errors or, more generally, non-local effects of the noise that are usually invisible from individual gate tomography. This is because our classical simulation algorithm assumes that individual gate tomography provides a fair description of the behavior of the noise, as it takes as inputs the noise maps of the gates extracted from tomography. Thus, if gate tomography does not fairly describe the noise, the results from simulation and experiment would disagree. To be concrete: if the noise model satisfies the general definition of perfect bias, beyond bit-flips, i.e., Definition 4, every gate applied on the measured register introduces noise that can impact the measurement outcomes. Furthermore, these are the only gates for which the noise can impact the measurement outcomes. This is because, with a perfect bias, no error produced outside of the measured register can propagate to the measured register (this is a consequence of how our circuits are built). Hence, we can first detect the presence of some correlated errors between the gates applied on the measured register, as they would violate the assumption that each gate can be described with a noise map independent from the others: this is a first example of a violation we can detect. We can also detect some violations of the locality assumption of the noise. If there are non-local effects in the noise, the noise maps of the gates applied on the measured register would, in general, not correctly describe the experimental outcomes. For instance, it could be that some operations done locally in the parallelization or data registers introduce unexpected errors in the measured one (because of crosstalk, for example). Our benchmarking will, in general, allow us to detect such effects. These are just some examples to help the reader understand what the benchmark can be useful for. Yet, we think that the most useful application of our protocol is its ability to detect the occurence of phase-flip errors produced at a higher rate than expected (we recall that having error-rates growing with the computer's size is an effect experimentally observed in superconducting qubits[57–59]). Indeed, such effects would in general lead to a mismatch between the classical simulation and the experiment for algorithms containing less gates than expected, i.e, a total number of gates, N, smaller than what (9) predicts ((9) quantifies the circuit's size for which our benchmarking is applicable). To be concrete, an experimentalist should choose the largest N so that the bound (9) is saturated. If $\Delta \geq 0$ (Δ quantifies how much the expected noise model is violated experimentally: it is formally defined in the paragraph "Principle of the benchmarking" in Theorem 4),

there is a violation of the assumption behind the noise model, indicating a potential threat to the scalability of the platform. For instance, for $\epsilon = 1/50$, our quantitatives estimates of supplemental V B indicate that $N = 10^6$ would work.

We can also provide a concrete example of a circuit that can be used to efficiently detect the production of phase-flip at a higher rate than expected. It would be the one implementing the controlled unitary $c_X U$, with $U = \otimes_{i=1}^{n} X_i$. This circuit matches the constraints of Theorem 1 (including its generalization to noisy identity gates mentioned in the paragraph that follows Theorem 1): we can thus implement it for the benchmarking. The reason why this circuit is useful is that it would, in general, make the measurement outcome sensitive to the introduction of phase-flip error on any of the qubits from the data register, after the preparation unitary B. This is because a phase-flip error occuring on any of the qubits from the data register would first propagate to the parallelization register (through the blue cNOTs implementing $W = U = \otimes_{i=1}^{n} X_i$ in Fig. 2), and finally to the measured register. Hence, phase-flip errors produced in B will, in general, modify the measurement outcome probability distribution, allowing to efficiently detect an excessive production of such errors. An experimentalist could then see if the circuit composed of bias-preserving gates of interest, encoded in the unitary B, does indeed produce more phase-flip errors than it should. Detecting such events is crucial for superconducting cat qubits given the fact their whole scalability strategy precisely relies on keeping negligible the phase-flip error rates also when used in large-scale circuits[65]. Note that if the goal is only to detect the production of phase-flips at a higher rate than expected, and not to see if other non-local effects are occurring (we can detect some of these other non-local effects with our benchmarking, see the previous paragraph), simpler circuits than the ones based on Fig. 2 could very likely be used for benchmarking. Assuming noiseless measurements, one could for instance measure the observable $X^{\otimes n}$ on the prepared $|\psi\rangle$ by directly measuring all the data qubits in the Pauli-X basis (i.e., in this case, one would remove the parallelization and measured register of Fig. 2). In presence of a perfect bias, the errors will commute with the measurements. Hence, the implementation of the circuit would give the same outcome as the classical simulation algorithm Theorem 2 (for $U = X^{\otimes n}$, $B|\psi_0\rangle = |\psi\rangle$). For this reason, we believe that a similar benchmarking protocol as Theorem 4 could be derived for this simpler circuit (a mismatch between the simulation and experiment would indicate a violation). However, we leave a rigorous proof of this guess for a future work (the proof should in particular acknowledge what happens if these measurements are noisy—something we acknowledged in Theorem 4 (See the paragraph preceding Theorem 3 and the noise of Definition 4)). This noise model makes Pauli-Y and Pauli-Z measurements noisy, which are the only ones we use in our circuit of Fig. 2)—and precisely quantify the violation with similar equations as (9) and (10).

