arXiv:2301.04173v3 [quant-ph] 20 Sep 2023

A novel approach to noisy gates for simulating quantum computers

Giovanni Di Bartolomeo [◇,1,2,*] Michele Vischi [◇,1,2,†] Francesco Cesa,[1,2,‡]
Roman Wixinger [3] Michele Grossi [4,§] Sandro Donadi [2] and Angelo Bassi [1,2]

[1] *Department of Physics, University of Trieste, Strada Costiera 11, 34151 Trieste, Italy*
[2] *Istituto Nazionale di Fisica Nucleare, Trieste Section, Via Valerio 2, 34127 Trieste, Italy*
[3] *Institute of Particle Physics and Astrophysics, ETH Zürich, Zürich, Switzerland*
[4] *European Organization for Nuclear Research (CERN), Geneva 1211, Switzerland*

We present a novel method for simulating the noisy behaviour of quantum computers, which allows to efficiently incorporate environmental effects in the driven evolution implementing the gates acting on the qubits. We show how to modify the noiseless gate executed by the computer to include any Markovian noise, hence resulting in what we will call a noisy gate. We compare our method with the IBM Qiskit simulator, and show that it follows more closely both the analytical solution of the Lindblad equation as well as the behaviour of a real quantum computer, where we ran algorithms involving up to 18 qubits; as such, our protocol offers a more accurate simulator for NISQ devices. The method is flexible enough to potentially describe any noise, including non-Markovian ones. The noise simulator based on this work is available as a python package at this link.

I. INTRODUCTION

Quantum computers are on the way; currently they manage between dozens and hundreds of qubits [1–5], which does not sound as an impressive number, yet it is already good enough to perform interesting tasks [6, 7]. As powerful as they promise to be, quantum computers are far from being ideal: since, as for any quantum system, they can hardly be isolated from the surrounding environment, they are prone to errors, which limit their capabilities. Like in the classical case, error correcting schemes have been developed [8–10] and first tests have been performed [11, 12], but to be implemented they require to the least thousands qubits, which are not available; for the time being, we have to cope with errors.

This stage of development is referred to as Noisy Intermediate-Scale Quantum (NISQ) [6, 13] era; the major aim of the research during this near-term period is to maximize the computational power of current devices in view of the long-term goal of fault-tolerant quantum computation [1].

It is clear that NISQ computers require a good understanding of how noises affect quantum circuits and, in order to do so, a proper modeling of the noises is needed. This requires essentially two major tasks: understanding the major sources of noise affecting the qubits, and writing better algorithms for simulating a given noise model on a classical computer. The present work deals with this second task.

To date, the simulations of noisy digital gate-based quantum computers is implemented by adding appropriate quantum operations before and after each ideal gate [14–16]: schematically, and working with the density matrix formalism, if an ideal (unitary) gate G is supposed to be executed, the noises affecting it are modeled by adding appropriate operations $\mathcal{E}_1$ ($\mathcal{E}_2$) mimicking the noise, before (after) the gate:

$$\rho \;\text{—}\boxed{\mathcal{E}_1}\text{—}\boxed{G}\text{—}\boxed{\mathcal{E}_2}\text{—}\; \rho' \,. \tag{1}$$

Such a modeling completely decouples the action of the controlled operation generating the gate G from that of the environment. This approximation works well if G acts almost instantaneously with respect to the noise, i.e. if the gate time t_g required to implement the gate is much smaller than the characteristic time scales of the system-environment interaction. For instance, in IBM's superconducting devices [17] $t_g \sim 10^{-8}s$, while typical environmental effects such as relaxation and phase damping have characteristic times of order $T_1, T_2 \sim 10^{-4}s$. This justifies why this approach has been implemented by the majority of available noise simulators of NISQ computers (see appendix H).

Yet this approach has some limitations. By separating the action of the gate from that of the noise, it does not represent a faithful description of what happens inside a computer, where the controlled action on the qubit(s) generating the gate and the environment act simultaneously and potentially affect each other. Therefore it is expected not to be fully accurate in describing a NISQ computer, especially when the number of gates and qubits is relatively large, which is actually the regime where simulations are more interesting.

In this article we propose an alternative approach, where the noise is *integrated* into the logical gates, in the sense that the resulting noisy gate is computed by solving for the dynamics generating it, with additional terms describing the noise added to it:

$$\rho \;\text{—}\boxed{\mathcal{G}}\text{—}\; \rho' \,, \tag{2}$$

◇ These authors contributed equally
* giovanni.dibartolomeo@phd.units.it
† michele.vischi@phd.units.it
‡ francesco.cesa@phd.units.it
§ michele.grossi@cern.ch

where in general $\mathcal{G} \neq \mathcal{E}_2 \circ G \circ \mathcal{E}_1$, and under standard assumptions (e.g., Markovianity) it gives an analytic expression for the solution of the Lindblad equation obtained with perturbative methods. Now $\mathcal{G}$ captures, within the limits of validity of Lindblad's equation, the entire physics occurring during the execution of each gate; not only it offers a more accurate description of the system and therefore a better protocol for circuit simulations, but also it helps to understand the different noises acting on the computer, especially in view of possible mitigation strategies. This new approach does not have any computational disadvantage with respect to (1).

As a note, Markovianity, which is the main physical assumption behind the Lindblad equation, and it is a very convenient working hypothesis, can be released in favour of more general noises [14, 18–22]; we will not touch on this possibility here, although the generalization of the approach here introduced is rather straightforward.

Both approaches (1) and (2) have a drawback if they are implemented at the density matrix level: the simulation will be slowed down quadratically as a function of the number of qubits. This drawback can be resolved for (1) by replacing the superoperations $\mathcal{E}_{1,2}$ acting on the density matrix with suitable stochastic operations acting on the state vector [23, 24]; in this way, the noisy algorithm becomes random and each single run of the simulation can be seen as a single run of the algorithm on the noisy quantum computer. The same strategy can be adopted for (2); one writes

$$|\psi\rangle \;\text{—}\boxed{\mathcal{G}_{\boldsymbol{\xi}}}\text{—}\; |\psi'_{\boldsymbol{\xi}}\rangle \;, \tag{3}$$

where $\mathcal{G}_{\boldsymbol{\xi}}$ is a stochastic gate, solution of a stochastic Schrödinger equation, incorporating both the controlled action generating the (otherwise ideal) gate G and the noise. Here $\boldsymbol{\xi}$ denotes a set of stochastic gaussian variables, and stresses the fact that $\mathcal{G}_{\boldsymbol{\xi}}$, and hence $|\psi'_{\boldsymbol{\xi}}\rangle$, are random; we will omit to indicate $\boldsymbol{\xi}$ in the rest of the paper. Physical quantities are obtained by averaging over the noise.

The general procedure therefore is the following. Given a noiseless algorithm, the corresponding noisy one is obtained by replacing each ideal gate with a noisy gate. The resulting noisy algorithm, which is stochastic, is repeated for different realizations of the random variables, as if they were different runs on a physical quantum computer. This produces a statistics of outcomes, to be compared with those of a real computer, or to be used to predict the behavior of a future NISQ device.

As such, as already mentioned, a first application of our approach is to predict the behaviour of NISQ devices, their potentialities and limitations. But its use goes beyond the NISQ-era horizon: by offering a more accurate modeling of the noise, it allows to better understand the physics underlying the functioning of a quantum computer and to enforce appropriate error mitigation schemes [25–27].

The rest of the paper details this program. We present the noisy gates method by designing it on the IBM superconducting computers [17] although the approach is general and can be used to describe any NISQ quantum platform, once the native gate set and the proper noise model are defined.

The paper is organized as follows. In Sec. II we review the main noises affecting superconducting qubits, and how they are described within the Lindblad's formalism; in Sec. III, IV and V we present the general derivation of the noisy gates, specializing it to the native single and two-qubits noisy gates of IBM devices. In Sec. VI we compare the structure of our algorithm with that of IBM Qiskit.

In Sec. VII we present the results of the simulations, which test our algorithm against that of Qiskit in reproducing the solution of the Lindlbad equation, as well as the outcomes from current IBM quantum computers: the simulations show that the proposed method is more accurate and precise compared to that of Qiskit in reproducing the Lindblad equation, with an average improvement between 50% and 90% and more.

The improvement in simulating the real device fluctuates between 10% and 30%, because the underlying noise model is not accurate enough, and also because the devices are not really stable; for a large number of qubits it becomes even lower because the number of runs of the device, which are necessary to recover a good statistic, is too high. In both cases, this is not a limitation of our algorithm, but of the physical model describing the computer.

We conclude our paper with some general remarks and an outlook.

II. REVIEW OF THE NOISE MODEL

The noises which are more relevant in the functioning of superconducting devices have already been characterized in literature [15, 16, 28]; in this section we briefly present them. With good approximation they are described by a Lindblad dynamics [29, 30]:

$$\frac{\mathrm{d}\rho_s}{\mathrm{d}s} = -\frac{i}{\hbar}\big[\mathrm{H}_s, \rho_s\big] + \mathfrak{D}(\rho_s); \tag{4}$$

here, H_s is the Hamiltonian of the system which implements the ideal gate, and $\mathfrak{D}(\rho)$ is a Lindblad term describing the effect of the environment. For convenience, we will describe the evolution with a time schedule $s \in [0,1]$, defined as $s = t/t_g$, where t_g is the duration of a gate.

Apart from state preparation and measurement (SPAM) errors, which happen at the very beginning and very end, during the execution of an algorithm there are two main sources of noise, namely, depolarization and relaxation [28, 31]. The first, which can be ascribed to the imperfections of the device, tends to bring the state towards the totally mixed one, $\mathbb{1}/\sqrt{N}$, where $N = 2^n$ and

n is the number of qubits; for the single qubit, this can be modeled by the following Lindblad term [15, 16],

$$\mathfrak{D}_d(\rho) = \gamma_d \sum_{k=1}^{3} \left[\sigma^k \rho \sigma^k - \rho\right], \tag{5}$$

where $\sigma^1 = \mathrm{X}$, $\sigma^2 = \mathrm{Y}, \sigma^3 = \mathrm{Z}$ are the standard Pauli matrices and $\gamma_d \geq 0$ is the rate at which depolarization occurs.

The second type of noise is due to the interaction of the physical qubits with the surrounding environment; in particular, due the thermalization towards an equilibrium with the environment, energy exchanges occur. In the scenario of interest, this induces the decay of a qubit towards the ground state $|0\rangle$, an effect which is also known as amplitude damping [15, 16]. This damping is characterized by a relaxation time T_1, which identifies the scales at which the initial state decays towards $|0\rangle$; it causes also a damping of the off-diagonal elements of the density matrix in terms of dephasing, which (if only amplitude damping is acting) has a characteristic time $2T_1$. However, at the same time also a contribution of pure dephasing must be taken in account, resulting in an effective dephasing rate $1/T_2 \geq 1/2T_1$. When also $T_1 \geq T_2$ holds (and this is the case of interest to us), the combined action of these two effects, that from now on we will refer to as relaxation or amplitude and phase damping, can be described by the following Lindblad term,

$$\mathfrak{D}_r(\rho) = \gamma_1 \left[\sigma^+ \rho \sigma^- - \frac{1}{2}\left\{\mathbb{P}^{(1)}, \rho\right\}\right] + \gamma_z \left[\mathrm{Z}\rho\mathrm{Z} - \rho\right], \tag{6}$$

where we use the convention $\sigma^{\pm} = (\mathrm{X} \pm i\mathrm{Y})/2$ and $\mathbb{P}^{(1)} = |1\rangle\langle 1|$ is the projector onto $|1\rangle$; the coefficients are related to the characteristic times as $\gamma_1 = t_g T_1^{-1}$ and $\gamma_z = t_g(2T_1 - T_2)/4T_1T_2$.

We will consider both sources of noise together, meaning that the Lindblad term is $\mathfrak{D}(\rho) = \mathfrak{D}_d(\rho) + \mathfrak{D}_R(\rho)$, which can be diagonalized in the canonical Lindblad form by standard procedures. Eventually one obtains the Lindblad term

$$\mathfrak{D}(\rho) = \epsilon^2 \sum_{k=1}^{3} \left[\mathrm{L}_k \rho \mathrm{L}_k^\dagger - \frac{1}{2}\left\{\mathrm{L}_k^\dagger \mathrm{L}_k, \rho\right\}\right], \tag{7}$$

where the non normalized Lindblad operators are

$$\mathrm{L}_1 = \sqrt{\frac{\lambda_1}{\lambda}}\sigma^-, \quad \mathrm{L}_2 = \sqrt{\frac{\lambda_2}{\lambda}}\sigma^+, \quad \mathrm{L}_3 = \sqrt{\frac{\lambda_3}{\lambda}}\mathrm{Z}; \tag{8}$$

here, we set $\lambda_1 = 2\gamma_d$, $\lambda_2 = 2\gamma_d + \gamma_1$, $\lambda_3 = \gamma_d + \gamma_z$ and $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, and we defined the parameter $\epsilon = \sqrt{\lambda}$. As mentioned in the Introduction, in the case of IBM's superconducting devices the typical order of magnitude of the decoherence times is $\sim 10^{-4}$ s; by contrast, the typical order of magnitude of the time to execute a gate is $t_g \sim 10^{-8}$ s, which is small compared to $T_{1,2}$; in particular, one has $\gamma_d, \gamma_1, \gamma_z \ll 1$, which leads to $\epsilon = \sqrt{\lambda} \ll 1$. This justifies the perturbative expansion we will implement later.

While terms of the form (7) describe the dissipation occurring at the single qubit level, one straightforward generalization to the multi-qubit case (the one we will consider in this work) is obtained via the direct sum

$$\mathfrak{D}(\rho) = \bigoplus_{k=1}^{n} \mathfrak{D}^{(k)}(\rho), \tag{9}$$

where the upper index (k) indicates that the Lindblad term (7) acts on the $k-$th qubit. Such a generalization is based on the assumption that single qubit noises are dominating, therefore neglecting cross talks and correlated noises [32]; they can straightforwardly be implemented in our noisy framework, and they will be the subject of future research. We stress that through Eq. (9) we already account for the fact that (for instance, on IBM's devices) multiple-qubit operations are more faulty than single qubit manipulations: when entangling gates are performed, single qubit noises act together, and errors therefore amplify.

Before proceeding, one further comment is in order. Casting the behaviour of a real quantum device in a theoretical model is a hard task, and the more accurate the model, the less general it is. As remarked in the Introduction, the purpose of this work is not that of finding the best noise model for a given quantum computer; rather, given a noise model, we are interested in the best way to simulate the device. The noise model we are considering here is therefore ultimately motivated by the fact that it is accurate enough to already give appreciable results in the simulations, but on the other hand it is also simple enough to efficiently enlighten our main points, and general enough to be readily extended to different platforms. It is understood that better results can be achieved only by specializing more the analysis on physical device to be considered.

III. GENERAL DERIVATION OF NOISY GATES

Let us consider the situation in which the computer executes a gate U_g on a set of n qubits. This is achieved by driving the system with an Hamiltonian H_s for $s \in [0, 1]$, which will induce some unitary evolution U_s, defined by $i\hbar \mathrm{d}\mathrm{U}_s/\mathrm{d}s = \mathrm{H}_s\mathrm{U}_s$, and such that $\mathrm{U}_{s=1} = \mathrm{U}_g$. However, if noises and imperfections are taken in account, this coherent evolution is replaced by a partially non coherent one, which under the assumptions of Markovianity (and complete positivity) is described by a master equation of the form (4) discussed in the previous section, with the Lindblad term given, in our case, by (9) and (7), which needs to be solved in place of the Schrödinger equation. We recall here that the coefficient ϵ is small, $\epsilon \ll 1$.

