Spectral quantum tomography

Helsen, Jonas; Battistel, Francesco; Terhal, Barbara M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

ARTICLE   OPEN

Spectral quantum tomography

Jonas Helsen [iD][1], Francesco Battistel[1] and Barbara M. Terhal[1,2]

We introduce spectral quantum tomography, a simple method to extract the eigenvalues of a noisy few-qubit gate, represented by a trace-preserving superoperator, in a SPAM-resistant fashion, using low resources in terms of gate sequence length. The eigenvalues provide detailed gate information, supplementary to known gate-quality measures such as the gate fidelity, and can be used as a gate diagnostic tool. We apply our method to one- and two-qubit gates on two different superconducting systems available in the cloud, namely the QuTech Quantum Infinity and the IBM Quantum Experience. We discuss how cross-talk, leakage and non-Markovian errors affect the eigenvalue data.

INTRODUCTION

A central challenge on the path towards large-scale quantum computing is the engineering of high-quality quantum gates. To achieve this goal, many methods that accurately and reliably characterize quantum gates have been developed. Some of these methods are scalable, meaning that they require an effort which scales polynomially in the number of qubits on which the gates act. Scalable protocols, such as randomized benchmarking,[1–8] necessarily give a partial characterization of the gate quality, for example, an average gate fidelity. Other protocols such as robust tomography[9] or gate-set tomography[10,11] trade scalability for a more detailed characterization of the gate. A desirable feature of all the above protocols is that they are resistant to state-preparation and measurement (SPAM) errors. The price of using SPAM-resistant (scalable) methods is that these protocols have significant experimental complexity and/or require assumptions on the underlying hardware to properly interpret their results.

In this work, we present spectral quantum tomography, a simple non-scalable method that extracts spectral information from noisy gates in a SPAM-resistant manner. To process the tomographic data and obtain the spectrum of the noisy gate, we rely on the matrix-pencil technique, a well-known classical signal processing method. This technique has been advocated in ref. [8] in the context of randomized benchmarking and has also been used in ref. [12] for processing data in the algorithm of quantum phase estimation. It has also been used, under the phrase "linear systems identification," in ref. [13] to predict the time evolution of quantum systems. While the matrix pencil technique leads to explicitly useful estimates of eigenvalues and their amplitudes, we note that the same underlying idea is used in the method of "delayed vectors," which has been proposed in ref. [14] to assess the dimensionality of a quantum system from its dynamics. This "delayed vectors" approach has been applied to assess leakage in superconducting devices in ref. [15]

The spectral information of a noisy gate $\mathcal{S}$, which approximates some target unitary U, is given by the eigenvalues of the so-called Pauli transfer matrix representing $\mathcal{S}$. These eigenvalues, which are of the form $\lambda = \exp(-\gamma)\exp(i\phi)$, contain information about the quality of the implemented gate. Intuitively, the parameter γ captures how much the noisy gate deviates from unitarity due to entanglement with an environment, while the angle ϕ can be compared to the rotation angles of the targeted gate U. Hence ϕ gives information about how much one over- or under-rotates. The spectrum of $\mathcal{S}$ can also be related to familiar gate-quality measures such as the average gate fidelity and the unitarity. Moreover, in the case of a noisy process modeled by a Lindblad equation, the spectrum can be easily related to the more familiar notions of relaxation and dephasing times.

The main advantage of spectral quantum tomography is its simplicity, requiring only the (repeated) application of a single noisy gate $\mathcal{S}$, as opposed to the application of a large set of gates as in randomized benchmarking, gate-set tomography, and robust tomography. Naturally, simplicity and low cost come with some drawback, namely, the method does not give information about the eigenvectors of the noisy gate, such as the axis around which one is rotating. However, information about the eigenvectors is intrinsically hard to extract in a SPAM-resistant fashion since SPAM errors can lead to additional rotations.[16] Another feature of spectral quantum tomography is that it can be used to extract signatures of non-Markovianity, namely, the phenomenon where the noisy gate $\mathcal{S}$ depends on the context in which it is applied (e.g., time of application, whether any gates have been applied before it). As we show in this paper, our method can be used to detect various types of non-Markovian effects, such as coherent revivals, parameter drifts, and Gaussian-distributed time-correlated noise. It is also possible to distinguish non-Markovian effects from qubit leakage. For these reasons, we believe that spectral quantum tomography adds a useful new tool to the gate-characterization toolkit. The method could also have future applications in assessing the performance of logical gates in a manner that is free of logical state preparation and measurement errors, see the "Discussion" section.

[1]QuTech, Delft University of Technology, P. O. Box 5046, 2600 GA Delft, the Netherlands and [2]JARA Institute for Quantum Information, Forschungszentrum Juelich, 52425 Juelich, Germany
Correspondence: Jonas Helsen (j.helsen@tudelft.nl)

RESULTS

Eigenvalues of trace-preserving completely positive (TPCP) maps

Take a unitary gate U on a d-dimensional space with $U|\psi_j\rangle = e^{i\phi_j}|\psi_j\rangle$. The corresponding TPCP map $\mathcal{S}_U(\rho) = U\rho U^\dagger$ has one trace-full eigenvector, namely, I with eigenvalue 1, as well as $d^2 - 1$ traceless eigenvectors. In particular, there are $d^2 - d$ traceless eigenvectors of the form $|\psi_j\rangle\langle\psi_l|$ for $j \neq l$ with eigenvalues $\exp(i(\phi_j - \phi_l))$ and $d - 1$ traceless eigenvectors of the form $|\psi_1\rangle\langle\psi_1| - |\psi_j\rangle\langle\psi_j|$ for $j = 2, \ldots, d$ with eigenvalue 1.