Overall, the important aspect of our benchmark is that it is scalable (even if measurements and state preparation are noisy), and allows to detect violation of the noise model in a manner that would not be visible from individual gate tomography, because some effects of the noise cannot be detected at this level, as we discussed at the beginning of section sketching the benchmarking protocol.

Finally, while our circuits are efficiently simulable, their strong noise-resilience, rigorously proven in presence of a Pauli bit-flip noise (i.e. the noise model described around (2), which is more restrictive than the one of Definition 4 for which the benchmark is applicable), makes natural to wonder if extensions of our work could lead to noise-resilient circuits also showing a computational interest: this work also shows what we analyzed on this question of fundamental interest. For this, we first notice that the set of bias-preserving gates, $\mathbb{B}_n$ can be composed of $c_X X$ gates, which combined to initial states $|0/1\rangle$ can generate arbitrary graph states (in the local X basis). It is known that a typical n-qubit stabilizer state exhibits strong multipartite entanglement[66]. This, together with the fact that arbitrary graph states are locally equivalent to stabilizer states[67], implies that, in general bias-preserving circuits can generate a rich family of highly entangled states (these graphs states are however not computationally useful for the specific

task adressed in our paper, see the section II B of the Supplemental Material). We can also implement certain non-Clifford gates, and for our task, $U \in \mathbb{U}_n^X$ can act non-trivially over all the qubits of the data register, there is no restriction on the number of gates in the preparation unitary B, and the circuit is scalable despite noisy measurements. The error rates of the gates could also grow with n, making our circuits noise-resilient against one of the main threats to the scalability[60] (see section IV A of the Supplemental Material).

Methods

Benchmarking protocol and simulation of the noisy circuits

In the last part of the results section, we gave a sketch of the benchmarking protocol for Pauli bit-flip biased noise. Here, we give the technical details behind the protocol while extending it beyond Pauli bit-flip noise. We also give the classical simulation algorithm, able to simulate the noisy circuit for this more general noise model. In general, the noise model describing biased qubits can have its Kraus operators describing the noise map of each gate, state preparation and measurements that do not exactly correspond to a Pauli bit-flip channel. We mean that the Kraus operators could be a linear combination of Pauli bit-flips (i.e., they would follow Definition 4: the bit-flip channel treated so far was a particular case of the expected noise model for such hardware). Additionally, there can be imperfections in the bias. It means that the Kraus operators describing the noise maps could have a non-zero Hilbert-Schmidt inner product with Pauli operators not belonging to $\mathbb{P}_n^X$. For instance, the Kraus operators could have a non-zero overlap with a multi-qubit Pauli P containing at least one Pauli Z, or one Pauli Y in the expression of its tensor product (such as $P = X \otimes Z$ for instance). If it occurs, we will say that the noise model also produces phase-flip errors. Our benchmarking protocol can detect violations of the assumed noise model in this more general case. The protocol is based on the fact that we can simulate the outcome of the Hadamard test in the *noisy* case, under the assumption that the bias is perfect but not necessarily Pauli (i.e., that it satisfies Definition 4). This classical simulation takes as input the noise model of each individual gate, approximated to that of a perfect bias. A deviation of this classical simulation from the experimental outcomes further than an error budget will imply that the noise model prescribed by individual gate tomography is not occurring experimentally, identifying a possible threat to the hardware scalability, due to collective effects occurring at the scale of the algorithm. This error budget is related to the quality of the approximation of the noise model taken from individual gate tomography with a perfect bias. The formal protocol is written in Theorem 4. We now state the definition and theorems we need.