In order to switch from the density matrix formalism to the state vector formalism, we perform a *linear stochastic*

unraveling of the Lindbald equation [19, 33–36]; specifically we consider the following Itô stochastic differential equation for the state vector: [37]

$$\mathrm{d}\,|\psi_s\rangle = \left[-\frac{i}{\hbar}\mathrm{H}_s\mathrm{d}s + \sum_{k=1}^{N^2-1}\left[i\epsilon\mathrm{dW}_{k,s}\mathrm{L}_k - \frac{\epsilon^2}{2}\mathrm{d}s\mathrm{L}_k^\dagger\mathrm{L}_k\right]\right]|\psi_s\rangle\,, \tag{10}$$

where $\mathrm{dW}_{k,s}$ are differentials of standard independent Wiener processes, i.e. stochastic infinitesimal increments such that $\mathbb{E}\big[\mathrm{dW}_{k,s}\big] = 0$ and $\mathbb{E}\big[\mathrm{dW}_{k,s}\mathrm{dW}_{k',s'}\big] = \delta_{k,k'}\mathrm{d}s$. Eq. (10) is an unraveling of the Lindblad equation in the sense that the density matrix obtained by averaging the pure states $|\psi_s\rangle\langle\psi_s|$ over the noise:

$$\rho_s = \mathbb{E}\Big[\,|\psi_s\rangle\langle\psi_s|\,\Big], \tag{11}$$

is a solution of Eq. (4). In this sense, Eqs. (10) and (4) have the same physical content; the advantage of the stochastic unraveling is that it allows to work with Schrödinger-like equations for the state vector.

One key property of Eq. (10) is that it is linear, and therefore it allows to write the solution as $|\psi_{s=1}\rangle = \mathrm{N}_g\,|\psi_0\rangle$, where N_g can be interpreted as a noisy random gate acting on the system. Since Eq. (10) in general does not preserve the norm of the state vector, the associated gate N_g is not unitary; this is a consequence of the chosen unraveling: one could have chosen norm-preserving unravelings [38, 39], which however are not linear and therefore do not allow for a gate-like formulation. The lack of norm preservation is not a problem since at the statistical level, i.e. when the average over the noise is taken as in (11), one recovers the Lindblad equation, which is trace preserving.

In general, Eq. (10) cannot be solved in a closed form [37, 40] except for few specific cases, for example when all operators commute. In Appendix A we show how an approximate solution to order $\mathcal{O}(\epsilon^2)$ can be derived, which results in the following expression for the noisy version of a noiseless gate U_g:

$$\mathrm{N}_g = \mathrm{U}_g e^{\Lambda} e^{\Xi}, \tag{12}$$

where we defined the deterministic operator:

$$\Lambda := -\frac{\epsilon^2}{2}\int_0^1 \mathrm{d}s \sum_{k=1}^{N^2-1}\left[\mathrm{L}_{k,s}^\dagger\mathrm{L}_{k,s} - \mathrm{L}_{k,s}^2\right] \tag{13}$$

and the stochastic one:

$$\Xi := i\epsilon\sum_{k=1}^{N^2-1}\int_0^1 \mathrm{dW_{k,s}}\mathrm{L}_{k,s}. \tag{14}$$

Note that in Eqs. (13) and (14), $\mathrm{L}_{k,s} = \mathrm{U}_s^\dagger\mathrm{L}_k\mathrm{U}_s$ are the Lindblad operators in the interaction picture, therefore the noiseless part of the dynamics U_s and the noisy one given by the Lindblad operators L_k do not factorize, as it might look from a naive understanding of Eq. (12).

As explained in Appendix A, we omitted the additional term $-(\epsilon^2/2)\sum_{k,l=1}^{N^2-1}\int_0^1 \mathrm{dW}_{k,s}\int_0^s \mathrm{dW}_{l,s'}\big[\mathrm{L}_{k,s},\mathrm{L}_{l,s'}\big]$ in Eq. (14), which in principle should contribute to order ϵ^2; this is legitimate because it is a nested Itô integral of non anticipating functions [37], and hence its stochastic average is 0. For this reason, it drops from all final averaged quantities, and therefore we can neglect it from the start.

Let us also point out that, in the cases of interest to us, the term (13) can always be exponentiated, so that we will always be able to directly calculate e^{Λ}.

The only stochastic term entering the noisy gate N_g is Ξ in Eq. (14), which is a function of several random variables ξ arising from the stochastic processes $\mathrm{W}_{k,s}$. Let us call $\mathrm{L}_{kij,s} = \mathrm{L}^+_{kij,s} + i\mathrm{L}^-_{kij,s}$ the ij-th matrix element of the jump operator $\mathrm{L}_{k,s}$ in the computational basis, divided in real (+) and imaginary (−) part, respectively. Then, each entry of the stochastic matrix is of the form $\Xi_{ij} = i\epsilon\sum_{k=1}^{N^2-1}\big[\xi^+_{kij} + i\xi^-_{kij}\big]$, where we defined the random variables

$$\xi^+_{kij} = \int_0^1 \mathrm{dW}_{k,s}\mathrm{L}^+_{kij,s}, \quad \xi^-_{kij} = \int_0^1 \mathrm{dW}_{k,s}\mathrm{L}^-_{kij,s}, \tag{15}$$

which, being Itô integrals of deterministic functions, are all normally distributed with zero mean, $\mathbb{E}\big[\xi^\pm_{kij}\big] = 0$, and variances $\mathbb{E}[(\xi^\pm_{kij})^2] = \int_0^1 \mathrm{d}s[\mathrm{L}^\pm_{kij}]^2$. Moreover, one can easily check that they are correlated with each other as

$$\mathbb{E}\big[\xi^\pm_{kij}\xi^\pm_{k'i'j'}\big] = \delta_{k,k'}\int_0^1 \mathrm{d}s\mathrm{L}^\pm_{kij,s}\mathrm{L}^\pm_{ki'j',s}. \tag{16}$$

The random variables giving Ξ its stochastic character may be defined in several other ways, and the best choice depends on the specific case of interest. In this section we presented one general strategy for defining them, but in practice this lead to an over estimation of the actual number of random variables needed. By straightforwardly counting, one has at most $2N^2(N^2-1)$ real gaussian random variables for a noisy gate acting on $n = \log_2 N$ qubits, each random variable being correlated with at most other $2N^2-1$ ones. In practice, however, we immediately point out that one shall expect neither the number of random variables, nor the number of correlations between them to really follow this scaling. This is mainly due to the fact that real quantum computers usually perform single and two qubit native gates, and single qubit noises are dominating. For instance, given (9), one can upper bound the number of random variables by $\sim 6N^2\log_2 N$. In the following sections, as we go through the construction of the native set of noisy gates for IBM's quantum computers, we shall make this claim more clear. Note that our derivation works for any choice of the starting Lindblad master equation, meaning that any Markovian noise model can be treated. In particular, while in this paper we specialize on the noise model described in the previous section, one can add device-motivated modifications (such as correlated noises and

leakages to upper levels in the case of IBM's platform); modifications of this kind are left to future research—as also the generalization of our derivation to non Markovian situations.

More details on the difference between our perturbative approximation and the one used in the standard approach (1) can be found in appendix B.

IV. SINGLE QUBIT NOISY GATES

IBM's superconducting devices implement single qubits operations with unitaries of the form $\mathrm{U}(\theta,\phi) = e^{-i\theta \mathrm{R}_{xy}(\phi)/2}$, where we set $\mathrm{R}_{xy}(\phi) = \cos(\phi)\mathrm{X} + \sin(\phi)\mathrm{Y}$; such gates are achieved by driving the system with the Hamiltonian [28, 41]

$$\mathrm{H}(\theta,\phi) = \frac{\theta\hbar}{2}\mathrm{R}_{xy}(\phi) \tag{17}$$

applied for a time $s = 1$ [42]. The Hamiltonian is driven by time-dependent pulses [28], so that in Eq. (17) one should actually consider $\theta \to \omega_s$, and set $\int_0^1 \mathrm{d}s\omega_s = \theta$. In this work we consider constant pulses for simplicity, being the generalization to general functions rather straightforward. It should be noted that the functional form of ω_s affects the action of the noises on the system, meaning that different pulse shapes might lead to smaller noise effects, i.e. error mitigation; this is a question left for future research.

The task now is to derive the noisy gates $\mathrm{N}(\theta,\phi)$ corresponding to the unitaries above, when depolarization and relaxation errors are both taken in account during the evolution.

We begin by computing the evolution of the jump operators in the interaction picture, obtaining the expressions:

$$\sigma_s^{\pm}(\theta,\phi) = \frac{e^{\pm i\phi}}{2}\Big[\mathrm{R}_{xy}(\phi) \pm i\mathrm{R}(2s\theta,\bar{\phi})\Big], \tag{18}$$

and

$$\mathrm{Z}_s = \mathrm{R}(2s\theta,\bar{\phi}), \tag{19}$$

where we defined $\mathrm{R}(\theta,\phi) = \cos(\theta/2)\mathrm{Z} + \sin(\theta/2)\mathrm{R}_{xy}(\phi)$ and for a generic angle α we set $\bar{\alpha} = \alpha + \pi/2$. Then, based on Eq. (12), we compute the deterministic, non unitary term $\Lambda(\theta,\phi)$. Since in the interaction picture the evolution is unitary, one sees that the term corresponding to $k = 3$ is always vanishing, and one has $\Lambda(\theta,\phi) = -\frac{1}{2}\int_0^1 \mathrm{d}s[\epsilon_1^2\sigma_s^+\sigma_s^- + \epsilon_2^2\sigma_s^-\sigma_s^+]$, where we set $\epsilon_k^2 \equiv \epsilon^2\lambda_k/\lambda$. Hence, we first calculate $\sigma_s^{\pm}\sigma_s^{\mp} = \mathrm{U}_s^{\dagger}\sigma^{\pm}\sigma^{\mp}\mathrm{U}_s$, and after integration we get

$$\int_0^1 \mathrm{d}s\sigma_s^{\pm}\sigma_s^{\mp} = \frac{1}{2}\Big[\mathbb{1} \pm \frac{\sin(\theta/2)}{\theta/2}\mathrm{R}(\theta,\bar{\phi})\Big], \tag{20}$$

where $\mathrm{R}(\theta,\phi) = \cos(\theta/2)\mathrm{Z} + \sin(\theta/2)\mathrm{R}_{xy}(\phi)$ and $\bar{\phi} = \phi + \pi/2$, so that one has

$$\Lambda(\theta,\phi) = -\frac{\epsilon_1^2+\epsilon_2^2}{4}\mathbb{1} - \frac{\epsilon_1^2-\epsilon_2^2}{4}\frac{\sin(\theta/2)}{\theta/2}\mathrm{R}\big(\theta,\bar{\phi}\big); \tag{21}$$

such an expression can be readily exponentiated, leading to

$$e^{\Lambda(\theta,\phi)} = e^{-\frac{\epsilon_1^2+\epsilon_2^2}{4}}\Big[\cosh F(\theta) - \mathrm{R}\big(\theta,\bar{\phi}\big)\sinh F(\theta)\Big], \tag{22}$$

where we defined $F(\theta) = \frac{\epsilon_1^2-\epsilon_2^2}{4}\frac{\sin(\theta/2)}{\theta/2}$.

Next, we turn to investigating the stochastic term, $\Xi(\theta,\phi)$. Here, it is convenient to define the following real stochastic variables:

$$\xi_{k,+} = \int_0^1 \mathrm{dW}_{k,s}\cos(s\theta), \quad \xi_{k,-} = \int_0^1 \mathrm{dW}_{k,s}\sin(s\theta), \tag{23}$$

whose variances are:

$$\mathbb{E}\big[\xi_{k,\pm}^2\big] = \frac{1}{2}\big[1 \pm \frac{\sin(2\theta)}{2\theta}\big], \tag{24}$$

while the correlations are:

$$\mathbb{E}\big[\xi_{k,+}\xi_{j,-}\big] = \frac{1-\cos(2\theta)}{4\theta}\delta_{kj}; \tag{25}$$

moreover, we define

$$\xi_{k,w} = \int_0^1 \mathrm{dW}_{k,s}, \tag{26}$$

such that $\mathbb{E}\big[\xi_{k,w}^2\big] = 1$, $\mathbb{E}\big[\xi_{k,+}\xi_{k,w}\big] = \sin(\theta)/\theta$ and $\mathbb{E}\big[\xi_{k,-}\xi_{k,w}\big] = \big[1-\cos(\theta)\big]/\theta$.
Summing all terms and re-arranging them conveniently, we arrive at the following expression

$$\Xi(\theta,\phi) = if_0\mathrm{Z} + if_1\mathrm{R}_{xy}(\phi) + if_2\mathrm{R}_{xy}(\bar{\phi}), \tag{27}$$

where we defined the following set of complex stochastic coefficients:

$$f_0 = \epsilon_3\xi_{3,+} - i\frac{e^{i\phi}\epsilon_2\xi_{2,-} - e^{-i\phi}\epsilon_1\xi_{1,-}}{2}, \tag{28}$$

$$f_1 = \frac{e^{i\phi}\epsilon_2\xi_{2,w} + e^{-i\phi}\epsilon_1\xi_{1,w}}{2}, \tag{29}$$

$$f_2 = \epsilon_3\xi_{3,-} + i\frac{e^{i\phi}\epsilon_2\xi_{2,+} - e^{-i\phi}\epsilon_1\xi_{1,+}}{2}. \tag{30}$$

Since these quantities are all combinations of gaussian random variables with the correlations previously discussed, they can be efficiently sampled with known algorithms; then, the stochastic matrix (27) can be assembled and numerically exponentiated. Multiplication by the deterministic term (22) and then by the noiseless gate $\mathrm{U}(\theta,\phi)$ eventually lead to the noisy gate $\mathrm{N}(\theta,\phi)$ for the single qubit, which, as shown only depends on 8 correlated gaussian variables.

V. TWO-QUBIT NOISY GATES

On IBM's quantum chips, two qubit gates are implemented by a driven cross resonance [28, 41, 43]; labeling with an upper index the qubit each operator

acts on, this consists in the execution of the unitary $\mathrm{U}^{(1,2)}(\theta,\phi) = e^{-i\theta \mathrm{Z}^{(1)}\otimes \mathrm{R}^{(2)}_{xy}(\phi)/2}$, which can be realised by driving the composite system with the Hamiltonian

$$\mathrm{H}^{(1,2)}(\theta,\phi) = \frac{\hbar\theta}{2}\mathrm{Z}^{(1)}\otimes \mathrm{R}^{(2)}_{xy} \tag{31}$$

for a duration $s = 1$, where, from now on, the tensor product symbol will be dropped, unless otherwise specified. In the proposed approach we take in consideration only noises acting on single qubits, so that the Lindblad term reads

$$\mathfrak{D}^{(1,2)}(\rho) = \epsilon^2 \sum_{i\in\{1,2\}}\sum_{k=1}^{3}\big[\mathrm{L}^{(i)}_k \rho \mathrm{L}^{(i)\dagger}_k - \frac{1}{2}\big\{\mathrm{L}^{(i)\dagger}_k \mathrm{L}^{(i)}_k, \rho\big\}\big], \tag{32}$$

where now ρ is the two-qubit statistical operator. The procedure for calculating the noisy gates is the same as in the single qubit case.