For general TPCP maps, it is convenient to use the Pauli transfer matrix formalism. For an n-qubit system ($d = 2^n$), consider the normalized set of Pauli matrices P_μ for $\mu = 0, \ldots, N$ with $N + 1 = 4^n = d^2$, where $P_0 = I/\sqrt{2^n}$ and the normalization is chosen such that $\mathrm{Tr}[P_\mu P_\nu] = \delta_{\mu\nu}$. For a TPCP map S acting on n qubits, the Pauli transfer matrix is then defined as

$$S_{\mu\nu} = \mathrm{Tr}[P_\mu \mathcal{S}(P_\nu)], \quad \mu, \nu = 0, \ldots, N. \tag{1}$$

The form of the Pauli transfer matrix S is[17]

$$\mathcal{S} \leftrightarrow S = \begin{pmatrix} 1 & 0 \\ \mathbf{s} & T^{\mathcal{S}} \end{pmatrix}, \tag{2}$$

where $T^{\mathcal{S}}$ is a real $N \times N$ matrix and $\mathbf{s}$ is a N-dimensional column vector. The 1 and 0s in the top row of the Pauli transfer matrix are due to the fact that $\mathcal{S}$ is trace-preserving. For a unital $\mathcal{S}$ that obeys $\mathcal{S}(I) = I$, the vector $\mathbf{s} = 0$.

A few properties are known of the eigenvalue–eigenvector pairs of S, i.e., the pairs $(\lambda, \mathbf{v})$ with $S\mathbf{v} = \lambda\mathbf{v}$:

- The eigenvalues of S, and thus the eigenvalues of T^S are 1, since the solutions of the equation $\det(S - \lambda I) = 0$ are the solutions of the equation $(1 - \lambda)\det(T^S - \lambda I) = 0$.
- The eigenvalues of S, and thus the eigenvalues of $T^{\mathcal{S}}$, come in complex–conjugate pairs. This is true because $T^{\mathcal{S}}$ is a real matrix.
- The eigenvalues of $T^{\mathcal{S}}$ (or S for that matter) have modulus <1, i.e. $|\lambda| \leq 1$ (see, e.g., Proposition 6.1 in ref. [18]).

If $T^{\mathcal{S}}$ is *diagonalizable* as a matrix, it holds that $T^{\mathcal{S}} = VDV^{-1}$ where D is a diagonal matrix and V a similarity transformation. Generically, $T^{\mathcal{S}}$ will be diagonalizable, in which case there are N eigenvalue–eigenvector pairs for T. A sufficient condition for diagonizability is, for example, that all the eigenvalues of $T^{\mathcal{S}}$ are distinct. In section "Single-qubit case with non-diagonalizable matrix T," we give examples and discuss what it means if $T^{\mathcal{S}}$ is not diagonalizable.

For some simple single-qubit channels, we can explicitly compute the spectrum. For instance, for a single-qubit depolarizing channel with depolarizing probability p, the eigenvalues of the submatrix T^S of the Pauli transfer matrix are $\{1 - p, 1 - p, 1 - p\}$. For a single-qubit amplitude-damping channel with damping rate p, they are $\{\sqrt{1-p}, \sqrt{1-p}, 1-p\}$.[11]

Relation to gate-quality measures. The eigenvalues of the Pauli transfer matrix of a noisy gate $\mathcal{S}$ can be related to several other known measures of gate quality, such as the average gate fidelity $\mathcal{F}(\mathcal{S}, U)$, the gate unitarity $u(\mathcal{S})$ and, for a single qubit ($n = 1$), the gate unitality.

The average gate fidelity is defined as $\mathcal{F}(\mathcal{S}, U) = \int d\phi\langle\phi|U^\dagger\mathcal{S}(|\phi\rangle\langle\phi|)U|\phi\rangle$. This fidelity relates directly to the entanglement fidelity $\mathcal{F}_{\mathrm{ent}}(\mathcal{S}, U)$ via $\mathcal{F} = \frac{\mathcal{F}_{\mathrm{ent}}d + 1}{d + 1}$,[19] where the entanglement fidelity is defined as

$$\mathcal{F}_{\mathrm{ent}}(\mathcal{S}, U) = \mathrm{Tr}[I \otimes U|\Psi\rangle\langle\Psi|I \otimes U^\dagger(I \otimes \mathcal{S})(|\Psi\rangle\langle\Psi|)],$$

where $|\Psi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^d |i, i\rangle$ is a maximally entangled state. Using

that $|\Psi\rangle\langle\Psi| = \frac{1}{d}\sum_{\mu=0}^N P_\mu \otimes P_\mu$ and $UP_\mu U^\dagger = \sum_\kappa T_{\mu\kappa}^{U^\dagger}P_\kappa$, we can write

$$\mathcal{F}_{\mathrm{ent}}(\mathcal{S}, U) = \frac{1}{d^2}\sum_\mu \mathrm{Tr}[UP_\mu U^\dagger\mathcal{S}(P_\mu)] = \frac{1}{d^2}\left(1 + \mathrm{Tr}[T^{U^\dagger}T^{\mathcal{S}}]\right).$$

Thus, for the (entanglement) fidelity of a noisy gate $\mathcal{S}$ with respect to the identity channel $U = I$, one has $\mathcal{F}_{\mathrm{ent}}(\mathcal{S}, I) = \frac{1}{d^2}\left(1 + \sum_i \lambda_i\right)$, implying a direct relation to the spectrum $\{\lambda_i\}$ of $T^{\mathcal{S}}$. A more interesting relation is how the eigenvalues of $T^{\mathcal{S}}$ bound the fidelity with respect to a targeted gate U. In section "Upper bound on the entanglement fidelity with the targeted gate," we prove that the entanglement fidelity can be upper bounded as

$$\mathcal{F}_{\mathrm{ent}}(\mathcal{S}, U) \leq \frac{1}{d^2}\left[1 + (d^2 - 1)\left(\sqrt{1 - \frac{\sum_j |\lambda_j|^2}{d^2 - 1}} + \xi_{\max}\right)\right], \tag{3}$$

where $\xi_{\max} = \frac{1}{d^2 - 1}|\sum_j \lambda_j^{\mathrm{ideal}}\lambda_j^*|$ with $\lambda_j^{\mathrm{ideal}}$ the eigenvalues of T^U with U the targeted unitary, ordered such that the sum $|\sum_j \lambda_j^{\mathrm{ideal}}\lambda_j^*|$ is maximal.

This upper bound is not particularly tight, but for the case of a single qubit we can make a much stronger numerical statement, see section "Upper bound on the entanglement fidelity with the targeted gate".