Definition 4. (Perfect bias). Let $\mathcal{G}$ be a quantum channel describing either a noiseless unitary gate belonging to the accessible gateset, or a single-qubit state preparation, and $\mathcal{E}_\mathcal{G}$ be its noisy implementation in the laboratory. We say that the noisy implementation of the gate (or state preparation), $\mathcal{E}_\mathcal{G}$, follows a perfectly biased noise model if $\mathcal{E}_\mathcal{G} = \mathcal{N}_\mathcal{G} \circ \mathcal{G}$, where the noise map $\mathcal{N}_\mathcal{G}$ is a CPTP (Completely Positive Trace Preserving) map that admits the following Kraus decomposition:

$$\mathcal{N}_\mathcal{G}(\rho) = \sum_j K_j^\mathcal{G} \rho \left(K_j^\mathcal{G}\right)^\dagger,$$
$$K_j^\mathcal{G} = \sum_{\alpha \subset \text{supp}(\mathcal{G})} c_\alpha^j X_\alpha, \tag{7}$$

In (7), we follow the notations introduced around (2). We also have: $\forall j, c_\alpha^j \in \mathbb{C}$ and $\sum_j (K_j^\mathcal{G})^\dagger K_j^\mathcal{G} = I_{\text{supp}(\mathcal{G})}$. A quantum map satisfying (7) will be said to be perfectly biased.

A noisy single-qubit measurement will be modeled as a perfect measurement, preceded by the application of a perfectly biased single-qubit CPTP map on the qubit being measured. This CPTP channel, called the noise map of the measurement, represents the noise brought by the measurement.

To make the formulation of the following theorems simple, we will assume that state preparation and measurements are noiseless. In the case all gates are bias-preserving (which is what we consider in all this paper) and if the noise model satisfies Definition 4, this assumption doesn't remove generality from our results. This is because the noise of the state preparation or measurement can be acknowledged by redefining the noise map of the following or preceding gate. The newly obtained noise map will still satisfy Definition 4.

Theorem 3. (Efficient classical simulation of a noisy Hadamard test under perfect bias). Let $B \in \mathbb{B}_n$, $U = W.V \in \mathbb{U}_n^X$, N_W, N_V satisfying the constraints of Theorem 1. Let $|\psi_0\rangle = |\phi_1\rangle|\phi_2\rangle \cdots |\phi_n\rangle$ be the initial product state for the data register. If $N_V = O(\log(n))$, if the noise model of each gate is perfectly biased according to Definition 4, if the total number of gates in the algorithm is in Poly(n), and if state preparation and measurements are noiseless (this is without loss of generality, see the comments before Theorem 3), there exists a randomized classical algorithm taking as input (I) classical specifications of the circuit implementing the Hadamard test for the chosen $(B, U, |\psi_0\rangle)$, (II) the quantum channel describing the noise model of each gate used in the computation, (III) the initial state $|\psi_0\rangle$ such that this algorithm efficiently and with a high probability computes an additive approximation to $Tr(P_1 \rho_X)$, where ρ_X is the reduced density matrix of the measured register at the end of the noisy implementation of the algorithm, and P_1 is a single-qubit Pauli matrix. We call $\mathcal{C}$ this classical approximation.

Specifically, we have:

$$\Pr\left(|Tr(P_1 \rho_X) - \mathcal{C}| \le \epsilon\right) \ge 1 - \delta, \tag{8}$$

while the running time is $T = O((1/\epsilon^2) \log(1/\delta) \times \text{poly}(n))$.

Theorem 4. (Benchmarking protocol). Consider a Hadamard test satisfying the constraints of Theorem 1. The circuit is implemented with N unitary gates represented by the unitary quantum channels $\{\mathcal{G}_i\}_{i=1}^N$. We call $\mathcal{N}_{\mathcal{G}_i}$ the noise map associated with $\mathcal{G}_i$ (extracted from individual gate tomography) and we assume state preparation and measurements to be noiseless (this is without loss of generality up to a redefinition of the noise maps of the quantum gates, see comments before Theorem 3).

Let $\mathcal{N}_{X,\mathcal{G}_i}$ be an approximation to $\mathcal{N}_{\mathcal{G}_i}$, such that $\mathcal{N}_{X,\mathcal{G}_i}$ has a noise model satisfying Definition 4.

Let ρ be the density matrix of the measured register at the end of the algorithm if the noise map of each $\mathcal{G}_i$ was $\mathcal{N}_{\mathcal{G}_i}$. Let ρ_X be the reduced density matrix of the measured register at the end of the algorithm if the noise map of each $\mathcal{G}_i$ was $\mathcal{N}_{X,\mathcal{G}_i}$. Let $\rho_{\exp}$ be the reduced density matrix of the measured register that exactly predicts the experimental outcomes. More precisely, we mean that implementing the measurements used in the Hadamard test on $\rho_{\exp}$ would exactly reproduce the measurement outcomes experimentally observed.