First, we compute the Lindblad operators on the first qubit ($i = 1$) in the interaction picture:

$$\sigma^{\pm(1)}_s = e^{\pm is\theta \mathrm{R}^{(2)}_{xy}(\phi)}\sigma^{\pm(1)}, \tag{33}$$

while $\mathrm{Z}^{(1)}_s = \mathrm{Z}^{(1)}$ remains constant as it commutes with the Hamiltonian. For the second qubit ($i = 2$), one has

$$\sigma^{\pm(2)}_s = \frac{e^{\pm i\phi}}{2}\big[\mathrm{R}^{(2)}_{xy}(\phi) \pm i\mathrm{Z}^{(1)}\mathcal{R}(2s\bar{\theta},\bar{\phi})\big], \tag{34}$$

and $\mathrm{Z}^{(2)}_s = \mathcal{R}(2s\theta,\bar{\phi})$, where we defined for convenience

$$\mathcal{R}(\theta,\phi) = \cos(\theta/2)\mathrm{Z}^{(2)} + \sin(\theta/2)\mathrm{Z}^{(1)}\mathrm{R}^{(2)}_{xy}(\phi). \tag{35}$$

The deterministic term $\Lambda(\theta,\phi)$, see Eq.(13), can be calculated straightforwardly, leading to

$$\Lambda(\theta,\phi) = -\frac{\epsilon_1^2+\epsilon_2^2}{2}\mathbb{1} - \frac{\epsilon_1^2-\epsilon_2^2}{4}\Big[\mathrm{Z}^{(1)} + \frac{\sin(\theta/2)}{\theta/2}\mathcal{R}(\theta,\bar{\phi})\Big]; \tag{36}$$

notice that again this term can be exponentiated analytically as all the terms involved commute; in particular, one has

$$e^{\Lambda(\theta,\phi)} = e^{-\frac{\epsilon_1^2+\epsilon_2^2}{2}}\Big[\cosh\big(\frac{\epsilon_1^2-\epsilon_2^2}{4}\big)\mathbb{1} - \mathrm{Z}^{(1)}\sinh\big(\frac{\epsilon_1^2-\epsilon_2^2}{4}\big)\Big]\times \\ \times\Big[\cosh F(\theta) - \mathcal{R}(\theta,\bar{\phi})\sinh F(\theta)\Big], \tag{37}$$

where $F(\theta)$ is the same function defined in the single qubit case.

In order to efficiently write the stochastic term $\Xi(\theta,\phi)$, it is convenient to define, in analogy with the single qubit case, the gaussian random variables

$$\xi^{(i)}_{k,+} = \int_0^1 \mathrm{dW}^{(i)}_{k,s}\cos(s\theta), \quad \xi^{(i)}_{k,-} = \int_0^1 \mathrm{dW}^{(i)}_{k,s}\sin(s\theta), \tag{38}$$

and

$$\xi^{(i)}_{k,w} = \int_0^1 \mathrm{dW}^{(i)}_{k,s}, \tag{39}$$

whose correlations are straightforward to calculate and mimic those already seen in Sec. IV. Then, we can separate $\Xi(\theta,\phi)$ in two parts as $\Xi^{(1)}(\theta,\phi) + \Xi^{(2)}(\theta,\phi)$; the first is equal to

$$\Xi^{(1)}(\theta,\phi) = \epsilon_3\xi^{(1)}_{3,w}\mathrm{Z}^{(1)} + \epsilon_1\big[\xi^{(1)}_{1,+} + i\xi^{(1)}_{1,-}\mathrm{R}^{(2)}_{xy}(\phi)\big]\sigma^{-(1)} + \\ + \epsilon_2\big[\xi^{(1)}_{2,+} + i\xi^{(1)}_{2,-}\mathrm{R}^{(2)}_{xy}(\phi)\big]\sigma^{+(1)}, \tag{40}$$

while the second part reads

$$\Xi^{(2)}(\theta,\phi) = if_w\mathrm{R}^{(2)}_{xy}(\phi) - f_-\mathrm{Z}^{(1)}\mathrm{Z}^{(2)} + f_+\mathrm{R}_{xy}(\bar{\phi}) + \\ + i\epsilon_3\xi^{(2)}_{3,+}\mathrm{Z}^{(2)} + i\epsilon^{(2)}_{3,+}\mathrm{Z}^{(1)}\mathrm{R}_{xy}(\bar{\phi}), \tag{41}$$

where we defined

$$f_w = \frac{1}{2}\big[\epsilon_1 e^{-i\phi}\xi^{(2)}_{1,w} + \epsilon_2 e^{i\phi}\xi^{(2)}_{2,w}\big] \tag{42}$$

and

$$f_\pm = \frac{1}{2}\big[\epsilon_1 e^{-i\phi}\xi^{(2)}_{1,\pm} - \epsilon_2 e^{i\phi}\xi^{(2)}_{2,\pm}\big]. \tag{43}$$

Again, as in the single qubit case, the stochastic matrix $\Xi(\theta,\phi)$ can be assembled by combining gaussian random variables, and hence it can be efficiently sampled and numerically exponentiated; this, combined with the term $\mathrm{U}(\theta,\phi)e^{\Lambda(\theta,\phi)}$, gives the noisy gate for two qubits.

VI. COMPARISON OF THE ALGORITHMS

It is instructive to compare the structure of our approach to noise simulation with that of noise simulators based on the standard approach in Eq. (1). As shown in appendix H all relevant quantum computing frameworks implement such standard approach, and we chose IBM's Qiskit as term of comparison since it is the most developed one; in the next section we will compare also their performances in simulating the Lindblad equation as well as a real quantum computer.

Both methods rely on the state vector formulation, with important differences though. According to the Qiskit documentation [24, 44] the noises are implemented by Kraus maps, which in the density matrix formalism read:

$$\mathcal{E}(\rho) = \sum_i \mathrm{K}_i\rho\mathrm{K}_i^\dagger, \tag{44}$$

where $\sum_i \mathrm{K}_i^\dagger\mathrm{K}_i = \mathbb{1}$. The map can be unraveled as a stochastic map on the state vector by imposing that, at a given time, $|\psi\rangle$ changes randomly as follows:

$$|\psi'\rangle = \frac{1}{\sqrt{p_j}}\mathrm{K}_j|\psi\rangle, \tag{45}$$

with probability:

$$p_j = |\langle\psi|\mathrm{K}_j^\dagger\mathrm{K}_j|\psi\rangle|^2. \tag{46}$$

The associate pseudo code is reported in Alg. 1.

Algorithm 1 QISKIT SIMULATION

Input: Initial state $|\psi_0\rangle$, a noiseless circuit $\mathrm{C} = \{\mathrm{U}^{(1)}, ..., \mathrm{U}^{(n_g)}\}$ composed by n_g gates $\mathrm{U}^{(i)}$ and number of samples N_s

for $0 \leq k \leq N_s$ **do**
  while $1 \leq i \leq n_g$ **do**
    compute $|\psi_k\rangle^{(i)} = \mathrm{U}^{(i)} |\psi_k\rangle^{(i-1)}$
    compute $p_j = |\langle\psi_k|^{(i)} \mathrm{K}_j^\dagger \mathrm{K}_j |\psi_k\rangle^{(i)}|^2$
    sample K_j operator from $\{p_j\}$
    update the state to $|\psi_k\rangle^{(i)} = \frac{1}{\sqrt{p_j}} \mathrm{K}_j |\psi_k\rangle^{(i)}$
  end
  compute $\rho_k = |\psi_k\rangle^{(n_g)} \langle\psi_k|^{(n_g)}$
end
Output: $\rho_f = \frac{1}{N_s} \sum_{k=1}^{N_s} \rho_k$

The time complexity of Alg.1 is primarily determined by the matrix vector multiplication step, exhibiting a complexity of $\mathcal{O}(2^{2n})$, where n is the number of qubits. The space complexity is dominated by the storage of the state vector and it scales as $\mathcal{O}(2^n)$. It has to be noted that when the Kraus operators are not unitary, as for relaxation, one needs to store the intermediate state vectors, which are necessary in order to compute the probabilites in Eq.(46). This operation has the same time and space complexity as those of the the previous step. (This can be avoided for mixed unitary error channels: probabilities are known and independent of the current state.)

Our noisy gates simulation instead is based on the algorithm summarized in Alg. 2.

Algorithm 2 NOISY GATES SIMULATION

Input: Initial state $|\psi_0\rangle$, a noiseless circuit $\mathrm{C} = \{\mathrm{U}^{(1)}, ..., \mathrm{U}^{(n_g)}\}$ composed by n_g gates $\mathrm{U}^{(i)}$ and number of samples N_s

for $0 \leq k \leq N_s$ **do**
  map a noisy circuit $\tilde{\mathrm{C}} = \{\mathrm{N}^{(1)}, ..., \mathrm{N}^{(n_g)}\}$ on C
  sample stochastic processes ξ inside noisy gates $\mathrm{N}^{(i)}$
  compute $|\psi_k\rangle = \mathrm{N}^{(n_g)} \dots \mathrm{N}^{(1)} |\psi_0\rangle$
  compute $\rho_k = |\psi_k\rangle \langle\psi_k|$
end
Output: $\rho_f = \frac{1}{N_s} \sum_{k=1}^{N_s} \rho_k$

The time complexity of Alg. 2 is again $\mathcal{O}(2^{2n})$, determined by the matrix vector multiplication step. Analogously, the space complexity is $\mathcal{O}(2^n)$. We notice that in Alg. 2 there is no need to perform the scalar product in Eq. (46). Moreover, all optimization to reduce the time complexity that are possible for the first step of Alg. 1 are also possible for Alg. 2. Finally, both algorithms perform samples of random numbers, but this operation has a constant scaling.

VII. SIMULATIONS

We now study the performances of our noisy gates method, and compare them with those of Qiskit's simulator [44]. First, in subsection VII A we test the two approaches against the solution of Lindblad equation (4), by studying a repeated application of IBM's native gate set. Then, in subsection VII B we compare the predictions of both methods with the behaviour of an actual quantum computer, by running the inverse QFT algorithm on the IBM's quantum processors ibmq_kolkata and ibmq_oslo. In appendix G we perform the same analysis by running the GHZ algorithm on ibm_oslo. All simulations are performed by using the noise model described in section II (see also appendices C and D). The implementation of the work proposed in this paper is open source and available as a python package at this link. It allows the user to run noisy simulations.

A. Comparison with the numerical solution of Lindblad equations

First, let us compare our method with the one implemented in the Qiskit simulator for the task of simulating the Lindblad equation. To this purpose, we simulate the same Lindblad equation with both methods, obtaining the density matrix ρ^{ng}, from the noisy gates simulation, and the density matrix ρ^{ibm} from the Qiskit simulation. We then benchmark the results with the density matrix σ obtained by directly solving numerically the Lindblad equation with Mathematica [45]. We compare these density matrices by computing the Hellinger distances $\mathcal{H}_\sigma^{\mathrm{ng}} = \mathcal{H}(\rho^{\mathrm{ng}}, \sigma)$, $\mathcal{H}_\sigma^{\mathrm{ibm}} = \mathcal{H}(\rho^{\mathrm{ibm}}, \sigma)$ where the Hellinger distance is defined by

$$\mathcal{H}(\rho, \sigma) = \frac{1}{\sqrt{2}} \sqrt{\sum_{k=1}^{N} \left(\sqrt{\rho_{kk}} - \sqrt{\sigma_{kk}}\right)^2}, \tag{47}$$

with ρ_{kk} (σ_{kk}) the diagonal elements of ρ (σ). Note that the Hellinger distance is a classical measure of the distance between the readout probability distributions: while it cannot be interpreted as a distance between quantum states (it does not take in account the coherences), it directly compares the concrete outputs of the real device, which are classical (the oucomes of Z measurements). In appendix F we also compute the fidelities $\mathcal{F}_\sigma^{\mathrm{ng}} = \mathcal{F}(\rho^{\mathrm{ng}}, \sigma)$ and $\mathcal{F}_\sigma^{\mathrm{ibm}} = \mathcal{F}(\rho^{\mathrm{ibm}}, \sigma)$. We run the simulations on both single and two qubit gates. Considering the native gate set of IBM's quantum computers, $\{\mathrm{R}_z(\phi), \mathrm{X}, \mathrm{SX}, \mathrm{CNOT}\}$, we remind that $\mathrm{R}_z(\phi)$ are implemented as virtual gates [28, 41], i.e. they are noiseless, and the CNOT gates are implemented by combining single qubit gates in Eq. (17) and CR gates in Eq. (31) [28, 41, 46]. Moreover, X and SX gates are both rotations around the X-axis for different values of θ, see Eq. (17). Thus for our purposes, it is sufficient to simulate the X, CR and CNOT gates affected by noises.

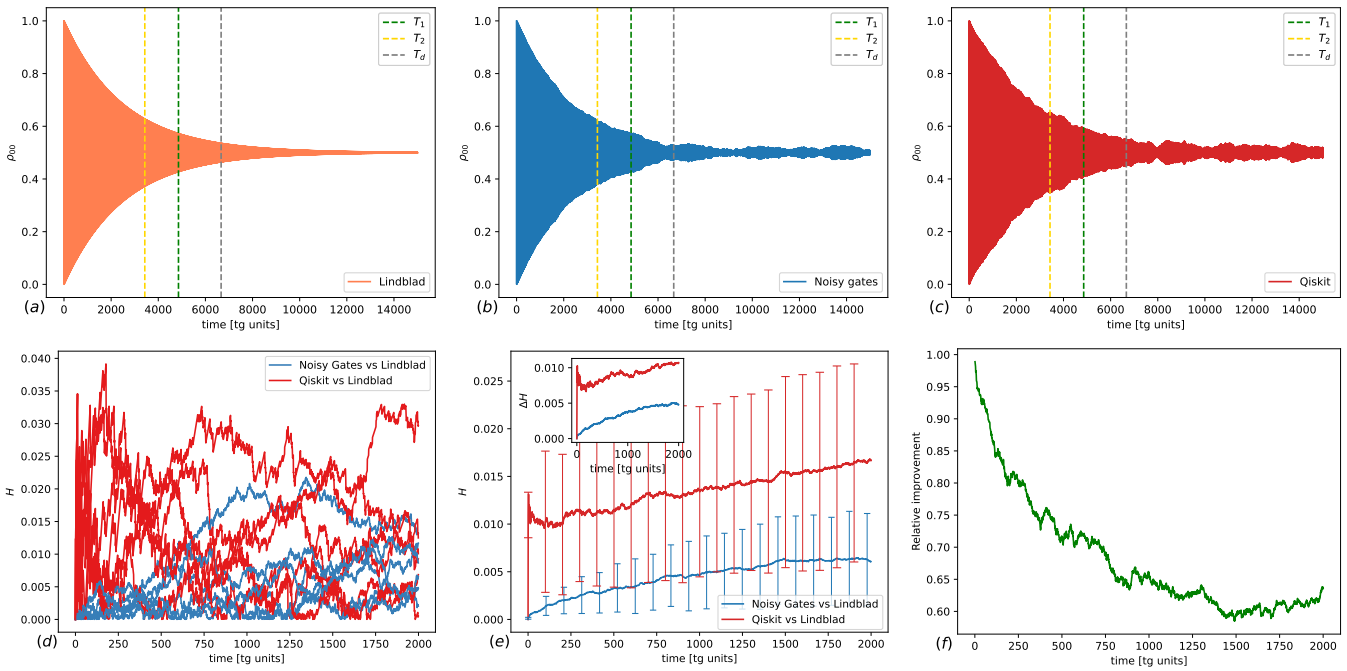

FIG. 1. Repetition of X gates. The three upper panels (a), (b), (c) show the time evolution of the $\rho_{00} = \langle 0|\rho|0\rangle$ entry of the density matrix. The numerical solution of the Lindblad equation is displayed in orange (a), that of the noisy gates simulation in blue (b), and that of the Qiskit simulation in red (c). The noisy gates and Qiskit simulations are obtained with 1000 samples, and qualitatively they reproduce the time evolution of the Lindblad equation. Vertical dashed lines in the three top panels represent the time scales of relaxation T_1 (green), T_2 (yellow) and depolarization T_d (grey). Panel (d) shows the Hellinger distances $\mathcal{H}_\sigma^{\rm ng}$, in blue, and $\mathcal{H}_\sigma^{\rm ibm}$, in red, as a function of time. Different curves are obtained from 100 independent runs of the two methods (for better readability only five are shown), where each simulation is obtained by averaging over 1000 samples. Panel (e) shows the mean of the Hellinger distances $\bar{\mathcal{H}}_\sigma^{\rm ng}$, and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$, obtained from the 100 independent runs, and vertical error bars show their standard deviations $\Delta\mathcal{H}_\sigma^{\rm ng}$, $\Delta\mathcal{H}_\sigma^{\rm ibm}$. The inset displays $\Delta\mathcal{H}_\sigma^{\rm ng}$ and $\Delta\mathcal{H}_\sigma^{\rm ibm}$ as functions of time. Panel (f) shows the relative improvement of the distance $\bar{\mathcal{H}}_\sigma^{\rm ng}$ with respect to $\bar{\mathcal{H}}_\sigma^{\rm ibm}$, calculated as $|\bar{\mathcal{H}}_\sigma^{\rm ibm} - \bar{\mathcal{H}}_\sigma^{\rm ng}|/\bar{\mathcal{H}}_\sigma^{\rm ibm}$. The fact that noises drive the system towards the maximally mixed state is the reason why the improvement decreases in time. The noisy gates and the standard approaches lead to the same predictions when one is close to decoherence times, as the noise is dominant over the unitary evolution. In the interesting regime $[0, 2000 \cdot t_g]$ before decoherence dominates, our improvement is always above 60%.