Another measure of gate quality, namely, the unitarity or the coherence of a channel[5] on a d-dimensional system, is defined as

$$u(\mathcal{S}) = \frac{d}{d - 1}\int d\phi\, \mathrm{Tr}\left[[\mathcal{S}'(|\phi\rangle\langle\phi|)]^\dagger\mathcal{S}'(|\phi\rangle\langle\phi|)\right], \tag{4}$$

where $\mathcal{S}'(\rho) := \mathcal{S}(\rho) - \mathrm{Tr}[\mathcal{S}(\rho)]I/\sqrt{d}$. A more convenient but equivalent definition is

$$u(\mathcal{S}) = \frac{1}{d^2 - 1}\mathrm{Tr}[T^{\mathcal{S}\dagger}T^{\mathcal{S}}] = \frac{1}{d^2 - 1}\sum_i \sigma_i(T^{\mathcal{S}})^2, \tag{5}$$

where $\{\sigma_i\}$ are the singular values of the matrix $T^{\mathcal{S}}$. The unitarity captures how close the channel is to a unitary gate. A lower bound on the unitarity is given by Proposition 2 in ref. [16]:

$$u(\mathcal{S}) \geq \frac{1 + \sum_{i=1}^{d^2-1}|\lambda_i|^2 - d}{d(d - 1)}, \tag{6}$$

where $\{\lambda_i\}$ are the eigenvalues of $T^{\mathcal{S}}$. For a single qubit, an upper bound on the unitarity can also be given in terms of a non-convex optimization problem, see section "Upper bound on the entanglement fidelity with the targeted gate".

The unitality of a TPCP map is defined as $1 - \|\mathbf{s}\|^2$ with $\mathbf{s}$ in Eq. (2). Specifically, for single-qubit channels, one can derive the bound[16]

$$\|\mathbf{s}\|^2 \leq 1 - |\lambda_1|^2 - |\lambda_2|^2 - |\lambda_3|^2 + 2\lambda_1\lambda_2\lambda_3. \tag{7}$$

Relation to relaxation and dephasing times. We consider the eigenvalues of a superoperator induced by a simple Lindblad equation modeling relaxation and decoherence of a driven qubit, as an example. We have a Lindblad equation with time-independent Lindbladian $\mathcal{L}$:

$$\dot{\rho} = \mathcal{L}(\rho). \tag{8}$$

The formal solution of Eq. (8) is given by $\rho(t) = e^{t\mathcal{L}}(\rho(t = 0))$, where $e^{t\mathcal{L}}$ is a TPCP map for every t. We are interested in the total evolution after a certain gate time τ and set $\mathcal{S}_\tau = e^{\tau\mathcal{L}}$. We assume a simple model in which a qubit evolves according to a Hamiltonian $H = (h_x X + h_y Y + h_z Z)/2$ and is subject to relaxation

and pure dephasing processes, according to the Lindbladian:

$$\mathcal{L}(\rho) = -i[H, \rho] + \frac{1}{T_1}\left(\sigma_-\rho\sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho\}\right) + \frac{1}{2T_\phi}(Z\rho Z - \rho).$$

We define the relaxation, respectively, dephasing rates $\Gamma_1 = 1/T_1$ and $\Gamma_2 = 1/T_2 = 1/(2T_1) + 1/T_\phi$. The Pauli transfer matrix $L^{\mathcal{L}}$ of $\mathcal{L}$ then takes the form

$$L^{\mathcal{L}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -\Gamma_2 & h_z & h_y \\ 0 & -h_z & -\Gamma_2 & h_x \\ \Gamma_1 & -h_y & -h_x & -\Gamma_1 \end{pmatrix}. \tag{9}$$

We will denote the eigenvalues of $L^{\mathcal{L}}$ by Ω_j for $j \in \{0, \ldots, 3\}$ and the eigenvalues of $\mathcal{S}_\tau$ by λ_j for $j \in \{0, \ldots, 3\}$. As expected, $\Omega_0 = 0$ implying that $\lambda_0 = e^0 = 1$ is an eigenvalue of $\mathcal{S}_\tau$. The other three eigenvalues of $L^{\mathcal{L}}$ can be found from the 3×3 submatrix in the lower-right corner. Here we consider some simple cases.

Case 1: $h_x = h_y = h_z = 0$. In this case, for $j = 1, 2, 3$ the three eigenvalues of $\mathcal{L}$ and $\mathcal{S}_\tau$ are clearly

$$\Omega_j \in \{-\Gamma_2, -\Gamma_2, -\Gamma_1\},$$
$$\lambda_j \in \{e^{-\Gamma_2\tau}, e^{-\Gamma_2\tau}, e^{-\Gamma_1\tau}\},$$

thus relating directly to the relaxation and dephasing rates.

Case 2: $h_x = h_y = 0$. In this case, we have

$$\Omega_j \in \{-\Gamma_2 + ih_z, -\Gamma_2 - ih_z, -\Gamma_1\},$$
$$\lambda_j \in \{e^{-\Gamma_2\tau}e^{ih_z\tau}, e^{-\Gamma_2\tau}e^{-ih_z\tau}, e^{-\Gamma_1\tau}\},$$

where we have separated the decaying part of the λ_j (corresponding to the real part of the Ω_j) and their phases (corresponding to the imaginary part). If we work in the rotating frame of the qubit, h_z can be understood as an over-rotation along the z axis, which would appear in the spectrum as an extra phase imparted to two of the eigenvalues. Again we see that the decaying part of the eigenvalues directly relates to the relaxation and dephasing rates.