Assume that there exist $\epsilon > 0$ such that

$$\max_i ||\mathcal{N}_{X,\mathcal{G}_i} - \mathcal{N}_{\mathcal{G}_i}|| \le \frac{\epsilon}{\sqrt{2N}} \tag{9}$$

is satisfied. Then, for any single-qubit Pauli P_1:

$$|Tr(\rho P_1) - Tr(\rho_X P_1)| \le \epsilon. \tag{10}$$

Principle of the benchmarking

The benchmarking protocol works as follow. $Tr(\rho_X P_1)$ corresponds to the outcome of the circuit with a noise model for each gate satisfying Definition 4. Hence, it can be classically estimated with Theorem 3. $Tr(\rho_{\exp} P_1)$ is accessed experimentally. If $|Tr(\rho_{\exp} P_1) - Tr(\rho_X P_1)| > \epsilon$, $\rho_{\exp}$ and ρ necessarily differ, indicating that the noise model predicted by individual gate tomography (the noise maps $\{\mathcal{N}_{\mathcal{G}_i}\}_{i=1}^N$) is not occuring experimentally. It

thus indicates the presence of collective effects in the noise, possibly threatening the scalability of the biased-noise qubits.

Noteworthy, $\Delta \equiv |Tr(\rho_{exp}P_1) - Tr(\rho_X P_1)| - \epsilon$ can quantify how strong the noise violation is at the scale of the whole circuit (if $\Delta > 0$). It is a consequence of the fact $|Tr(\rho_{exp}P_1) - Tr(\rho P_1)| \geq \Delta$ (the larger Δ, the larger the violation).

Data availability
No data have been generated for this publication.

Code availability
No code has been generated for this publication.