Single qubit simulations. We first simulate a repetition of X gates, each of which can be obtained by setting $\theta = \pi$ and $\phi = 0$ in Eq. (17); we initialize the qubit in $|0\rangle$ and we use the qubit noise parameters of ibmq_manila (more details on the device can be found in appendix E). We evolve the state of the qubit for a time $T = \mathcal{N} t_g$, with $\mathcal{N} = 15000$. In the upper panels of Fig. 1 we plot the time evolution of the population of the ground state, $\rho_{00} = \langle 0|\rho|0\rangle$, as obtained with the three methods. In the noiseless case, ρ_{00} should oscillate between 0 and 1 with period $2t_g$, as at each step of t_g a complete X rotation is performed; in the presence of noises, the oscillations are damped due to the relaxation of the qubit, while the depolarization drives probabilities towards the asymptotic value $\rho_{00} \to 0.5$.

Both our simulation and that obtained using Qiskit's simlulator qualitatively reproduce this behaviour. In Fig. 1 we have also highlighted with dashed vertical lines the characteristic times of relaxation and depolarization (see the caption); for times approaching these values the state is not a reliable quantum state anymore, as the density matrix becomes completely mixed. Given this consideration, in the lower plots we stop at $\mathcal{N} = 2000$.

In order to inspect which of the two models reproduces more accurately and precisely the Lindblad evolution, we have run 100 independent simulations with both the noisy gates simulator and the Qiskit simulator, computing for each run the Hellinger distances $\mathcal{H}_\sigma^{\rm ng}$, $\mathcal{H}_\sigma^{\rm ibm}$. We computed the means over the 100 independent simulations, $\bar{\mathcal{H}}_\sigma^{\rm ng}$, $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ and the standard deviations $\Delta\mathcal{H}_\sigma^{\rm ng}$, $\Delta\mathcal{H}_\sigma^{\rm ibm}$. These quantities are shown in the lower panels of Fig. 1. During the relevant time interval $[0, T]$ the Hellinger distance of the noisy gates simulator is closer to zero, than that obtained with the Qiskit simulator. Both results are compatible within the error bars, however the standard deviations associated to the noisy gates simulations are significantly smaller than those associated to the Qiskit simulations, as also highlighted in the inset of Fig. 1 (e). We notice that the difference between $\bar{\mathcal{H}}_\sigma^{\rm ng}$ and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ is of the order $\sim 10^{-3} - 10^{-2}$, and this corresponds to a relative improvement, calculated as $|\bar{\mathcal{H}}_\sigma^{\rm ibm} - \bar{\mathcal{H}}_\sigma^{\rm ng}|/\bar{\mathcal{H}}_\sigma^{\rm ibm}$, in the range from 90% to 60% as time

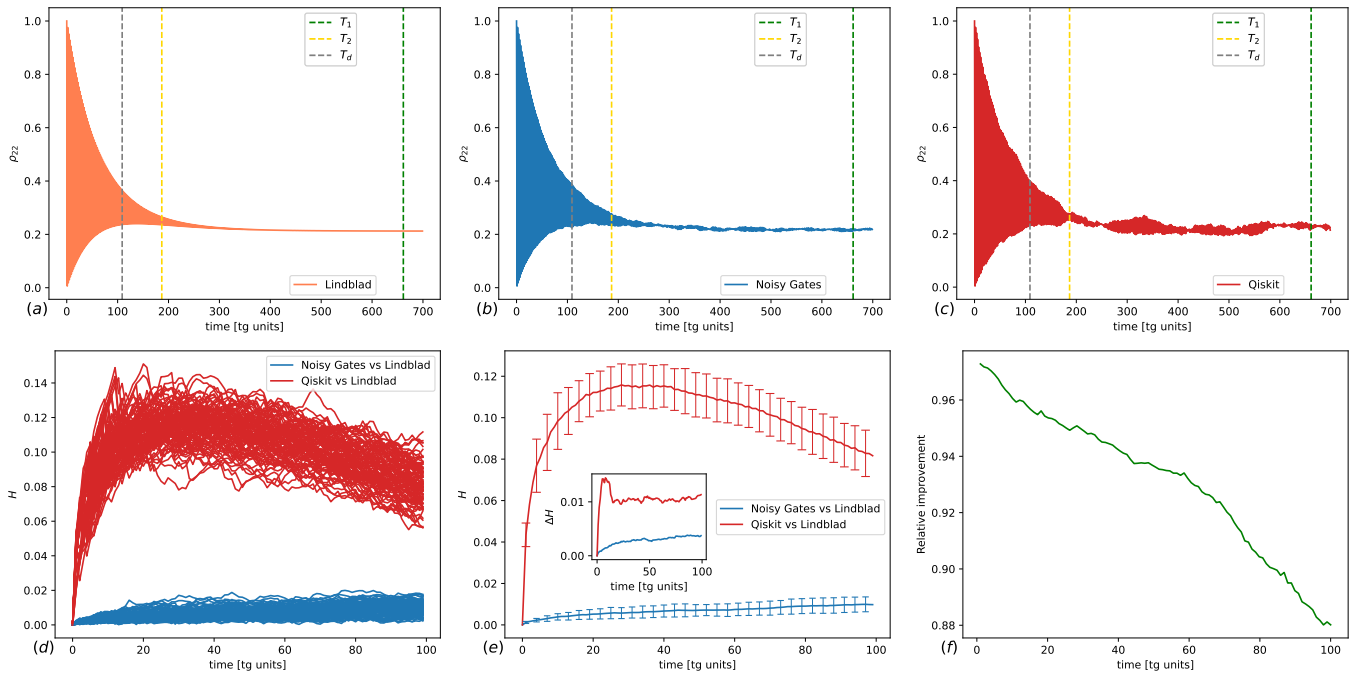

FIG. 2. Repetition of CR gates. The three upper panels (a), (b), (c) show the time evolution of the ρ_{22} entry of the density matrix for the CR gate with $\theta = \pi$ and $\phi = 0$. Colors have the same meaning as for Fig. 1. Vertical dashed lines represent the time scales of relaxation, T_1 (in green) and T_2 (in yellow) of the target qubit, and depolarization T_d (grey). The noisy gates simulations reproduce qualitatively better the time evolution obtained from the direct numerical solution of the Lindblad equation. Panels (d) and (e) display the Hellinger distances $\mathcal{H}_\sigma^{\rm ng}$, in blue, and $\mathcal{H}_\sigma^{\rm ibm}$, in red, as a function of time, for a repetition of CR gates. The plots have the same meaning as for Fig.1. Panel (f) shows the relative improvement of the distance $\bar{\mathcal{H}}_\sigma^{\rm ng}$ with respect to $\bar{\mathcal{H}}_\sigma^{\rm ibm}$, calculated as $|\bar{\mathcal{H}}_\sigma^{\rm ibm} - \bar{\mathcal{H}}_\sigma^{\rm ng}|/\bar{\mathcal{H}}_\sigma^{\rm ibm}$. The fact that noises drive the system towards the maximally mixed state is the reason why the improvement decreases in time. The noisy gates and the standard approaches lead to the same predictions when one is close to decoherence times, as the noise is dominant over the unitary evolution. In the interesting regime $[0, 100 \cdot t_g]$ our improvement is always above 88%.

increases. The relative improvement is shown in Fig. 1 (f). The fact that noises drive the system towards the maximally mixed state is the reason why the improvement decreases over time. The noisy gates and the standard approaches lead to the same predictions when one is close to decoherence times. Indeed after such times the strength of the noise is dominant over the unitary evolution, or the Hamiltonian contribution is negligible with respect to the Lindblad term (see Eq.(4)), which is the same in the two approaches. In the interesting regime $[0, T]$ our improvement is always above 60%. In appendix F we repeat a similar analysis for the fidelities.

Two qubits simulations. Next, we simulate a repetition of Cross-resonance gates as defined in Eq. (31), where we choose $\phi = 0$ and $\theta = \pi$. We initialize the system in the state $|10\rangle$ and we use the qubit noise parameters of ibmq_manila. In the three upper panels of Fig. 2 we show the time evolution of the entry $\rho_{22} = \langle 10|\rho|10\rangle$; the x-axis is normalized in terms of the two-qubit gate time t_g. The two-qubit state goes asymptotically towards the completely mixed state as ρ_{22} reaches the asymptotic value 0.25. The probability ρ_{22}, which in the ideal case should flip between one and zero, is again damped over time by relaxation effects. Again, we have highlighted with vertical dashed lines the characteristic time scales of the noises, showing only the T_1 and T_2 values of the target qubit as representative values. The depolarizing error is the dominant one, spoiling the quantum state already after $\sim$ 100 CR gates; for this reason, the the lower panels we consider a total duration $\mathcal{N} \sim 100$. As before, we report the Hellinger distances, showing the different results of 100 independent simulations together with their mean and standard deviation in the three lower panels of Fig. 2.

As in the single qubit case, within the relevant time interval $[0, T]$ the Hellinger distances obtained with the noisy gates simulations are closer to zero than those obtained with the Qiskit simulator. However now, differently from the single qubit case, the two results are not compatible within error bars. Moreover the difference between $\bar{\mathcal{H}}_\sigma^{\rm ng}$ and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ is now of the order $\sim 10^{-1}$. This corresponds to a relative improvement in the range from 90% to 88% as time increases, shown in Fig. 2 (f). In the interesting regime $[0, T]$ our improvement is always above 88%. We notice that in Fig. 2 (e) the value of $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ approaches that of $\bar{\mathcal{H}}_\sigma^{\rm ng}$ for times close to 100 CR gate times. The reason why this happens is the same explained above for the single qubit case. In appendix F

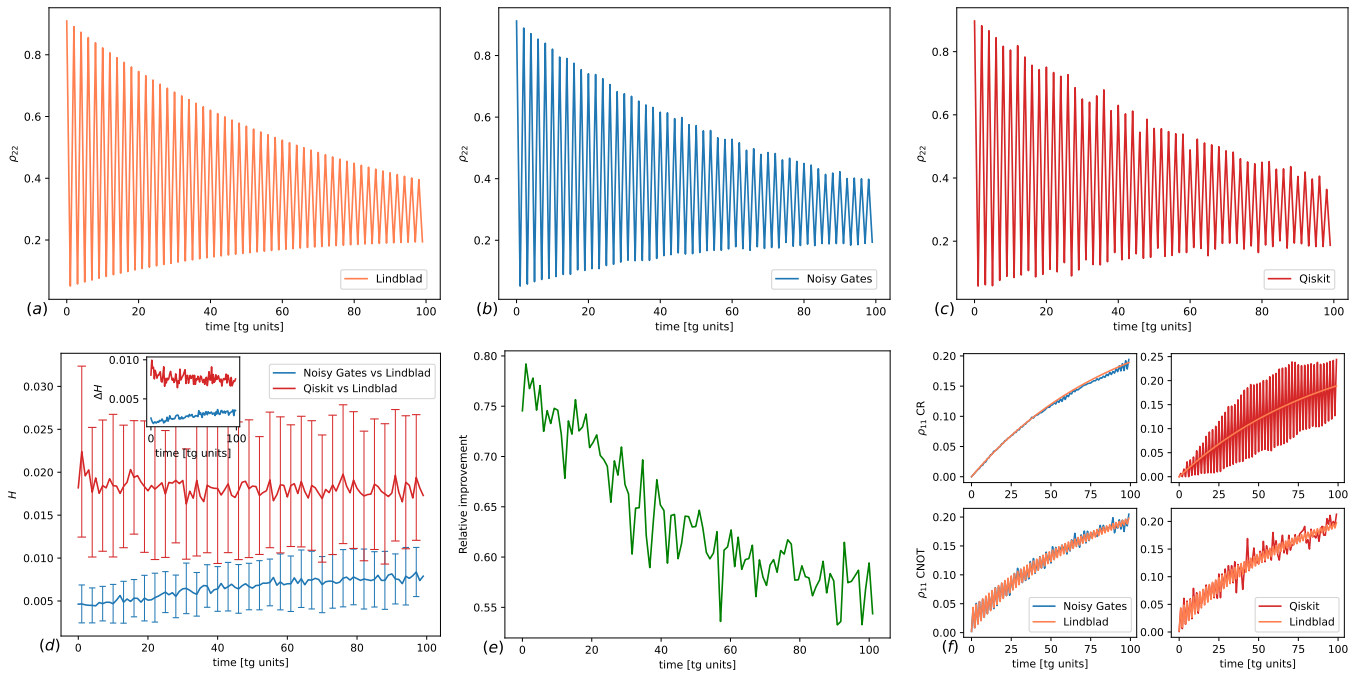

FIG. 3. Repetition of CNOT gates. The three upper panels (a), (b), (c) show the time evolution of the ρ_{22} entry of the density matrix for the CNOT gate. Colors have the same meaning as for Fig. 1. The noisy gates simulations reproduce qualitatively better the time evolution obtained from the direct numerical solution of the Lindblad equation. Panel (d) displays mean of the Hellinger distances $\bar{\mathcal{H}}_\sigma^{\rm ng}$, in blue, and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$, in red, and their standard deviations as functions of time. Panel (e) shows the relative improvement. The fact that noises drive the system towards the maximally mixed state is again the reason why the improvement decreases in time: the noisy gates and the standard approaches lead to the same predictions when one is close to decoherence times, as the noise is dominant over the unitary evolution. In the interesting regime $[0, 100 \cdot t_g]$ our improvement is always above 55%. The upper subplots of panel (f) show the time evolution of the ρ_{11} entry of the density matrix for the same sequence of CR gates in Fig. 2 and the lower subplots show the time evolution of the ρ_{11} entry of the density matrix for the sequence of CNOT gates. Colors have the same meaning as for Fig. 1. For the CR gates, the Qiskit simulation of ρ_{11} is visibly different from the Lindblad evolution, thus explaining the higher improvement of the noisy gates simulation in the Hellinger distance in Fig. 2.

we repeat a similar analysis for the fidelities.

We then perform the analysis for a repetition of CNOT gates, for an initial state given by $|10\rangle$ and qubit noise parameters of ibmq_quito (see appendix E). We notice that in this simulation we implement each CNOT gate directly without expressing it as a combination of single qubit gates and CR gates, as it is done in IBM devices. We make this choice because in this way it is easier to solve numerically the target Lindblad equation. At each time step of the evolution we simulate a circuit with an increasing number of CNOT gates and measurements at the end. Thus we add SPAM channels (see appendices C and D) to model measurements errors. This allows to extend the analysis to runs on real hardware, that involve measurements, as we will show later in subsection VII B. In the three upper panels of Fig. 3 we show the time evolution of the $\rho_{22} = \langle 10|\rho|10\rangle$ entry of the density matrix. The relevant time interval is again given by a total duration of $\mathcal{N} \sim 100$ gates. Indeed the depolarizing error in this case spoils the quantum state after ~ 120 CNOT gates. Fig. 3 (d) shows the mean of the Hellinger distances $\bar{\mathcal{H}}_\sigma^{\rm ng}$ (in blue) and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ (in red) and their standard deviations $\Delta\mathcal{H}_\sigma^{\rm ng}$ and $\Delta\mathcal{H}_\sigma^{\rm ibm}$, also shown in the inset. Once more, within the relevant time interval $[0, T]$ the Hellinger distances obtained with the noisy gates simulations are closer to zero than those obtained with the Qiskit simulator and the difference between $\bar{\mathcal{H}}_\sigma^{\rm ng}$ and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ is of the order $\sim 10^{-2}$. This corresponds to a relative improvement in the range from 80% to 55% as time increases. This is shown in Fig. 3 (e). In the interesting regime $[0, T]$ the relative improvement is always above 55%.