Case 3: $h_y = h_z = 0$. This case shows that over-rotations can also modify the decay strength of the eigenvalues. We analyze the eigenvalues as a function of h_x. From $L^{\mathcal{L}}$ in Eq. (9), we see that $\Omega_1(h_x) = -\Gamma_2$ for all h_x. For the other eigenvalues, we have

$$\Omega_{2,3}(h_x) = -\frac{1}{2}\left(\Gamma_1 + \Gamma_2 \pm \sqrt{(\Gamma_1 - \Gamma_2)^2 - 4h_x^2}\right). \tag{10}$$

We see that if $|h_x| < |\Gamma_1 - \Gamma_2|/2 \equiv h_x^{cr}$, only the moduli of λ_2 and λ_3 are affected as compared to Case 1, in other words, λ_2 and λ_3 only decay with no extra phases. On the contrary, the phases of these eigenvalues become non-zero when the driving is sufficiently strong: $|h_x| > h_x^{cr}$. It implies that, if we look at the dynamics induced by the Lindblad equation, real oscillations, not only decay, will be present as a function of τ. Hence, these two scenarios represent, respectively, the overdamped ($|h_x| < h_x^{cr}$) and underdamped regime ($|h_x| > h_x^{cr}$), similar to the dynamics of a vacuum-damped qubit-oscillator system, see, e.g., ref. [20] At $|h_x| = h_x^{cr}$, the system is critically damped and $L^{\mathcal{L}}$ does not have four linearly independent eigenvectors, meaning that the Pauli transfer matrix of $\mathcal{S}_\tau$ is not diagonalizable. In this case, the dynamics also has a linear dependence on t besides the exponential decay with t, see the discussion in section "Single-qubit case with non-diagonalizable matrix T".

Spectral tomography

In this section, we describe the spectral tomography method, which estimates the eigenvalues of $\mathcal{S}$, where $\mathcal{S}$ is a TPCP implementation of a targeted unitary gate.

We model state-preparation errors as a perfect preparation step followed by an unknown TPCP map $\mathcal{N}_{prep}$. Similarly, measurement errors are modeled by a perfect measurement preceded by an unknown TPCP map $\mathcal{N}_{meas}$. We assume that, when we apply the targeted gate k times, an accurate model of the resulting noisy dynamics is $\mathcal{S}^k$. The spectral tomography method can be applied without this assumption but the interpretation of its results is more difficult, see section "Leakage and non-Markovian noise" for a discussion. The method works by constructing the following signal function, for $k = 0, 1, \ldots, K$ for some fixed K:

$$g(k) = \sum_{\mu=1}^{N} \text{Tr}\left[P_\mu \mathcal{N}_{meas} \circ \mathcal{S}^k \circ \mathcal{N}_{prep}(P_\mu)\right]. \tag{11}$$

Gathering the data to estimate $g(k)$ requires (1) picking a traceless n-qubit Pauli P_μ, (2) preparing an n-qubit input state in one of the 2^n basis states corresponding to this chosen Pauli, (3) applying the gate k times and measuring in the same chosen Pauli basis, and (4) repeating (1–3) over different Paulis, basis states, and experiments to get good statistics. As in standard process tomography,[21] one takes linear combinations of the estimated probabilities for the outcomes to construct an estimator of a Pauli operator on a Pauli input. This gives an estimate of $g(k)$ for a fixed k. Repeating this process for $k \in \{0, \ldots, K\}$, we reconstruct the entire signal function. In section "Resources to relaxation and dephasing times," we discuss the cost of doing these experiments as compared to randomized benchmarking.

Let us now examine how $g(k)$ depends on the eigenvalues of the matrix T. When there are no SPAM errors, that is, $\mathcal{N}_{meas}$ and $\mathcal{N}_{prep}$ are identity channels, we have

$$g^{\text{NO SPAM}}(k) = \sum_{\mu=1}^{N} (T^k)_{\mu\mu} = \text{Tr}[T^k] = \sum_{j=1}^{N} \lambda_j^k, \tag{12}$$

where $\{\lambda_j\}$ are the eigenvalues of T. The last step in this equality follows directly when T is diagonalizable, but it can alternatively be proved using the so-called Schur triangular form of T (we give this proof in section "Single-qubit case with non-diagonalizable matrix T").

When $\mathcal{N}_{meas}$ and $\mathcal{N}_{prep}$ are not identity channels, we have

$$g(k) = \text{Tr}\left[T_{meas}T^kT_{prep}\right] = \text{Tr}\left[A_{SPAM}D^k\right] = \sum_{j=1}^{N} A_j\lambda_j^k, \tag{13}$$

where T_{meas} and T_{prep} are, respectively, the T-submatrices of the Pauli transfer matrix of $\mathcal{N}_{meas}$ and $\mathcal{N}_{prep}$. Here we assume that $T = VDV^{-1}$ is diagonalizable and the matrix $A_{SPAM} = V^{-1}T_{prep}T_{meas}V$ captures the SPAM errors. One may expect that A_{SPAM} is close to the identity matrix in the typical case of low SPAM errors, in particular one may expect that $A_j \neq 0$ for all j so that all eigenvalues of T are present in the signal $g(k)$.

In principle, one could take more tomographic data and consider a full matrix-valued signal $c_{\mu\nu}(k) = \text{Tr}[P_\mu\mathcal{N}_{meas} \circ \mathcal{S}^k \circ \mathcal{N}_{prep}(P_\nu)]$ instead of only Eq. (11). This requires doing many more experiments and there is no clear advantage in terms of the ability to determine the spectrum.