References
1. Farhi, E., Goldstone, J. & Gutmann, S. A quantum approximate optimization algorithm. *arXiv* https://arxiv.org/abs/1411.4028 (2014).
2. Amaro, D. et al. Filtering variational quantum algorithms for combinatorial optimization. *Quantum Sci. Technol.* **7**, 015021 (2022).
3. Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.* https://epubs.siam.org/doi/10.1137/S0097539795293172 (2006).
4. Häner, T., Jaques, S., Naehrig, M., Roetteler, M. & Soeken, M. Improved quantum circuits for elliptic curve discrete logarithms. In *Post-Quantum Cryptography*, 425–444 (Springer, Cham, Switzerland, 2020).
5. Chakrabarti, S. et al. A threshold for quantum advantage in derivative pricing. *Quantum* **5**, 463 (2021).
6. Rebentrost, P. & Lloyd, S. Quantum computational finance: quantum algorithm for portfolio optimization. *Künstliche Intelligenz* **38**, 327–338 (2018).
7. Bauer, B., Bravyi, S., Motta, M. & Chan, G. K.-L. Quantum algorithms for quantum chemistry and quantum materials science. *Chem. Rev.* **120**, 12685–12717 (2020).
8. Cao, Y. et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).
9. Ma, H., Govoni, M. & Galli, G. Quantum simulations of materials on near-term quantum computers. *npj Comput. Mater.* **6**, 1–8 (2020).
10. Cao, Y., Romero, J. & Aspuru-Guzik, A. Potential of quantum computing for drug discovery. *IBM J. Res. Dev.* **62**, 6:1–6:20 (2018).
11. Zinner, M. et al. Quantum computing's potential for drug discovery: early stage industry dynamics. *Drug Discov. Today* **26**, 1680–1688 (2021).
12. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
13. Campbell, E. T., Terhal, B. M. & Vuillot, C. Roads towards fault-tolerant universal quantum computation. *Nature* **549**, 172–179 (2017).
14. Gottesman, D. An introduction to quantum error correction and fault-tolerant quantum computation. arXiv (2009). 0904.2557.
15. Grassl, M. & Rötteler, M. Quantum error correction and fault tolerant quantum computing. In *Encyclopedia of Complexity and Systems Science*, 7324–7342 (Springer, New York, NY, New York, NY, USA, 2009).
16. Terhal, B. M. Quantum error correction for quantum memories. *Rev. Mod. Phys.* **87**, 307–346 (2015).
17. Aliferis, P., Gottesman, D. & Preskill, J. Quantum accuracy threshold for concatenated distance-3 codes. *arXiv* https://arxiv.org/abs/quant-ph/0504218 (2005).
18. Fowler, A. G., Mariantoni, M., Martinis, J. M. & Cleland, A. N. Surface codes: towards practical large-scale quantum computation. *Phys. Rev. A* **86**, 032324 (2012).
19. Aharonov, D., Gao, X., Landau, Z., Liu, Y. & Vazirani, U. A polynomial-time classical algorithm for noisy random circuit sampling. *arXiv* https://arxiv.org/abs/2211.03999 (2022).
20. Gao, X. & Duan, L. Efficient classical simulation of noisy quantum computation. *Phys. Rev. Lett.* **126**, 210502 (2018).
21. Aharonov, D., Ben-Or, M., Impagliazzo, R. & Nisan, N. Limitations of noisy reversible computation. *arXiv* https://arxiv.org/abs/quant-ph/9611028 (1996).
22. De Palma, G., Marvian, M., Rouzé, C. & França, D. S. Limitations of variational quantum algorithms: a quantum optimal transport approach. *PRX Quantum* **4**, 010309 (2022).
23. Stilck França, D. & García-Patrón, R. Limitations of optimization algorithms on noisy quantum devices. *Nat. Phys.* **17**, 1221–1227 (2021).
24. Oh, C., Jiang, L. & Fefferman, B. On classical simulation algorithms for noisy boson sampling. *arXiv* https://arxiv.org/abs/2301.11532 (2023).
25. Brod, D. J. & Oszmaniec, M. Classical simulation of linear optics subject to nonuniform losses. *Quantum* **4**, 267 (2020).
26. García-Patrón, R., Renema, J. J. & Shchesnovich, V. Simulating boson sampling in lossy architectures. *Quantum* **3**, 169 (2019).
27. Chen, S., Cotler, J., Huang, H.-Y. & Li, J. The complexity of NISQ. *Nat. Commun.* **14**, 6001 (2022).
28. Zhou, Y., Stoudenmire, E. M. & Waintal, X. What limits the simulation of quantum computers? *Phys. Rev. X* **10**, 041038 (2020).
29. Temme, K., Bravyi, S. & Gambetta, J. M. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.* **119**, 180509 (2017).
30. Bonet-Monroig, X., Sagastizabal, R., Singh, M. & O'Brien, T. E. Low-cost error mitigation by symmetry verification. *Phys. Rev. A* **98**, 062339 (2018).
31. 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* **4**, 257 (2020).
32. Strikis, A., Qin, D., Chen, Y., Benjamin, S. C. & Li, Y. Learning-based quantum error mitigation. *PRX Quantum* **2**, 040330 (2021).
33. Cai, Z. et al. Quantum error mitigation. *arXiv* https://arxiv.org/abs/2210.00921 (2022).
34. Berg, E. v. d., Minev, Z. K., Kandala, A. & Temme, K. Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors. *arXiv* https://arxiv.org/abs/2201.09866 (2022).
35. Qin, D., Chen, Y. & Li, Y. Error statistics and scalability of quantum error mitigation formulas. *arXiv* https://arxiv.org/abs/2112.06255 (2021).
36. Koczor, B. Exponential error suppression for near-term quantum devices. *Phys. Rev. X* **11**, 031057 (2021).
37. Quek, Y., França, D. S., Khatri, S., Meyer, J. J. & Eisert, J. Exponentially tighter bounds on limitations of quantum error mitigation. *Nat. Phys.* **20**, 1648–1658 (2022).
38. Takagi, R., Tajima, H. & Gu, M. Universal sampling lower bounds for quantum error mitigation. arXiv https://arxiv.org/abs/2208.09178 (2022).
39. Takagi, R., Endo, S., Minagawa, S. & Gu, M. Fundamental limits of quantum error mitigation. *npj Quantum Inf.* **8**, 1–11 (2022).
40. Tan, A. K., Liu, Y., Tran, M. C. & Chuang, I. L. Error correction of quantum algorithms: arbitrarily accurate recovery of noisy quantum signal processing. *arXiv* https://inspirehep.net/literature/2625270 (2023).
41. Eddins, A. et al. Doubling the size of quantum simulators by entanglement forging. *PRX Quantum* **3**, 010309 (2022).
42. Singal, T., Maciejewski, F. B. & Oszmaniec, M. Implementation of quantum measurements using classical resources and only a single ancillary qubit. *npj Quantum Inf.* **8**, 82 (2022).
43. Smith, J. & Mosca, M. Algorithms for quantum computers. In *Handbook of Natural Computing*, 1451–1492 (Springer, Berlin, Germany, 2012).
44. Bharti, K. et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **94**, 015004 (2022).