By looking at Figs. 2 (e) and 3 (d), we notice that the improvement in the Hellinger distance gained by using the noisy gates approach is much higher for CR gates with respect to CNOT gates. The reason why this happens is clarified in Fig. 3 (f). The panel consists of two upper subplots showing the time evolution of the ρ_{11} entry of the density matrix for the CR gates and two lower subplots showing the time evolution of the ρ_{11} entry of the density matrix for the CNOT gates. Similarly to the convention used above, orange curves are obtained with the numerical solution of the Lindblad equation, blue curves are obtained with the noisy gates simulations

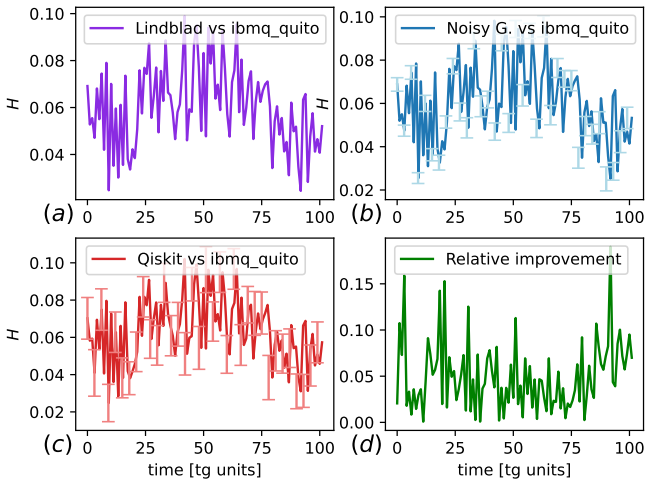

FIG. 4. Repetition of CNOT gates. Panel (a) shows the Hellinger distance $\mathcal{H}_\sigma^\chi$ between the Lindblad evolution and ibmq_quito for the repetition of CNOT gates. Panel (b) shows the mean Hellinger distance $\bar{\mathcal{H}}_\chi^{\rm ng}$ and the standard deviations between the noisy gates simulation and ibmq_quito. Panel (c) shows the mean Hellinger distance $\bar{\mathcal{H}}_\chi^{\rm ibm}$ and the standard deviations between the Qiskit simulation and ibmq_quito. Panel (d) shows the relative improvement calculated as $|\bar{\mathcal{H}}_\chi^{\rm ibm} - \bar{\mathcal{H}}_\chi^{\rm ng}|/\bar{\mathcal{H}}_\chi^{\rm ibm}$. The relative improvement is around 10%. The smaller relative improvement with respect to those shown in the previous figures, is mainly due to additional noises present in ibmq devices, i.e. crosstalks, correlated noises and coherent errors.

and red curves are obtained with Qiskit simulations. The noisy gates simulations make good predictions for both gate sequences, as the blue curves follow closely the orange curves. On the other hand, the Qiskit simulation for the CR gates is visibly different from the numerical solution of the Lindblad equation. This might be due to the fact that the CR gate is a block diagonal matrix with $X(\theta)$ in the upper block and $X(-\theta)$ in the lower block while the CNOT gate is block diagonal with an identity in the upper block and $X(\theta)$ in the lower block. The identity in the CNOT might lead to a lower influence of noises on the ρ_{00} and ρ_{11} entries of the density matrix. These observations explain why the Hellinger distances obtained with the noisy gates in different simulations are very good and similar to each other, while the Hellinger distance obtained with Qiskit is better for the CNOT with respect to the CR. Nevertheless, the noisy gates approach always outperforms the standard one by a significant amount, as shown by the relative improvements.

B. Comparison with the behaviour of a real quantum computer

Now, we inspect the performances of the noisy gates approach when trying to reproduce the behaviour of a real quantum computation. To this purpose, we first extend the analysis of the CNOT gates sequence in subsection VII A, and then we focus on the inverse Quantum Fourier Transform (QFT†). When dealing with a real hardware, we must take into account that the noise model we implement in this analysis (see section II) might not be accurate enough in describing the device, and that different quantum devices might behave very differently from one another. As we will show, despite the choice of a simple noise model and the instability of ibmq devices, our approach is still able to outperform the standard one also when compared with the real hardware.

CNOT simulations. We run the sequence of CNOT gates of subsection VII A on ibmq_quito, available on the cloud and comprising 7 superconducting transmon qubits [47] (see appendix E) and we reconstruct the density matrix χ obtained from the physical device, to be compared with the density matrices $\rho^{\rm ng}$, $\rho^{\rm ibm}$ and σ obtained for the CNOT simulations discussed in subsection VII A. We remark again that for the CNOT simulations of subsection VII A, we implemented each CNOT gate directly without expressing it as a combination of single qubit gates and CR gates, as it is done in IBM devices, because in this way it is easier to solve numerically the target Lindblad equation. We create a list of circuits, each consisting of an increasing number of CNOT gates, and measure each circuit 1000 times to obtain the output probability distributions, thus deriving the evolution of the outcome probabilities as the number of gates increases. As noted above, since each circuit involves measurements we added a SPAM error to model measurement errors.

The Hellinger distance $\mathcal{H}_\sigma^\chi = \mathcal{H}(\chi,\sigma)$ between the Lindblad evolution and the evolution obtained with ibmq_quito is shown in Fig. 4 (a). This distance is three to tens time larger with respect to $\bar{\mathcal{H}}_\sigma^{\rm ng}$ and $\bar{\mathcal{H}}_\sigma^{\rm ibm}$ that are shown in Fig. 3 (d). While the standard approach and the noisy gates approach have a certain level of agreement with the Lindblad equation, the latter is deviating from the quantum hardware by a significantly higher level. This is also the reason why it is not possible to appreciate the difference between the mean Hellinger distance $\bar{\mathcal{H}}_\chi^{\rm ng} = \bar{\mathcal{H}}(\rho^{\rm ng},\chi)$ of the noisy gates with ibmq_quito and the mean Hellinger distance $\bar{\mathcal{H}}_\chi^{\rm ibm} = \bar{\mathcal{H}}(\rho^{\rm ibm},\chi)$ of Qiskit with ibmq_quito, as shown in Fig. 4 (b) and Fig. 4 (c). Fig. 4 (d) shows the relative improvement with respect to the device, which is calculated as $|\bar{\mathcal{H}}_\chi^{\rm ibm} - \bar{\mathcal{H}}_\chi^{\rm ng}|/\bar{\mathcal{H}}_\chi^{\rm ibm}$. The relative improvement is around 10%. The smaller relative improvement with respect to those shown in the previous subsection is only to a small extend due to the fact that we do not decompose CNOT gates. The main reason, as we explain when discussing the simulations of the QFT (see below), is that additional noises are present in ibmq devices, i.e. crosstalks, correlated noises and coherent errors [48, 49]. The simple noise model that we consider in this work does not take such noises into account.

QFT simulations The (QFT†) is a subroutine of many important quantum algorithms, as for example the Shor's algorithm [50, 51]. An important feature of QFT† is that

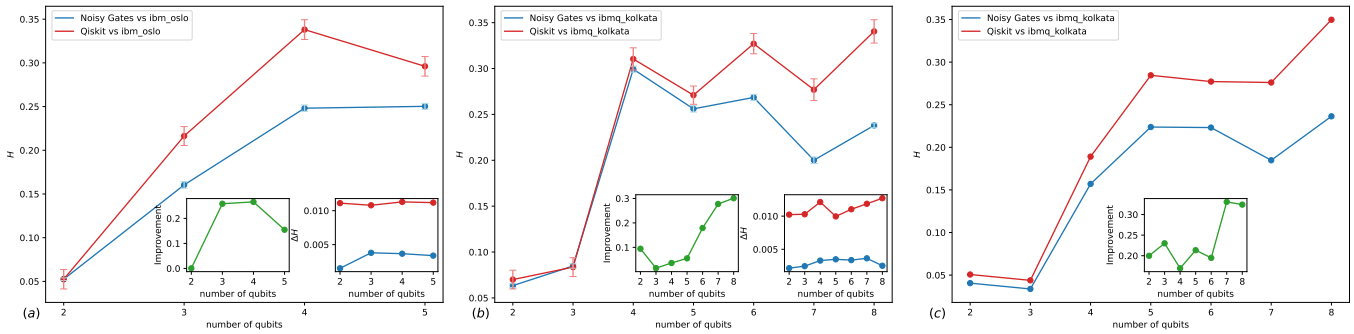

FIG. 5. Quantum Fourier Transform. Panel (a) shows the Hellinger distances between the noisy gate approach and ibmq_oslo, and between the Qiskit simulator and ibmq_oslo, when executing the $\mathrm{QFT}^\dagger$ algorithm for $n = 2, \ldots, 5$ qubits. Each value is the mean of 100 independent simulations for the noisy gates, in blue, and for the Qiskit simulations, in red. The left inset shows the relative improvement, calculated as $|\bar{\mathcal{H}}_\chi^{\mathrm{ibm}} - \bar{\mathcal{H}}_\chi^{\mathrm{ng}}|/\bar{\mathcal{H}}_\chi^{\mathrm{ibm}}$, while the right inset shows the standard deviations as functions of the number of qubits. Panel (b) is the same as (a), with ibmq_oslo replaced by ibmq_kolkata and the number of qubit going up to 8. In panel (c) the comparison presented in (b) has been repeated a second time on ibmq_kolkata; in this case, only a single simulation of 1000 samples is considered. The inset shows the relative improvement.

the circuit for n qubits is readily extendable to $n+1$ qubits; thus we can efficiently test the robustness of the method as the circuit's width and depth increase. We run $\mathrm{QFT}^\dagger$ for $n = 2, \ldots, 5$ on ibmq_oslo and for $n = 2, \ldots, 18$ on ibmq_kolkata. These devices are available on the cloud, comprising respectively 7 and 27 superconducting transmon qubits [47], see appendix E for further details. We set as input of $\mathrm{QFT}^\dagger$ the state $|+\rangle^{\otimes n}$, obtained by applying a layer of Hadamard gates on each qubit initialized in $|0\rangle$. In this way the ideal output of $\mathrm{QFT}^\dagger$ should be $|0\rangle^{\otimes n}$. Runs on real quantum computers are performed by taking 1000 shots, i.e. measurements. We also run the corresponding noisy gates and Qiskit simulations. (In appendix G we perform a similar analysis for the GHZ algorithm [52].)

Implementing $\mathrm{QFT}^\dagger$ circuit on ibmq devices requires to transpile the circuit into their native gate set. We have defined a custom noise model in Qiskit, by adding after each gate of the transpiled circuit the depolarizing and relaxation channels, and the SPAM channel before measurements, see appendix C. Similarly, in the noisy gates simulation each gate is replaced with its noisy version according to the noise model in section II. During idle-times of qubits we put the relaxation noise gates (see appendix D) in order to take into account the stand-by times of the physical qubits; before measurements, we apply SPAM noise gates (see appendix D) which accounts for read-out errors. In these simulations the CNOT gates inside the circuits are decomposed in terms of single-qubit and CR gates, as in ibmq devices.

In order to measure the performance of different approaches in simulating the behaviour of the quantum computer, we look at their distance with the outcomes of the real device; this is achieved by computing the Hellinger distance between the probability distributions, and it can be done without performing full tomography on the quantum states, which scales exponentially with the number of qubits and becomes unfeasible for the current simulations.

In Fig. 5 (a) we plot the average values of $\mathcal{H}_\chi^{\mathrm{ng}} = \mathcal{H}(\rho^{\mathrm{ng}}, \chi)$, $\mathcal{H}_\chi^{\mathrm{ibm}} = \mathcal{H}(\rho^{\mathrm{ibm}}, \chi)$ as the number of qubits n increases from 2 to 5, where now the diagonal elements of χ are the outcome probabilities of ibmq_oslo. Fig. 5 (b) displays again the average values of $\mathcal{H}_\chi^{\mathrm{ng}} = \mathcal{H}(\rho^{\mathrm{ng}}, \chi)$, $\mathcal{H}_\chi^{\mathrm{ibm}} = \mathcal{H}(\rho^{\mathrm{ibm}}, \chi)$ up to 8 qubits, where now the diagonal elements of χ come from ibmq_kolkata. As shown in Fig. 5 (c) we compute again $\mathcal{H}_\chi^{\mathrm{ng}}$, $\mathcal{H}_\chi^{\mathrm{ibm}}$ to test the stability of ibmq_kolkata in different runs.

As in the previous section, we have run 100 independent simulations, each including 1000 samples, for both methods and for each n, in order to compute the standard deviations $\Delta\mathcal{H}_\chi^{\mathrm{ng}}$, $\Delta\mathcal{H}_\chi^{\mathrm{ibm}}$ shown in the insets of Fig. 5. Only for Fig. 5 (c) we have run a single simulation of 1000 samples, thus standard deviations are not present. We notice that for every n we get $\bar{\mathcal{H}}_\chi^{\mathrm{ng}} < \bar{\mathcal{H}}_\chi^{\mathrm{ibm}}$ and $\Delta\mathcal{H}_\chi^{\mathrm{ng}} < \Delta\mathcal{H}_\chi^{\mathrm{ibm}}$. The relative improvement, shown in green in the insets of of Fig. 5, changes significantly between different devices and also for the same device but in different moments, namely with different noise parameters, meaning that the performances of such devices are not very stable. For example at $n = 3$, in the left panel the relative improvement is $\sim 25\%$, in the central panel it is $\sim 5\%$ and in the right panel it is $\sim 25\%$. The highest relative improvement obtained with the run on ibmq_oslo is $\sim 30\%$ and for runs on ibmq_kolkata is $\sim 35\%$.

The results show that our method is more accurate than existing ones. Actually, it reproduces the Lindblad dynamics better (Figures 1 and 2) than the dynamics of the quantum devices. The reason, mentioned before, is that quantum devices are affected by additional and more complicated noises, which are not taken into account by the noise model we are using; we stress again that to find a better noise model is not the scope of this work, and will be subject of future research.