Signal analysis or matrix-pencil method for extracting eigenvalues. In this section, we review the classical signal-processing method that reconstructs, from the (noisy) signal $g(k) = \sum_{j=1}^{N} A_j\lambda_j^k$ for $k = 0, \ldots, K$, an estimate for the eigenvalues λ_j and the amplitudes A_j. Note that we have $g(k) \in \mathbb{R}$ due to Eq. (11). Not surprisingly, this signal-processing method has been employed and reinvented in a variety of scientific fields. We implement the so-called ESPRIT analysis described in ref. [22] but see also ref. [23]. In the context of spectral tomography, we know that the signal $g(k)$ will in principle contain N eigenvalues (which are possibly degenerate). However,
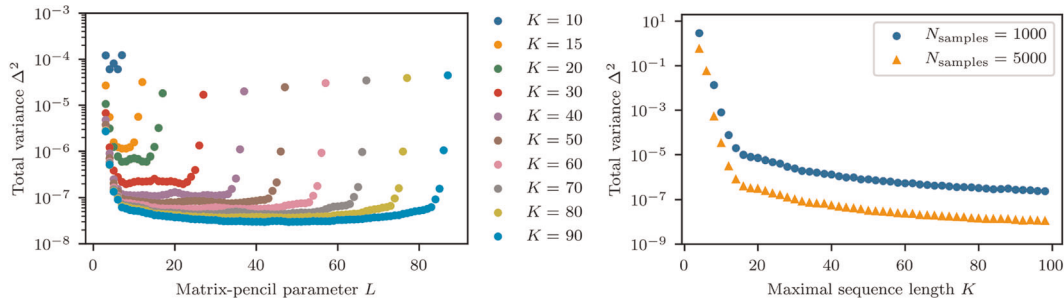
Fig. 1 Preliminary study of the numerical accuracy of the matrix-pencil method as a function of L, K and N_{samples}. (Left) We use the matrix-pencil method with different Ls and Ks to estimate the eigenvalues of a random single-qubit channel, for $N_{\text{samples}} = 1000$. On the vertical axis, we give the variance in the estimate of the eigenvalues: $\Delta^2 = \frac{1}{3}(\sum_{j=1}^{N=3} |\lambda_j - \lambda_j^{\text{est}}|^2)$. We see that, as long as the matrix-pencil parameter L is chosen away from 0 or K, the accuracy of the reconstructed signal is nearly independent of L. Furthermore, we see that higher K's can achieve a lower Δ^2. (Right) We generate a random single-qubit channel and set $L = K/2$. We plot Δ^2 as a function of K for two different values of $N_{\text{samples}} = 1000$ and $N_{\text{samples}} = 5000$, showing how a larger N_{samples} suppresses the total variance. We see that for constant N_{samples} the accuracy of the method increases rapidly at first when K is increased, but it increases more slowly if K is already large. This can be explained by the fact that the signal decreases exponentially in K and so data points for large K have much lower signal-to-noise ratio. For both figures, random channels were generated using QuTip's random TPCP map functionality, and measurement noise was approximated by additive Gaussian noise with standard deviation equal to $1/\sqrt{N_{\text{samples}}}$

we can vary the number of eigenvalues we use to fit the signal to see whether a different choice than N gives a significantly better fit. This is relevant in particular when the implemented gate contains leakage or non-Markovian dynamics, see section "Leakage and non-Markovian noise".

We require at least $K \geq 2N - 2$ in order to determine the eigenvalues accurately. This implies that, for a single-qubit gate with $N = 3$, we need at least $K = 4$, and for a two-qubit gate with $N = 15$, we need at least $K = 28$. However, the signal $g(k)$ has sampling noise due to a bounded N_{samples} and in practice it is good to choose K larger than strictly necessary to make the reconstruction more robust against noise. We study the effect of varying K in Fig. 1 (left panel).

The method goes as follows and relies on picking a so-called pencil parameter L.

Let us assume for now that each $g(k)$ is learned without sampling noise. One constructs a $(K - L + 1) \times (L + 1)$-dimensional data matrix Y as

$$
Y = \begin{pmatrix}
g(0) & g(1) & \dots & g(L) \\
g(1) & g(2) & \dots & g(L+1) \\
g(2) & \vdots & & \vdots \\
\vdots & & & \vdots \\
g(K-L) & \dots & \dots & g(K)
\end{pmatrix}
$$

$$
= \sum_{j=1}^{N} A_j \begin{pmatrix}
1 & \lambda_j & \dots & \lambda_j^L \\
\lambda_j & \lambda_j^2 & \dots & \lambda_j^{L+1} \\
\lambda_j^2 & \vdots & & \vdots \\
\vdots & & & \vdots \\
\lambda_j^{K-L} & \dots & \dots & \lambda_j^K
\end{pmatrix}.
$$

(14)

Note that $\text{rank}(Y) \leq N$ since Y is a sum of at most N rank-1 matrices when there are N eigenvalues. Consider two submatrices of Y: the matrix G_0 is obtained from Y by deleting the last column of Y, while the matrix G_1 is obtained by deleting the first column of Y. When $L = \frac{K}{2}$, the matrices G_0 and G_1 are square matrices of dimension $M = \frac{K}{2} + 1$. For this choice of L, the smallest value of K so that $M = N$ is $2N - 2$. We seek a time-shift matrix $\mathfrak{T}$ such that $\mathfrak{T}G_0 = G_1$. When $M \geq N$, there certainly exists a matrix $\mathfrak{T}$ such that

for all $j \in \{1, \dots, N\}$:

$$
\mathfrak{T}\begin{pmatrix} 1 \\ \lambda_j \\ \vdots \\ \lambda_j^M \end{pmatrix} = \lambda_j \begin{pmatrix} 1 \\ \lambda_j \\ \vdots \\ \lambda_j^M \end{pmatrix}.
$$

(15)

Furthermore, if G_0^{-1} exists, which is the case when $\text{rank}(G_0) = M$, this matrix $\mathfrak{T}$ will be uniquely given as $G_1 G_0^{-1}$. Hence, in this case there is a unique matrix $\mathfrak{T}$, obtained by constructing $G_1 G_0^{-1}$ from the data, which is guaranteed to have $\{\lambda_j\}$ as eigenvalues. When the pencil parameter $L > \frac{K}{2}$, one needs to ensure that there are at least N rows of the matrix Y: if not, $\mathfrak{T}$ would be of dimension $< N$, not giving N eigenvalues. This implies $K \geq N + L - 1$.

The general method for a non-square Y, which includes an additional sampling-noise reduction step, then goes as follows. The choice for N in the procedure can be varied from its minimal value equal to $d^2 - 1$ to a larger value, depending on a goodness-of-fit.

1. Construct a singular-value decomposition of the matrix Y, i.e., $Y = R_1 \Sigma R_2^T$ and replace the diagonal matrix Σ by a diagonal matrix Σ_{clean} with only the largest N singular values. Let $Y_{\text{clean}} = R_1 \Sigma_{\text{clean}} R_2^T$. This step is to reduce sampling noise.
2. Take the submatrices G_0 and G_1 of Y_{clean}.
3. Compute $\mathfrak{T} = G_1 G_0^+$, where G_0^+ is the Moore–Penrose pseudo-inverse of the matrix G_0 so that $\mathfrak{T}$ is a matrix with at most N non-zero eigenvalues.
4. Compute the eigenvalues of $\mathfrak{T}$: these will be the estimates λ_j^{est} of λ_j for all $j \in \{1, \dots, N\}$. Formally, the linear matrix pencil is $G_0 - \lambda G_1$ and the eigenvalues of this matrix pencil, i.e., the values where $\det(G_0 - \lambda G_1) = 0$, are the λ_j^{est}.