45. Puri, S. et al. Bias-preserving gates with stabilized cat qubits. *Sci. Adv.* **6**, eaay5901 (2020).
46. Guillaud, J. & Mirrahimi, M. Error rates and resource overheads of repetition cat qubits. *Phys. Rev. A* **103**, 042413 (2021).
47. Chamberland, C. et al. Building a fault-tolerant quantum computer using concatenated cat codes. *PRX Quantum* **3**, 010329 (2022).
48. Lescanne, R. et al. Exponential suppression of bit-flips in a qubit encoded in an oscillator. *Nat. Phys.* **16**, 509–513 (2020).
49. Xu, Q., Iverson, J. K., Brandão, F. G. S. L. & Jiang, L. Engineering fast bias-preserving gates on stabilized cat qubits. *Phys. Rev. Res.* **4**, 013082 (2022).
50. Vidal, G. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.* **91**, 147902 (2003).
51. Aaronson, S. & Gottesman, D. Improved simulation of stabilizer circuits. *Phys. Rev. A* **70**, 052328 (2004).
52. Jozsa, R. & Linden, N. On the role of entanglement in quantum-computational speed-up. *Proc. R. Soc. Lond. A.* **459**, 2011–2032 (2003).
53. Aharonov, D., Jones, V. & Landau, Z. A polynomial quantum algorithm for approximating the Jones polynomial. *Algorithmica* **55**, 395–421 (2009).
54. Gokhale, P. et al. Quantum fan-out: circuit optimizations and technology modeling. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 276–290 (IEEE, 2021).
55. Jasper, L. & Samuel, S. 'distinguishing (discrete) distributions' (2020). https://cs.brown.edu/courses/csci1951-w/lec/lec%2011%20notes.pdf. (2023)
56. Polla, S., Anselmetti, G.-L. R. & O'Brien, T. E. Optimizing the information extracted by a single qubit measurement. *arXiv* https://arxiv.org/abs/2207.09479 (2022).
57. Zhao, P. et al. Quantum crosstalk analysis for simultaneous gate operations on superconducting qubits. *PRX Quantum* **3**, 020301 (2022).
58. Krinner, S. et al. Benchmarking coherent errors in controlled-phase gates due to spectator qubits. *Phys. Rev. Appl.* **14**, 024042 (2020).
59. Sevilla, J. & Riedel, C. J. Forecasting timelines of quantum computing. *arXiv* preprint arXiv:2009.05045 (2020).
60. Fellous-Asiani, M., Chai, J. H., Whitney, R. S., Auffèves, A. & Ng, H. K. Limitations in quantum computing from resource constraints. *PRX Quantum* **2**, 040335 (2021).
61. Murali, P., McKay, D. C., Martonosi, M. & Javadi-Abhari, A. Software mitigation of crosstalk on noisy intermediate-scale quantum computers. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, 1001–1016 (2020).
62. Agarwal, A., Lindoy, L. P., Lall, D., Jamet, F. & Rungger, I. Modelling non-markovian noise in driven superconducting qubits. *Quant. Sci. Technol.* **9**, 035017 (2024).
63. Knill, E. et al. Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
64. Pelofske, E., Bärtschi, A. & Eidenbenz, S. Quantum volume in practice: what users can expect from NISQ devices. *IEEE Trans. Quantum Eng.* **3**, 1–19 (2022).
65. Gouzien, É., Ruiz, D., Le Régent, F.-M., Guillaud, J. & Sangouard, N. Performance analysis of a repetition cat code architecture: Computing 256-bit elliptic curve logarithm in 9 hours with 126 133 cat qubits. *Phys. Rev. Lett.* **131**, 040602 (2023).
66. Smith, G. & Leung, D. Typical entanglement of stabilizer states. *Phys. Rev. A* **74**, 062314 (2006).
67. Van den Nest, M., Dehaene, J. & De Moor, B. Graphical description of the action of local Clifford transformations on graph states. *Phys. Rev. A* **69**, 022316 (2004).
68. Moore, C. & Nilsson, M. Parallel quantum computation and quantum codes. *SIAM J. Comput.* **31**, 799–815 (2001).

Acknowledgements

Author contributions

M.F.A. formulated the problem behind the paper and proposed the first version of noise-resilient circuits. M.F.A. and M.N. made calculations to understand how noise propagates through the various gates used in the circuit to generalize the circuits. M.O. derived the classical simulation algorithm, which was then generalized by MN and MFA to also simulate the noisy dynamics. M.F.A., M.N., and M.O. formulated the benchmarking protocol. A.S. and C.D. analyzed entanglement properties that can be produced in the circuit. All authors participated in writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41534-025-01054-5.

Correspondence and requests for materials should be addressed to Marco Fellous-Asiani.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.