The simulations on ibmq_kolkata have been extended to 18 qubits to test the computational scalability of the

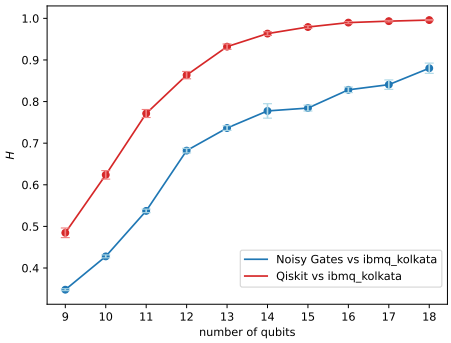

FIG. 6. Mean Hellinger distances $\bar{\mathcal{H}}_\chi^{\rm ng}$ and $\bar{\mathcal{H}}_\chi^{\rm ibm}$ and their standard deviations $\Delta\mathcal{H}_\chi^{\rm ng}$ and $\Delta\mathcal{H}_\chi^{\rm ibm}$ from $n = 9$ to $n = 18$ qubits for the QFT† executed on ibmq_kolkata. Since for $n \geq 9$ the depth of the circuit is such that noises make the resulting probability distribution very flat, 1000 runs of the circuit on the quantum device, which returns a single computational basis state in each run, are not sufficient to reconstruct faithfully the probability distribution over the 2^n basis states; by increasing the number of qubits, a 0 probability is associated to an increasing number of basis states. The simulator instead does does not suffer from this limitation: in each run, it returns a non-zero value for each possible output. Then the respective probability distributions differ more and more, and this is the reason why the Hellinger distance rapidly increases. Despite the fact that the number of runs of the device is not sufficient to derive clear conclusions, we notice that the noisy gates approach still performs better than the standard one.

noisy gates simulator. In Fig. 6 we show the mean Hellinger distances $\bar{\mathcal{H}}_\chi^{\rm ng}$ and $\bar{\mathcal{H}}_\chi^{\rm ibm}$ and their standard deviations $\Delta\mathcal{H}_\chi^{\rm ng}$ and $\Delta\mathcal{H}_\chi^{\rm ibm}$ from $n = 9$ to $n = 18$ qubits: simulations apparently become rapidly bad, since the Hellinger distance approaches 1, its maximum value. There is a clear reason behind that, which does not represent a limitation of our simulator. First of all, for such an high number of qubits, the depth of the transpiled circuit is so large that noises dominate [53] and the resulting probability distributions are very flat. Then, to recover a faithful probability distribution over the 2^n possible outcomes by the quantum device, which returns a single outcome in each run, the number of circuit runs must be significantly larger than 2^n. Therefore, for $n \geq 9$ a number of runs equal to 1000 is not sufficient (and increasing this number becomes soon impractical): the output distribution from the device is increasingly dominated by 0's, while our simulator returns (in general) a non-zero probability for each output state: this makes the Hellinger distances of Fig. 6 approach 1. Nevertheless, also in this case the noisy gates approach performs better than the standard one, even if the number of runs of the quantum device are not enough to properly recover the full probability distribution.

As a final remark, we stress that we obtain better results with respect to Qiskit, despite the fact that we have chosen the simplest time dependent pulse shape in the Hamiltonians (see Eq. (17) and Eq. (31)).

VIII. CONCLUSIONS AND OUTLOOK

We have developed a novel approach, called noisy quantum gates, to improve classical simulations of NISQ computers: it is based on integrating the noise into the gates, rather than keeping gates and noise as two separate dynamics. We have shown that our approach is very successful in simulating the Lindblad dynamics, with a relative improvement between 50% and 90% and more, compared with the standard gate-noise separation method.

When compared against real quantum devices, the improvement fluctuates between 10% and 30%; this is largely due to the fact that the underlying noise model is too simple to accurately represent the dynamics of the device, as discussed in connection to the simulation of the CNOT gate. This is not a weakness of the noisy gate approach here presented, but of the underlying noise model, which we used since it is rather standard in the literature.

There is a number of potential improvements that can be straightforwardly implemented; all of them require an update of the noise model, not of the simulation strategy, which is already very good. First of all, there are likely additional single-qubit errors which should be taken into account, for example those induced by the driving pulses. Secondly, in the present work we considered only non-correlated single-qubit errors, but the method can easily accommodate also correlated two-qubits errors [48, 54] by introducing proper correlated noises into the stochastic equations. Another possible extension of the approach is to add in the Hamiltonians small interactions between adjacent qubits in order to mimic cross talk errors [49, 55]. Last, the current version of the noisy gates approach relies on the Lindblad equation that works in the Markovian limit; this is reflected in the fact that we used stochastic equations based on white noises. The approach can be generalized to non-Markovian dynamics by using colored noises, as already discussed in the literature in different contexts [19–22].

Furthermore, our approach is also useful for other purposes that go beyond plane error analysis. For example, the shape of the pulse in the driving Hamiltonians, (see Eq. (17) and Eq. (31)), can affect the noise. In our work we chose for simplicity a rectangular shape, but usually in real devices different shapes can be used, for example Gaussian ones. Consequentially, a natural application of our approach is error mitigation [56, 57], by optimizing the parameters of the pulse in order to minimize the effect of the noise [58–60]; the optimization can be performed for example by exploiting machine learning techniques, to find the best pulse parameters, which can be tested on real quantum hardware.

In this work we specified our approach to the native gate set and noise model of IBM devices; clearly the approach is general and can be used to describe in principle any NISQ platform.

ACKNOWLEDGMENTS

G.D.B. and M.V. thank A. Gundhi, L. Pintucci and L.L. Viteritti for useful discussions. F.C., G.D.B. and M.V. acknowledge the financial support from University of Trieste and INFN; F.C. acknowledges also the financial support of PON Ricerca e Innovazione 2014-2020 (D.M. 1061, 10.08.21) and QTI (Quantum Telecommunications Italy). M.G. is supported by CERN through the CERN Quantum Technology Initiative. S.D. acknowledges the financial support from INFN. A.B. acknowledges financial support from the EIC Pathfinder project QuCoM (GA no. 101046973), the PNRR PE National Quantum Science and Technology Institute (PE0000023), the University of Trieste and INFN.

Appendix A: Derivation of the approximate solution

In this appendix, let us show how the approximate solution in Eq. (12) to Eq. (10) can be rigorously derived to order $\mathcal{O}(\epsilon^2)$. We propose two different methods.

1. Perturbative expansion in the interaction picture

As a first proof, let us perform the stochastic unraveling in the interaction picture, hence defining the state quantum trajectory at any time as $|\psi_s\rangle = \mathrm{U}_s\,|\phi_s\rangle$, where the state vector $|\phi_s\rangle$ is the solution, at time s, of the Itô equation

$$\mathrm{d}\,|\phi_s\rangle = \left[i\epsilon \sum_{k=1}^{N^2-1} \mathrm{dW}_{k,s}\mathrm{L}_{k,s} - \frac{\epsilon^2}{2}\sum_{k=1}^{N^2-1} \mathrm{d}s\mathrm{L}^\dagger_{k,s}\mathrm{L}_{k,s}\right]|\phi_s\rangle\,; \tag{A1}$$

here, $\mathrm{dW}_{k,s}$ are defined as in the main text, and we defined the jump operators in the interaction picture, $\mathrm{L}_{k,s} := \mathrm{U}_s^\dagger \mathrm{L}_k \mathrm{U}_s$. Then, by dividing the time interval $s \in [0,1]$ in infinitesimal steps of width $1/M$ and taking the limit $M \to \infty$, formally the solution to Eq. (10) can be written as

$$\mathrm{N} = \mathrm{U}_g \lim_{M\to\infty} \mathrm{N}_M, \tag{A2}$$

where we defined $\mathrm{N}_M := \prod_{m=0}^{M-1} \exp\left[\epsilon \mathrm{B}_m + \frac{\epsilon^2}{2}\mathrm{A}_m\right]$, with

$$\mathrm{A}_m = -\frac{1}{M}\sum_{k=1}^{N^2-1}\left[\mathrm{L}^\dagger_{k,m/M}\mathrm{L}_{k,m/M} - \mathrm{L}^2_{k,m/M}\right] \tag{A3}$$

and

$$\mathrm{B}_m = i\sum_{k=1}^{N^2-1} \mathrm{L}_{k,m/M}\int_{m/M}^{(m+1)/M} \mathrm{dW}_{k,s}. \tag{A4}$$

For general purposes (and, in particular, for ours) N can not be calculated analytically; hence, we show how to obtain a general form to the second order in ϵ (i.e., to first order in λt_g). First, let us prove that the following approximation holds:

$$\mathrm{N}_M = e^{\frac{\epsilon^2}{2}\mathcal{A}_M} e^{\epsilon\mathcal{B}_M + \frac{\epsilon^2}{2}\mathcal{C}_M} + \mathcal{O}(\epsilon^3), \tag{A5}$$

where we defined $\mathcal{B}_M = \sum_{k=0}^{M-1}\mathrm{B}_k$, $\mathcal{A}_M = \sum_{k=0}^{M-1}\mathrm{A}_k$, and

$$\mathcal{C}_M = \sum_{k=0}^{M-1}\sum_{j=0}^{k}\left[\mathrm{B}_k,\mathrm{B}_\mathrm{j}\right] = \sum_{k=0}^{M-1}\left[\mathrm{B}_k,\mathcal{B}_k\right]. \tag{A6}$$

The proof follows by induction. First, one can straightforwardly check that Eq. (A5) holds for $M = 0$; then, suppose it holds for $\tilde{M} = M-1$. Since by definition $\mathrm{N}_M = e^{\epsilon\mathrm{B}_M + \frac{\epsilon^2}{2}\mathrm{A}_M}\mathrm{N}_{M-1}$, applying the inductive hypothesis one can see that

$$\begin{aligned}\mathrm{N}_M &= 1 + \epsilon\mathcal{B}_M + \frac{\epsilon^2}{2}\left[\mathcal{A}_M + \mathcal{B}_M^2 + \mathcal{C}_M\right] + \mathcal{O}(\epsilon^3)\\ &= e^{\frac{\epsilon^2}{2}\mathcal{A}_M} e^{\epsilon\mathcal{B}_M + \frac{\epsilon^2}{2}\mathcal{C}_M} + \mathcal{O}(\epsilon^3),\end{aligned} \tag{A7}$$

which concludes the proof. Then, inserting (A5) in the formal expression for N, one can perform the limit $M \to \infty$, ending up with $\mathrm{N} = \mathrm{U}_g e^{\Lambda} e^{\Xi}$, where we defined

$$\Lambda = -\frac{\epsilon^2}{2}\int_0^1 \mathrm{d}s \sum_{k=1}^{N^2-1}\left[\mathrm{L}^\dagger_{k,s}\mathrm{L}_{k,s} - \mathrm{L}^2_{k,s}\right] \tag{A8}$$

and

$$\Xi = i\epsilon\sum_{k=1}^{N^2-1}\int_0^1 \mathrm{dW}_{k,s}\mathrm{L}_{k,s} - \frac{\epsilon^2}{2}\mathcal{C}; \tag{A9}$$

here, $\mathcal{C} = \sum_{k,l=1}^{N^2-1}\int_0^1 \mathrm{dW}_{k,s}\int_0^s \mathrm{dW}_{l,s'}\left[\mathrm{L}_{k,s},\mathrm{L}_{l,s'}\right]$. As explained in the main text, this term can actually be dropped at second order in ϵ, leading to the expressions given in (13) and (14).

2. Small noise expansion

A second approach makes use of a perturbative method known as *small noise expansion* or *asymptotic perturbative expansion* [37]. For simplicity, let us consider the SDE with one single Lindblad operator,

$$d\,|\psi_s\rangle = \left[-\frac{i}{\hbar}\mathrm{H}_s\mathrm{d}s + i\epsilon\mathrm{L}\mathrm{dW}_s - \frac{\epsilon^2}{2}\mathrm{L}^\dagger\mathrm{L}\mathrm{d}s\right]|\psi_s\rangle\,, \tag{A10}$$

the generalization to N^2-1 Lindblad operators being straightforward, and let us set the following ansatz:

$$|\psi_s\rangle = |\psi^0_s\rangle + \epsilon\,|\psi^1_s\rangle + \epsilon^2\,|\psi^2_s\rangle + \dots \tag{A11}$$

Substituting this ansatz into Eq.(A10) and equating terms with the same power of ϵ, up to second order we

get a system of SDEs:

$$\begin{aligned}
\mathrm{d}\left|\psi_s^0\right\rangle &= -\frac{i}{\hbar}\mathrm{H}_s\left|\psi_s^0\right\rangle \mathrm{d}s \\
\mathrm{d}\left|\psi_s^1\right\rangle &= -\frac{i}{\hbar}\mathrm{H}_s\left|\psi_s^1\right\rangle \mathrm{d}s + iL\left|\psi_s^0\right\rangle \mathrm{dW}_s \\
\mathrm{d}\left|\psi_s^2\right\rangle &= -\frac{i}{\hbar}\mathrm{H}_s\left|\psi_s^2\right\rangle \mathrm{d}s + iL\left|\psi_s^1\right\rangle \mathrm{dW}_s - \frac{1}{2}L^\dagger L\left|\psi_s^0\right\rangle \mathrm{d}s,
\end{aligned} \tag{A12}$$

which must be solved with the initial conditions $\left|\psi_0^0\right\rangle = |\psi_0\rangle$. The zeroth order differential equation is the deterministic one given by the Hamiltonian evolution alone, hence its solution is simply $\left|\psi_s^0\right\rangle = \mathrm{U}_s|\psi_0\rangle$. The first order SDE is an example of a time-dependent Ornstein-Uhlenbeck process [37]: the solution is

$$\left|\psi_s^1\right\rangle = i\mathrm{U}_s\mathrm{S}_s|\psi_0\rangle\,, \tag{A13}$$

where we defined $\mathrm{S}_s = \int_0^s \mathrm{dW}_\tau \mathrm{L}_\tau$. Finally, the solution to the second order SDE is

$$\left|\psi_s^2\right\rangle = -\mathrm{U}_s\int_0^s\left[\frac{1}{2}\mathrm{L}_s^\dagger\mathrm{L}_s\mathrm{d}s + \mathrm{L}_s\mathrm{S}_s\mathrm{dW}_s\right]|\psi_0\rangle\,, \tag{A14}$$

where $\mathrm{L}_s = \mathrm{U}_s^\dagger L \mathrm{U}_s$. Then, the solution at order ϵ^2 is given by $|\psi_1\rangle = \mathrm{N}|\psi_0\rangle + \mathcal{O}(\epsilon^3)$, where the evolution operator is $\mathrm{N} = \mathrm{U}_g\mathrm{N}'$, with

$$\mathrm{N}' = \left[\mathbb{1} + \epsilon\mathrm{S}_1 - \epsilon^2\int_0^1\left[\frac{1}{2}\mathrm{L}_s^\dagger\mathrm{L}_s\mathrm{d}s + \mathrm{L}_s\mathrm{S}_s\mathrm{dW}_s\right]\right]. \tag{A15}$$

In order to evaluate the solution in the form given in the main text, we make use of the following equality:

$$\int_0^\tau \mathrm{dW}_s\mathrm{L}_s\mathrm{S}_s = \frac{1}{2}\left[\mathrm{S}_s^2 + \int_0^\tau \mathrm{dW}_s[\mathrm{L}_s,\mathrm{S}_s] - \int_0^\tau \mathrm{d}s\mathrm{L}_s^2\right] \tag{A16}$$

obtained by using the Itô rule [37] for each entry of the stochastic matrices. Substituting this expression into Eq.(A15), we get to second order:

$$\begin{aligned}
\mathrm{N}' &= \mathbb{1} + i\epsilon\mathrm{S}_1 - \frac{\epsilon^2}{2}\left[\mathrm{S}_1^2 + \int_0^1 \mathrm{d}s\Big[\mathrm{L}_s^\dagger - \mathrm{L}_s\Big]\mathrm{L}_s + \mathcal{C}\right] = \\
&= e^\Lambda e^\Xi + \mathcal{O}(\epsilon^3),
\end{aligned}$$

where Λ, Ξ and $\mathcal{C} = \int_0^1 \mathrm{dW}_s\big[\mathrm{L}_s,\mathrm{S}_s\big]$ are the same quantities defined in the main text.

Appendix B: Comparison of the approximations

We focus on the main differences between the standard approximation made in error analysis against the one considered in the noisy gates approach.

Given the following Lindblad master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t = -\frac{i}{\hbar}[H_t,\rho_t] + \gamma\mathfrak{L}\left[\rho_t\right], \tag{B1}$$

where $\mathfrak{L}\left[\rho_t\right] = L\rho_t L^\dagger - \frac{1}{2}\{L^\dagger L,\rho_t\}$, let's move to the interaction picture by defining $\chi_t = U_{t,t_0}^\dagger\rho_t U_{t,t_0}$ and $\chi_{t_0} = \rho_{t_0}$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\chi_t = \gamma\mathfrak{L}(t)\left[\chi_t\right], \tag{B2}$$

where $\mathfrak{L}(t)\left[\chi_t\right] = U_{t,t_0}^\dagger\mathfrak{L}\left[\rho_t\right]U_{t,t_0}$.