We have first applied this method on the signal $g(k)$ of a randomly chosen single-qubit channel: by varying K and L we want to understand the role of the matrix-pencil parameter L and the choice for a larger K. The results are shown in Fig. 1 (left panel). Note that the chosen Ks are quite far above the bound $K \geq N + L - 1$ to effectively suppress sampling noise. For each K, there is a flat region in L where Δ^2 is roughly constant. In the remainder, we will choose $L = K/2$, putting ourselves in the middle of this region. Figure 1 (right panel) shows how increasing N_{samples} lowers the total variance of the estimated eigenvalues.
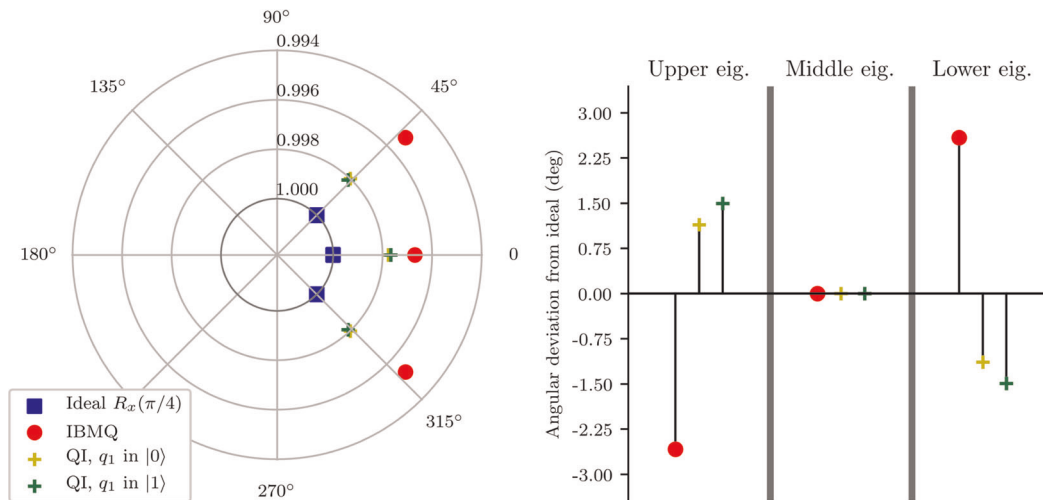
Fig. 2 (Left) Spectral footprints for single-qubit $R_x(\pi/4)$ gates on the *ibmqx4* (IBMQ) and the Quantum Infinity (QI) chips at $K = 50$, $L = 30$, and $N_{samples} = 8192$. The modulus of the eigenvalues is plotted in the radial direction, and in particular, it decreases from the center to the outside and it is equal to 1 on the (most inner) black circumference. The angular coordinate corresponds to the phase of the eigenvalues. (Right) Precise value of the deviation of the phases of the three eigenvalues from the ideal ones

An additional processing step is the determination of the (complex) amplitudes $\{A_j\}$. Viewing $g(k)$ as a set of $K + 1$ inner products between the vector $(A_1, ..., A_N)$ and the linearly-independent vectors $(\lambda_1^k, \lambda_2^k, ..., \lambda_N^k)$, it is clear that, given perfect knowledge of $g(k)$, the $\{A_j\}$ are uniquely determined when $K + 1 \geq N$. Since $g(k)$ is known with sampling noise, the $\{A_j\}$ can be found by solving the least-squares minimization problem $\min_{A_j} \sum_k |g(k) - \sum_j A_j(\lambda_j^{est})^k|^2$. The optimal values in this minimization A_j^{est} and λ_j^{est} together form the reconstructed signal $g^{est}(k)$ and the error is given by

$$\epsilon_N^{rms} = \left(\frac{1}{K+1}\sum_{k=0}^{K}|g(k) - g_N^{est}(k)|^2\right)^{1/2}. \quad (16)$$

Resources. It is interesting to consider the amount of experiments that must be done to perform spectral quantum tomography. One must estimate the function $g(k)$ defined in Eq. (13). This reconstruction process requires running $2^n \times N \times (K+1)$ different experiments and repeating each experiment $N_{samples}$ times. For a single-qubit gate, we need $6(K+1)$. Note that, while the number of experiments scales exponentially with qubit number (not surprising for a tomographic protocol), the number of experiments needed for performing spectral tomography on single- and two-qubit gates is comparable to the number of experiments that must be performed in randomized benchmarking on one or two qubits (which provides only average gate information). In randomized benchmarking, one must sample M random sequences for each sequence length $k \in [0:K]$, yielding $M \times (K+1)$ experiments. This M is independent of the number of qubits.[24] In experiments, M is often chosen between $M \approx 40$[25,26] at the low end and $M \approx 150$ at the higher end.[27] Values of K reported in randomized benchmarking experiments are also comparable to (or even higher than, see ref. [25] where $K \approx 300$ is considered) the values of K used for single- and two-qubit spectral tomography (section "Spectral tomography on two superconducting chips").

Spectral tomography on two superconducting chips
We have executed the spectral tomography method on a single-qubit $\pi/4$ rotation around the x axis: $R_x(\pi/4) = \exp(-i\pi X/8)$. For this gate, the ideal matrix $T^{R_x(\pi/4)}$ should have eigenvalues 1, $\exp(i\pi/4)$, and $\exp(-i\pi/4)$. We execute this gate on two different systems

available in the cloud: the two-qubit Quantum Infinity provided by the DiCarlo group at QuTech (for internal QuTech use) and the *ibmqx4* (IBM Q5 Tenerife) available at https://quantumexperience. ng.bluemix.net/qx/editor. The results of this experiment are shown in Fig. 2 (left panel) in a polar plot, which we refer to as the "spectral footprint" of the gate. For clarity, in Fig. 2 (right panel) we have also plotted the phase deviation from ideal for the implemented gates.

On the two-qubit (q_0 and q_1) Quantum Infinity chip, we perform the single-qubit gate experiment on q_0 twice to study cross-talk: in one case, the undriven qubit q_1 on the chip is in state $|0\rangle$, in the other case q_1 is in state $|1\rangle$. Since the residual off-resonant qubit coupling, mediated through a common resonator, is non-zero, we observe a small difference between these two scenarios, see Fig. 2. For the Quantum Infinity chip, when q_1 is $|0\rangle$ we estimate $\lambda_j^{est} \in \{0.691 + 0.719i, 0.691 - 0.719i, 0.997\}$, while $\lambda_j^{est} \in \{0.687 + 0.7239i, 0.687 - 0.724i, 0.998\}$ when q_1 is $|1\rangle$. Using the single-qubit fidelity bound given in section "Upper bound on the entanglement fidelity with the targeted gate", we can compute that the fidelity with respect to the targeted gate $R_x(\pi/4)$ can be no more than 0.999 regardless of the state of q_1. We can also compute upper and lower bounds on the unitarity (see sections "Eigenvalues of Trace-Preserving Completely Positive (TPCP) maps" and "Upper bound on the entanglement fidelity with the targeted gate") which yields $0.994 \leq u \leq 0.996$ regardless of the state of q_1.

Regarding the *ibmqx4* chip, the data are taken when all other qubits are in state $|0\rangle$. The reconstructed eigenvalues $\lambda_j^{est} \in \{0.735 + 0.671i, 0.735 - 0.671i, 0.996\}$ turn out to be lower in magnitude. From these numbers, we can conclude that the fidelity to the target gate is no higher than 0.998 and the unitarity lies between 0.988 and 0.991.

For all these numbers, a two-way 95% confidence interval (for both real and imaginary parts) deviates by <0.005 from the quoted values. The confidence intervals are obtained through a Wild resampling bootstrap with Gaussian kernel.[28]

We have considered whether the data can be better fitted with $>N = 3$ eigenvalues. For each experiment, we fit the data using N eigenvalues with $N \in \{4, ..., 15\}$ and we test whether there is a significant increase in goodness-of-fit using a standard F-test [ref. [29], section 2.1.5]. For no experiment and value of N does the resultant p value drop <0.05, leading us to conclude that increasing the number of eigenvalues does not significantly increase the accuracy of the fit.

We have also executed a two-qubit CNOT gate on *ibmqx4* (Fig. 3). The T matrix of the ideal CNOT gate has 15 eigenvalues and a very degenerate spectrum: 6 eigenvalues are equal to -1 and 9 eigenvalues are equal to 1, but our data, taking $K = 50$, show that a best fit is obtained using 4 instead of 2 eigenvalues. Using an F-test shows that the goodness-of-fit is significantly improved using 4 eigenvalues rather than 2 or 3, whereas adding more eigenvalues beyond 4 does not significantly improve the goodness-of-fit ($p > 0.05$). We have not tried using larger K (which may lead to a resolution of more eigenvalues) since this would break the requirement that our experiments are executed as a single job performed in a short amount of time on the IBM Quantum Experience. The eigenvalues are $\lambda_j^{\text{est}} \in \{0.939 + 0.059i, 0.938 - 0.059i, -0.961 + 0.067i, -0.961 - 0.067i\}$, all with a 95% confidence interval $< \pm 3 \times 10^{-3}$ for both real and imaginary parts. It is important to note that these 4 eigenvalues, coming in 2 complex–conjugate pairs, *cannot be the spectrum* of a two-qubit TPCP map $\mathcal{S}$, for the following reasons. As observed in section "Eigenvalues of Trace-Preserving Completely Positive (TPCP) maps," the submatrix $T^{\mathcal{S}}$ of the Pauli transfer matrix of $\mathcal{S}$ is a real matrix of odd ($4^2 - 1 = 15$) dimension. Since any complex eigenvalues of a real matrix come in conjugate pairs, $T^{\mathcal{S}}$ must have at least one real eigenvalue. Moreover, the data cannot be explained by allowing for leakage, as any eigenvalues associated with a small amount of leakage must have small associated amplitude, as we discuss in section "Leakage and non-Markovian noise". This is not the case for the eigenvalues plotted in Fig. 3 as all their amplitudes have comparable magnitude $A^{\text{est}} \in \{3.34 - 1.70i, 3.34 + 1.70i, 1.57 + 0.91i, 1.57 - 0.91i\}$. In section "Frame mismatch accumulation," we propose a simple model based on a frame mismatch accumulation that qualitatively reproduces these eigenvalues. This model is not stochastic but coherent, and it violates the assumption that the applied CNOT gate can be fully modeled as a TPCP map. A possible physical mechanism producing a frame mismatch accumulation can be a drift in an experimental parameter.

We do not compute bounds on the fidelity or unitarity of the CNOT gate since the bounds in section "Relation to gate-quality measures" do not apply when the evolution is non-Markovian.

Leakage and non-Markovian noise

In this section, we consider how spectral tomography behaves under error models that violate the assumptions that go into Eq. (13).

Three common mechanisms for gate inaccuracy are (1) crosstalk, meaning the gate depends on or affects the state of other "spectator" qubits; (2) leakage, meaning that the dynamics of the gate acts outside of the computational qubit subspace and (3) non-Markovian dynamics, meaning that the assumption that k applications of the noisy gate are equal to $\mathcal{S}^k$ for some TPCP map $\mathcal{S}$ is incorrect. Characterizing gates with respect to these features is important for assessing their use in multi-gate/multi-qubit devices for the purpose of quantum error correction or plainly reliable quantum computing.[4,30]

One can see that all three scenarios are due to the dynamics taking place in a larger Hilbert space than the targeted computational qubit space. In the case of leakage, the larger space is an extension of the computational space, while in the other cases the larger space is the tensor product of the computational space with the state space of spectator qubits (1), as explored in section "Spectral tomography on two superconducting chips," or other quantum or classical degrees of freedom in the environment (3).