The formal solution of Eq. (B2) is

$$\chi_t = \mathrm{T}\left[e^{\gamma\int_{t_0}^t \mathrm{d}s\mathfrak{L}(s)}\right]\chi_{t_0}, \tag{B3}$$

where $\mathrm{T}\left[\cdot\right]$ is the time ordering. Thus in the Schrödinger picture we can write the formal solution of Eq. (B1) as

$$\rho_t = U_{t,t_0}\mathrm{T}\left[e^{\gamma\int_{t_0}^t \mathrm{d}s\mathfrak{L}(s)}\right]\rho_{t_0}U_{t,t_0}^\dagger. \tag{B4}$$

- *Standard approximation* - The main approximation that can be found in the literature is to separate the Hamiltonian dynamics from the noise one [15, 16]. This choice is based on the observation that in general in quantum devices $\omega >> \gamma$, where ω is the pulse frequency of the Hamiltonian. Thus, the noise dynamics can be seen as frozen with respect to the faster Hamiltonian one. It means that in Eq. (B4) one assumes

$$\mathfrak{L}(t) \simeq \mathfrak{L} \tag{B5}$$

getting

$$\rho_t \simeq U_{t,t_0}e^{\gamma\mathfrak{L}\cdot(t-t_0)}\rho_{t_0}U_{t,t_0}^\dagger. \tag{B6}$$

We notice that indeed in Eq. (B6) the two dynamics are independent.

- *Noisy gates approximation* - Also in this case the approximation is based on $\omega >> \gamma$, but we assume that γ is not small enough to completely separate the dynamics. An example of this can be seen in the devices of IBM where the noise evolution can be influenced in a non-negligible manner by the pulse of the drive Hamiltonian [46, 61]. Thus we make a first order approximation over γ in Eq. (B4)

$$\mathrm{T}\left[e^{\gamma\int_{t_0}^t \mathrm{d}s\mathfrak{L}(s)}\right] \simeq 1 + \gamma\int_{t_0}^t \mathrm{d}s\mathfrak{L}(s), \tag{B7}$$

and we get

$$\rho_t \simeq U_{t,t_0}\left(1 + \gamma\int_{t_0}^t \mathrm{d}s\mathfrak{L}(s)\right)\rho_{t_0}U_{t,t_0}^\dagger. \tag{B8}$$

In Eq. (B8) the noise depends on the Hamiltonian dynamics through $\mathfrak{L}(s)$. We stress that the perturbative solution of the SDE in the noisy gates model reproduce density matrices of the form of Eq. (B8).

Appendix C: Kraus maps used in Qiskit simulations

The error channels that we included in the custom Qiskit noise model are a composition of depolarization and relaxation after the gates and bitflip before measurements. With relaxation we mean the amplitude and phase damping channel. In this appendix we show the corresponding Kraus maps for these channels that are used in Qiskit simulations through Alg. 1 in section VI.

1. State Preparation and measurement (SPAM)

This kind of error is usually described as a bit flip channel that acts on a single qubit [16]. Hence, its Kraus representation reads:

$$\mathcal{E}(\rho) = (1-p)\rho + pX\rho X, \tag{C1}$$

where ρ is the density matrix of a single qubit, X is the x-Pauli matrix and p is the probability of having a flip of the states of the computational basis. The probability p that we used in the simulations is the readout error provided as a calibration parameter for IBM devices, see appendix E.

2. Depolarization

Depolarization drives the qubit towards the maximally mixed state [16] and models incoherent gate infidelities. Its Kraus representation reads:

$$\mathcal{E}(\rho) = (1-\frac{3}{4}p)\rho + \frac{p}{4}X\rho X + \frac{p}{4}Y\rho Y + \frac{p}{4}Z\rho Z, \tag{C2}$$

where ρ is the density matrix of a single qubit, X, Y, Z are the Pauli matrices and $p/4$ is the equal probability of having a bit flip, a phase flip or a bit and phase flip of the states of the computational basis. The probability p that we used in the simulations is the gate error provided as a calibration parameter for IBM devices, see appendix E.

3. Amplitude and phase damping (Relaxation)

The amplitude-damping channel describes the decay $|1\rangle \to |0\rangle$ due to the interaction with the environment; on the other hand, phase-damping represents the process in which phase coherences decay over time. Here we briefly call relaxation the combination of both effects. The Kraus representation is given by [15, 16]

$$\mathcal{E}(\rho) = K\rho K + p_1\sigma^+\rho\sigma^- + p_z\mathcal{P}_1\rho\mathcal{P}_1, \tag{C3}$$

where we defined

$$K = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p_1-p_z} \end{pmatrix}; \tag{C4}$$

as usual, $\sigma^+ = |0\rangle\langle 1|$, $\sigma^- = |1\rangle\langle 0|$ and $\mathcal{P}_1 = |1\rangle\langle 1|$. Moreover, $p_1 = 1-e^{-t/T_1}$ is the probability of amplitude damping, T_1 being the relaxation time (the time it takes for the qubit to decay in the ground state), and $p_z = (1-p_1)p_{pd}$, where $p_{pd} = 1-e^{-t/T_{pd}}$ and $T_{pd} = T_1T_2/(2T_1 - T_2)$, T_2 being the decoherence time. We mention that the time scales T_1 and T_2 are related as $T_2 \leq 2T_1$. The times T_1 and T_2 that we used in the simulations are directly provided as calibration parameters for IBM devices, see appendix E.

Appendix D: Noise gates for Spam and Relaxation on idle qubits

In this section we address SPAM and relaxation noises on idle qubits, where the corresponding noise gates can be derived exactly [33, 62, 63]. We do not consider depolarization error on idle qubits, because this channel is used to model incoherent gate infidelities.

1. Noise gate for SPAM

The Kraus map of SPAM is in Eq.(C1) of appendix C. Assuming a behaviour in time of the form $p = (1-e^{-2t/T})/2$ for a characteristic time $T = \gamma^{-1}$, one gets the corresponding Lindblad master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t = \gamma(X\rho_t X - \rho_t). \tag{D1}$$

The associated stochastic differential equation is

$$\mathrm{d}|\psi_t\rangle = \left[i\sqrt{\gamma}X\mathrm{d}W_t - \frac{\gamma}{2}\mathrm{d}t\right]|\psi_t\rangle. \tag{D2}$$

This equation is analytically solvable with standard methods [37, 63], and thus we can exactly evaluate the corresponding noise gate as

$$N^{\mathrm{SPAM}}(t,t_0) = e^{i\sqrt{\gamma}X\bar{W}(t,t_0)}, \tag{D3}$$

where $\bar{W}(t,t_0) := \int_{t_0}^{t}\mathrm{d}W_s$. In this case, the noise gate happens to be unitary, thus we can interpret it as a stochastic Schrödinger evolution due to the presence of the Wiener process $\bar{W}(t,t_0)$. In the simulations we can directly sample $\bar{W}(t,t_0)$ from a Gaussian distribution with mean $\mathbb{E}[\bar{W}(t,t_0)] = 0$ and variance $\mathbb{E}[\bar{W}^2(t,t_0)] = t-t_0$.

2. Noise gate for amplitude and phase damping (Relaxation)

The Kraus map of the amplitude and phase damping is in Eq.(C2) of appendix C. Defining $\gamma_1 = 1/T_1$, $\gamma_{pd} = 1/T_{pd}$, the corresponding Lindblad equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t = \gamma_1\sigma^+\rho_t\sigma^- - \frac{\gamma_1}{2}\{\mathcal{P}_1,\rho_t\} + \frac{\gamma_{pd}}{4}(Z\rho_t Z - \rho_t), \tag{D4}$$

and the stochastic term of the relative Itô equation reads:

$$\mathrm{d}\mathcal{W} = i\sqrt{\gamma_1}\sigma^+\mathrm{d}W_{t,1} - \frac{\gamma_1}{2}\mathcal{P}_1\mathrm{d}t + i\sqrt{\frac{\gamma_{pd}}{4}}Z\mathrm{d}W_{t,2} - \frac{\gamma_{pd}}{8}\mathrm{d}t. \tag{D5}$$

With this stochastic term the Itô equation is analytically solvable [40] and we get the following non-unitary noisy gate

$$N^{\mathrm{relax}}(t,t_0) = \begin{pmatrix} e^{i\alpha\bar{W}_2(t,t_0)} & iS(t,t_0)e^{i\alpha\bar{W}_2(t,t_0)} \\ 0 & e^{-\frac{\gamma_1}{2}(t-t_0)}e^{-i\alpha\bar{W}_2(t,t_0)} \end{pmatrix}, \tag{D6}$$

where we defined for simplicity $\alpha := \sqrt{\gamma_{pd}/4}$, and

$$S(t,t_0) = \sqrt{\gamma_1}\int_{t_0}^{t} e^{-\frac{\gamma_1}{2}(s-t_0)}e^{-2i\alpha\bar{W}_2(s,t_0)}\mathrm{d}W_{s,1} \tag{D7}$$

is a complex stochastic Itô process. In principle, such a term is problematic in view of a simulation, since it is not easy to sample. To understand this, look for instance at the real part,

$$S_R(t,t_0) = \sqrt{\gamma_1}\int_{t_0}^{t} e^{-\frac{\gamma_1}{2}(s-t_0)}\cos\Big(2\alpha\bar{W}_2(s,t_0)\Big)\mathrm{d}W_{s,1}; \tag{D8}$$

this is an Itô integral of a stochastic function, and it is not easy to derive its probability distribution; thus, sampling $S(t,t_0)$ may be problematic. We can avoid such a difficulty by adequately substituting $N^{\mathrm{relax}}(t,t_0)$ with some modified noisy gate, which is equivalent to the former once the average is carried out, in the sense that Eq. (C3) still holds even if the new noisy gate is not a solution of the unraveling (D5) anymore. For instance, it is straightforward to verify that this holds for the following choice:

$$\tilde{N}^{\mathrm{relax}}(t,t_0) = \begin{pmatrix} e^{i\alpha\bar{W}_2(t,t_0)} & i\tilde{S}(t,t_0)e^{-i\alpha\bar{W}_2(t,t_0)} \\ 0 & e^{-\frac{\gamma_1}{2}(t-t_0)}e^{-i\alpha\bar{W}_2(t,t_0)} \end{pmatrix}, \tag{D9}$$

with the definition

$$\tilde{S}(t,t_0) = \sqrt{\gamma_1}\int_{t_0}^{t} e^{-\frac{\gamma_1}{2}(s-t_0)}\mathrm{d}W_{s,1}; \tag{D10}$$

i.e., one always has that

$$\mathbb{E}\big[N^{\mathrm{relax}}\left|\psi\right\rangle\left\langle\psi\right|N^{\mathrm{relax}\dagger}\big] = \mathbb{E}\big[\tilde{N}^{\mathrm{relax}}\left|\psi\right\rangle\left\langle\psi\right|\tilde{N}^{\mathrm{relax}\dagger}\big]. \tag{D11}$$

The difference is that now the process $\tilde{S}(t,t_0)$ is just the Itô integral of a deterministic function, hence we know that it must have a Gaussian statistics [37], which makes it more convenient for a simulation.

Appendix E: Device parameters

For the simulations in Sec. VII and in appendix G we used the device parameters provided by IBM. Here we report the average value of such parameters.

ibmq_manila contains 5 fixed-frequency transmons qubits [47], with median fundamental transition frequency of 4.962 GHz and median anharmonicity of −0.34358 GHz. The median qubit lifetime T_1 of the qubits is 149.11μs, the median coherence time T_2 is 44.43μs and the median readout error is 0.0217. The single qubit gate error varies between 1.975×10^{-4} and 6.138×10^{-4}, while the CNOT error varies between 7.072×10^{-3} and 1.125×10^{-2}, depending on the specific connection. In the simulations that reproduce the Lindblad equations, parameters of qubits zero and one were used. ibmq_kolkata contains 27 fixed-frequency transmons qubits, with median fundamental transition frequency of 5.102 GHz and median anharmonicity of −0.34345 GHz. The median qubit lifetime T_1 of the qubits is 127.39μs, the median coherence time T_2 is 86.41μs and the median readout error is 0.0132. The single qubit gate error varies between 1.443×10^{-4} and 5.410×10^{-3}, while the CNOT error varies between 4.214×10^{-3} and 1×10^{-2}, depending on the specific connection. The qubits, which are used to run QFT† algorithm, belong to the list [0,1,4,7,10,12,15,18,21,23,24,25,22,19,16,14,11,8,5,3,2]. ibmq_quito contains 7 fixed-frequency transmons qubits, with the median fundamental transition frequency of 5.164 GHz and median anharmonicity of −0.3315 GHz. The median qubit lifetime T_1 of the qubits is 105.84μs, the median coherence time T_2 is 84.05μs and the median readout error is 0.044. The single qubit gate error varies between 3.054×10^{-4} and 6.929×10^{-4}, while the CNOT error varies between 9.682×10^{-3} and 1.463×10^{-2}, depending on the specific connection. The qubits, which are used for CNOT gate sequence are 0 and 1. ibmq_oslo contains 7 fixed-frequency transmons qubits, with the median fundamental transition frequency of 5.046 GHz and median anharmonicity of −0.3429 GHz. The median qubit lifetime T_1 of the qubits is 128.12μs, the median coherence time T_2 is 58.57μs and the median readout error is 0.0216. The single qubit gate error varies between 1.648×10^{-4} and 6.698×10^{-4}, while the CNOT error varies between 6.471×10^{-3} and 2.067×10^{-2}, depending on the specific connection. The qubits, which are used to run GHZ algorithm, belong to the list [0,1,3,5,4].

Appendix F: Plots of the fidelities of the X and CR gates Lindblad simulations

Here in Fig 7 and Fig. 8 we show the plots of the fidelities obtained from the simulations in Sec. VII. The fidelity is defined as:

$$\mathcal{F}(\rho,\sigma) = \left(\mathrm{Tr}\sqrt{\sigma^{1/2}\rho\sigma^{1/2}}\right)^2. \tag{F1}$$

We notice that when one considers only diagonal density matrices, the fidelity is called Hellinger fidelity and it is related to the Hellinger distance as $\mathcal{F} = (1-\mathcal{H}^2)^2$. The Hellinger fidelity is not a proper mathematical distance,

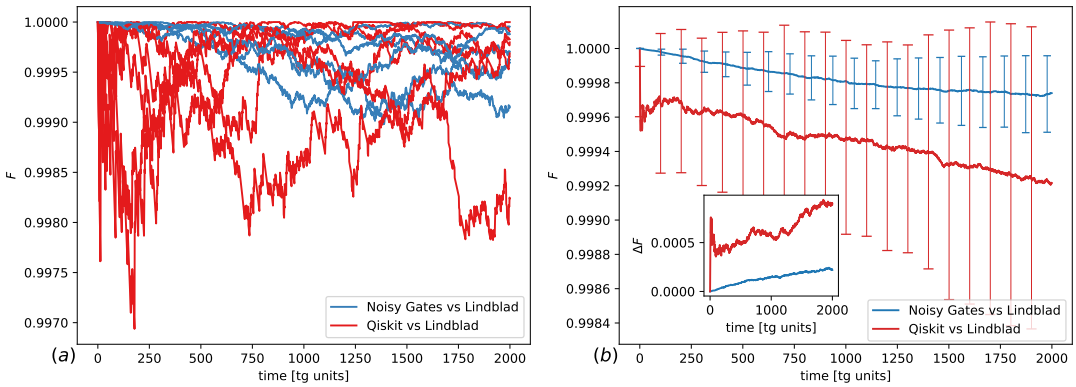

FIG. 7. Fidelities $\mathcal{F}_\sigma^{\rm ng}$, in blue, and $\mathcal{F}_\sigma^{\rm ibm}$, in red, as a function of time, for a repetition of X gates. On panel (a), the fidelities obtained from 100 independent runs of the two methods are pictured (for better readability only five are shown), where each simulation is obtained by averaging over 1000 samples. On panel (b), the means $\bar{\mathcal{F}}_\sigma^{\rm ng}$, $\bar{\mathcal{F}}_\sigma^{\rm ibm}$ of the same simulations and their standard deviations $\Delta\mathcal{F}_\sigma^{\rm ng}$, $\Delta\mathcal{F}_\sigma^{\rm ibm}$ are displayed. The inset shows the standard deviations $\Delta\mathcal{F}_\sigma^{\rm ng}$, $\Delta\mathcal{F}_\sigma^{\rm ibm}$ as functions of time.

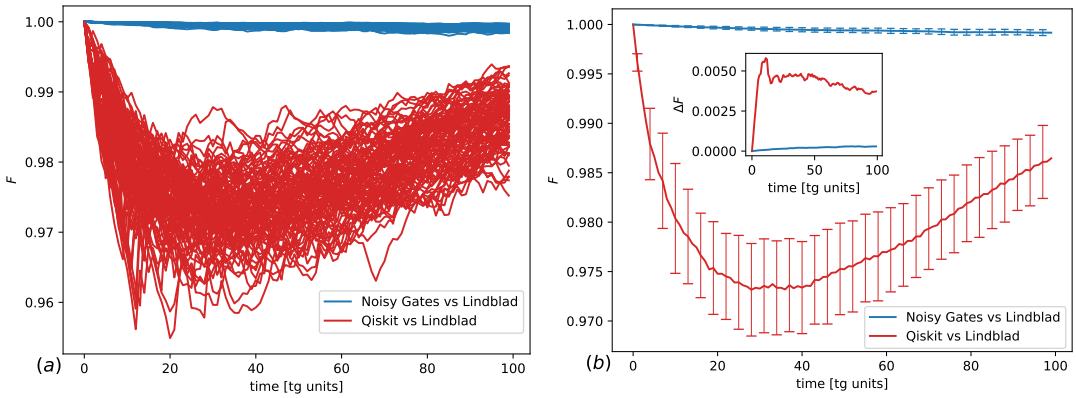

FIG. 8. Fidelities $\mathcal{F}_\sigma^{\rm ng}$, in blue, and $\mathcal{F}_\sigma^{\rm ibm}$, in red, as a function of time, for a repetition of CR gates. Panels (a) and (b) have the same meaning as for Fig.7.

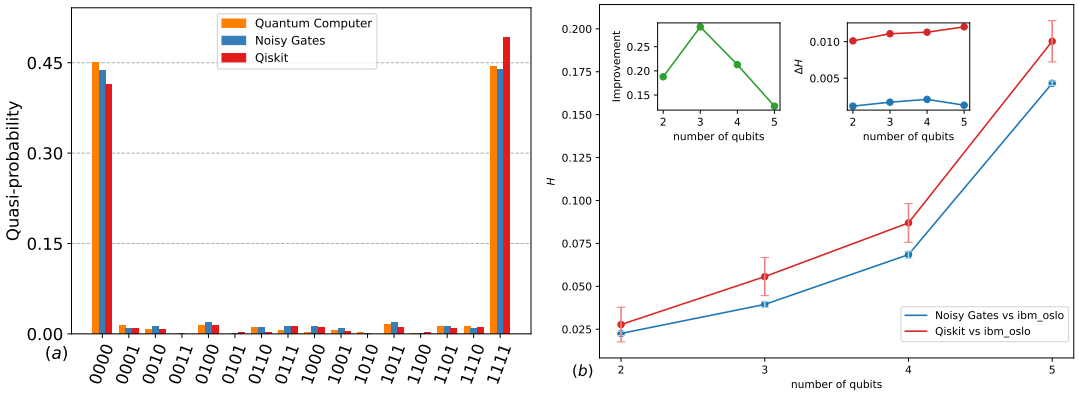

FIG. 9. On panel (a), probabilities histograms for 4 qubits of a single independent simulation of the GHZ algorithm. In orange the results for ibmq_oslo, in blue for the noisy gates and in red for the Qiskit simulator. On panel (b), Hellinger distance for the GHZ algorithm for $n = 2, \ldots, 5$ qubits. Each value is the mean of 100 independent simulations for the noisy gates, in blue, and for the Qiskit simulations, in red. The left inset shows the relative improvement, calculated as $|\bar{\mathcal{H}}_\sigma^{\rm ibm} - \bar{\mathcal{H}}_\sigma^{\rm ng}|/\bar{\mathcal{H}}_\sigma$ibm, while the right inset shows the standard deviations as functions of the number of qubits.

thus in the main text we used the Hellinger distance. As one can see the results are consistent with those in Sec. VII.

Appendix G: GHZ simulations

In this appendix we report the results of the analysis of the GHZ algorithm in order to inspect the performances of the noisy gates approach when trying to reproduce the behaviour of a real quantum computer. We run GHZ for n = 2, ... , 5 on ibmq_oslo and we set as input the state $|0\rangle^{\otimes n}$. Runs on real quantum computer are performed by taking 1000 shots, i.e. measurements. We also run the corresponding classical simulations. We use the same custom noise model defined in Sec. VII in the $QFT^{\dagger}$ case. The resulting probability histograms for 4 qubits of a single independent simulation is reported in Fig. 9. We notice that, as for the $\mathrm{QFT}^{\dagger}$ case, for every n we get $\bar{\mathcal{H}}^{\mathrm{ng}} < \bar{\mathcal{H}}^{\mathrm{ibm}}$ and $\Delta\mathcal{H}^{\mathrm{ng}} < \Delta\mathcal{H}^{\mathrm{ibm}}$, see Fig. 9.

Appendix H: Comparison between relevant quantum computing frameworks on noisy simulations

In the following we report a table with a list of relevant quantum computing frameworks where we verify whether they support noise simulation (NS) and if so wheter they implement the approach descibed in Eq. (1) of Sec. I that we call standard approach (SA).

Company	Name	Ref.	NS	SA
IBM	Qiskit	[64]	Yes	Yes
Rigetti	pyQuil	[23]	Yes	Yes
Quantinuum	t|*ket*⟩	[65]	Yes	Yes
Xanadu	Pennylane	[66]	Yes	Yes
Xanadu	Strawberry Field	[67]	Yes	Yes
Microsoft	Azure Quantum	[68]	No	-
Microsoft	LIQUI|⟩	[69]	Yes	Yes
Google	Cirq	[70]	Yes	Yes
Google	TensorFlow Quantum	[71]	Yes	Yes
Intel	Intel QS	[72]	Yes	Yes
Baidu	Paddle Quantum	[73]	Yes	Yes
Amazon	Braket	[74]	Yes	Yes
-	ProjectQ	[75]	No	-
-	QiBO	[76]	Yes	Yes
-	QCL	[77]	No	-
-	Quipper	[78]	No	-
-	Quirk	[79]	No	-
-	SilQ	[80]	No	-

TABLE I. List of relevant quantum computing frameworks. In the fourth column is specified whether the corresponding framework support noise simulation (NS) while in the fifth column we specify whether the noise simulation is based on the approach described in Eq. (1) of Sec. I that we call standard approach (SA).

[1] J. Chow, O. Dial, and J. Gambetta, IBM Research Blog (2021).
[2] D. Kielpinski, C. Monroe, and D. J. Wineland, Nature **417**, 709 (2002).
[3] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Nature **574**, 505 (2019).
[4] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, *et al.*, Physical review letters **127**, 180501 (2021).
[5] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, *et al.*, Science **370**, 1460 (2020).
[6] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, *et al.*, Reviews of Modern Physics **94**, 015004 (2022).
[7] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, *et al.*, Nature Reviews Physics **3**, 625 (2021).
[8] A. M. Steane, Physical Review A **68**, 042322 (2003).
[9] A. G. Fowler, A. M. Stephens, and P. Groszkowski, Physical Review A **80**, 052312 (2009).
[10] D. A. Lidar and T. A. Brun, *Quantum error correction* (Cambridge university press, 2013).
[11] Google Quantum AI, Nature **614**, 676 (2023).
[12] V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. Brock, A. Ding, L. Frunzio, *et al.*, arXiv preprint arXiv:2211.09116 (2022).
[13] J. Preskill, Quantum **2**, 79 (2018).
[14] H.-P. Breuer, F. Petruccione, *et al.*, *The theory of open quantum systems* (Oxford University Press on Demand, 2002).
[15] M. A. Nielsen and I. L. Chuang, Quantum computing and quantum information (2000).
[16] G. Benenti, G. Casati, D. Rossini, and G. Strini, *Principles of Quantum Computation and Information: A Comprehensive Textbook* (World Scientific, 2019).
[17] Ibm quantum compute resources (2022).
[18] S. L. Adler and A. Bassi, Journal of Physics A: Mathematical and Theoretical **40**, 15083 (2007).
[19] A. Bassi, Physical Review A **67**, 062101 (2003).
[20] S. Maniscalco and F. Petruccione, Physical Review A **73**, 012111 (2006).
[21] W. T. Strunz, L. Diósi, N. Gisin, and T. Yu, Physical Review Letters **83**, 4909 (1999).
[22] J. Gambetta and H. Wiseman, Journal of Optics B: Quantum and Semiclassical Optics **6**, S821 (2004).
[23] pyQuill (2022).
[24] Qiskit notebook (2022).
[25] E. van den Berg, Z. K. Minev, A. Kandala, and K. Temme, arXiv e-prints , arXiv (2022).
[26] K. Temme, S. Bravyi, and J. M. Gambetta, Physical review letters **119**, 180509 (2017).
[27] S. Zhang, Y. Lu, K. Zhang, W. Chen, Y. Li, J.-N. Zhang, and K. Kim, Nature communications **11**, 587 (2020).

[28] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Applied Physics Reviews **6**, 021318 (2019).
[29] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, Journal of Mathematical Physics **17**, 821 (1976).
[30] G. Lindblad, Communications in Mathematical Physics **48**, 119 (1976).
[31] K. Georgopoulos, C. Emary, and P. Zuliani, Physical Review A **104**, 062432 (2021).
[32] A. Ash-Saki, M. Alam, and S. Ghosh, in *Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design* (2020) pp. 25–30.
[33] K. Jacobs and P. L. Knight, Physical review A **57**, 2301 (1998).
[34] M. Caiaffa, A. Smirne, and A. Bassi, Physical Review A **95**, 062101 (2017).
[35] K. Jacobs, *Quantum measurement theory and its applications* (Cambridge University Press, 2014).
[36] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control* (Cambridge university press, 2009).
[37] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, Vol. 3 (springer Berlin, 1985).
[38] A. Bassi and G. Ghirardi, Physics Reports **379**, 257 (2003).
[39] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht, Reviews of Modern Physics **85**, 471 (2013).
[40] L. Arnold, New York (1974).
[41] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Physical Review A **96**, 022330 (2017).
[42] The native single qubit gates chosen by IBM are X and SX, which are rotations around the x−axis obtained by fixing $\phi = 0$ in Eq. (17). Rotations around the z−axis are implemented as *virtual* gates, since they are mimicked by the software and are not associated to a physical action on the device [41].
[43] C. Rigetti and M. Devoret, Physical Review B **81**, 134507 (2010).
[44] Ibm qiskit (2022).
[45] W. R. Inc., Mathematica, Version 13.1, champaign, IL, 2022.
[46] T. Alexander, N. Kanazawa, D. J. Egger, L. Capelluto, C. J. Wood, A. Javadi-Abhari, and D. C. McKay, Quantum Science and Technology **5**, 044006 (2020).
[47] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Physical Review A **76**, 042319 (2007).
[48] C. D. Wilen, S. Abdullah, N. Kurinsky, C. Stanford, L. Cardani, G. d'Imperio, C. Tomei, L. Faoro, L. Ioffe, C. Liu, *et al.*, Nature **594**, 369 (2021).
[49] P. Zhao, K. Linghu, Z. Li, P. Xu, R. Wang, G. Xue, Y. Jin, and H. Yu, PRX Quantum **3**, 020301 (2022).
[50] P. W. Shor, in *Proceedings 35th annual symposium on foundations of computer science* (Ieee, 1994) pp. 124–134.
[51] L. Ruiz-Perez and J. C. Garcia-Escartin, Quantum Information Processing **16**, 1 (2017).
[52] D. M. Greenberger, in *Compendium of quantum physics* (Springer, 2009) pp. 258–263.
[53] E.g. for $n = 9$ around 50 CNOTs are performed on almost every pair of connected qubits while for $n = 18$ more than 100 CNOTs. As one can see in Figs. 2 and 3, when more than 100 CR or CNOT gates are applied, the total execution time is larger than the decoherence times. Thus, the state of the system approaches rapidly the maximally mixed state.
[54] J. Preskill, arXiv preprint arXiv:1207.6131 (2012).
[55] M. Sarovar, T. Proctor, K. Rudinger, K. Young, E. Nielsen, and R. Blume-Kohout, Quantum **4**, 321 (2020).
[56] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien, arXiv preprint arXiv:2210.00921 (2022).
[57] S. Endo, S. C. Benjamin, and Y. Li, Physical Review X **8**, 031027 (2018).
[58] Z. Liang, H. Wang, J. Cheng, Y. Ding, H. Ren, Z. Gao, Z. Hu, D. S. Boning, X. Qian, S. Han, *et al.*, in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2022) pp. 556–565.
[59] M. Ibrahim, H. Mohammadbagherpoor, C. Rios, N. T. Bronn, and G. T. Byrd, arXiv preprint arXiv:2211.00350 (2022).
[60] S. Greenaway, F. Petiziol, H. Zhao, and F. Mintert, arXiv preprint arXiv:2211.16322 (2022).
[61] A. R. R. Carvalho, H. Ball, M. J. Biercuk, M. R. Hush, and F. Thomsen, Physical Review Applied **15**, 064054 (2021).
[62] A. Bassi and D.-A. Deckert, Physical Review A **77**, 032323 (2008).
[63] K. Jacobs, *Stochastic processes for physicists: understanding noisy systems* (Cambridge University Press, 2010).
[64] M. S. Anis, H. Abraham, R. A. AduOffei, G. Agliardi, M. Aharoni, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, S. Anagolum, *et al.*, Qiskit/qiskit (2021).
[65] S. Sivarajah, S. Dilkes, A. Cowtan, W. Simmons, A. Edgington, and R. Duncan, Quantum Science and Technology **6**, 014003 (2020).
[66] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. AkashNarayanan, A. Asadi, *et al.*, arXiv preprint arXiv:1811.04968 (2018).
[67] N. Killoran, J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook, Quantum **3**, 129 (2019).
[68] Azure Quantum (2020).
[69] D. Wecker and K. M. Svore, arXiv preprint arXiv:1402.4467 (2014).
[70] Cirq (2022).
[71] M. Broughton, G. Verdon, T. McCourt, A. J. Martinez, J. H. Yoo, S. V. Isakov, P. Massey, R. Halavati, M. Y. Niu, A. Zlokapa, *et al.*, arXiv preprint arXiv:2003.02989 (2020).
[72] G. G. Guerreschi, J. Hogaboam, F. Baruffa, and N. P. Sawaya, Quantum Science and Technology **5**, 034007 (2020).
[73] Paddle Quantum (2020).
[74] Amazon Braket (2020).
[75] D. S. Steiger, T. Häner, and M. Troyer, Quantum **2**, 49 (2018).
[76] S. Efthymiou, S. Ramos-Calderer, C. Bravo-Prieto, A. Pérez-Salinas, D. García-Martín, A. Garcia-Saez, J. I. Latorre, and S. Carrazza, Quantum Science and Technology **7**, 015018 (2021).
[77] B. Ömer, International Journal of Theoretical Physics **44**, 943 (2005).
[78] A. S. Green, P. L. Lumsdaine, N. J. Ross, P. Selinger,

and B. Valiron, in *Proceedings of the 34th ACM SIGPLAN conference on Programming language design and implementation* (2013) pp. 333–342.
[79] Quirk (2014).
[80] B. Bichsel, M. Baader, T. Gehr, and M. Vechev, in *Proceedings of the 41st ACM SIGPLAN Conference on Programming Language Design and Implementation* (2020) pp. 286–300.