Leakage. Let us consider how gate leakage affects the signal $g(k)$, making the analysis for one or two *qutrits*. One can choose an operator basis for the qutrit space such as the basis of the 8
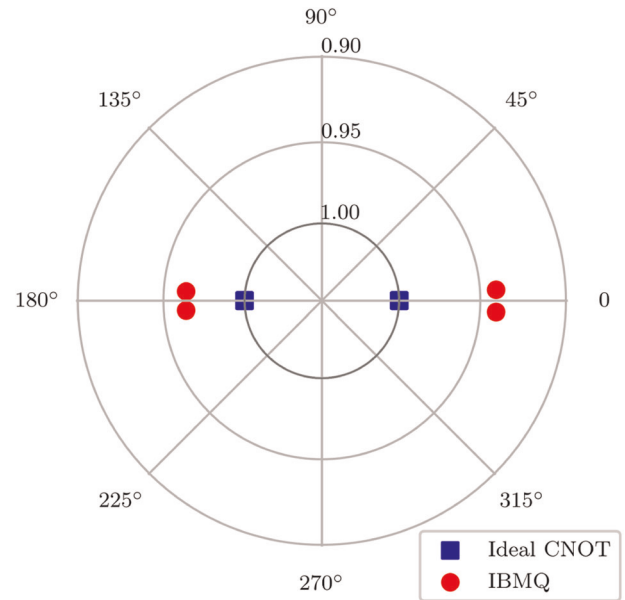
Fig. 3 Spectral footprint of the CNOT gate for $K = 50$ and $N_{\text{samples}} = 8192$. Even though the CNOT gate has only two (degenerate) eigenvalues, we find that the spectrum of the noisy gate can be best described using four distinct eigenvalues. The fact that none of them are real suggests that the data cannot be due to the repeated execution of the same noisy gate. In section "Frame mismatch accumulation," we propose a simple coherent non-Markovian model that offers a possible mechanism for the absence of real eigenvalues

traceless (normalized) Gell–Mann matrices σ_μ^{GM} for $\mu = 1, \ldots, 8$, together with the normalized identity $\sigma_0^{\text{GM}} = \frac{1}{\sqrt{3}} I_3$. For a single qutrit, we can consider the "Pauli" transfer matrix in this Gell–Mann basis, i.e., $S_{\mu\nu}^{\text{GM}} = \text{Tr}[\sigma_\mu^{\text{GM}} \mathcal{S}(\sigma_\nu^{\text{GM}})]$ and its submatrix T^{GM}.

For a single qutrit, the signal $g^{\text{NO SPAM}}(k)$ in Eq. (12) then equals $\text{Tr}_{\text{comp}}[(T^{\text{GM}})^k]$ where $\text{Tr}_{\text{comp}}[A]$ represents the trace over a 3×3 submatrix of A, corresponding to the Gell–Mann matrices, which act like X, Y, and Z in the two-dimensional computational space. In other words, we can see the matrix T^{GM} as being composed of blocks:

$$T^{\text{GM}} = \begin{pmatrix} T_{\text{comp}} & T_{\text{seep}} \\ T_{\text{leak}} & T_{\text{beyond}} \end{pmatrix}, \tag{17}$$

where the upper-left block is the submatrix whose trace we take in $g^{\text{NO SPAM}}(k)$. In the absence of other noise sources, T^{GM} corresponds to the evolution of a unitary gate and (assuming it is diagonalizable) it can be diagonalized by a rotation V as $T^{\text{GM}} = VDV^{-1}$, where D is a diagonal matrix with all the eigenvalues $\{\lambda_j\}$. If we assume that leakage is low, meaning that T_{leak} and T_{seep} have small norm of $O(\epsilon)$, then at lowest order in ϵ the diagonalizing transformation V will be block-diagonal, i.e., $V \approx V_{\text{comp}} \oplus V_{\text{beyond}}$. This means that $g^{\text{NO SPAM}}(k) = \text{Tr}_{\text{comp}}\left[(T^{\text{GM}})^k\right] = \text{Tr}_{\text{comp}}\left[VD^kV^{-1}\right] \approx \sum_{j=1}^{3} \lambda_j^k + O(\epsilon)$. Thus, at lowest order, the signal will have large amplitude on three relevant eigenvalues of the spectrum of T^{GM} and these eigenvalues could have been perturbatively shifted from their ideal location by low leakage. If the leakage is stronger, we can more generally write

$$g^{\text{NO SPAM,LEAK}}(k) = \sum_{j=1}^{8} \tilde{A}_j \lambda_j^k, \quad \tilde{A}_j = \langle \sigma_j | V^{-1} \Pi_{\text{comp}} V | \sigma_j \rangle. \tag{18}$$

Here $|\sigma_j\rangle$ is a vector notation for one of the eight Gell–Mann matrices σ_j and Π_{comp} is the projector onto the basis spanned by the three Gell–Mann matrices, which are the Paulis in the computational space. From this expression, we see that the effect
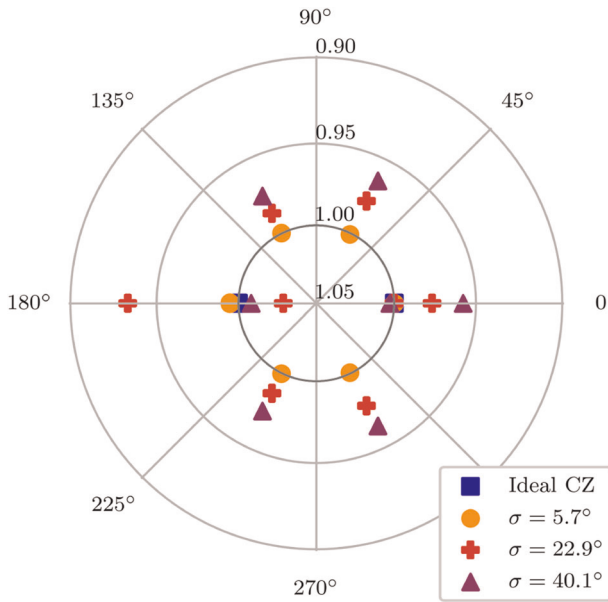
Fig. 4 Spectral footprint of a simulated CZ gate affected by non-Markovian noise quantified by σ, see Eq. (19). For each σ, we use an F-test (p value 0.01) to find the number of eigenvalues that best fit the simulated $g^{NO\ SPAM}(k)$ with $K = 50$. We find, respectively 7, 12, and 11 eigenvalues for $\sigma = 5.7°$, $22.9°$, and $40.1°$ (here we show only the eigenvalues with modulus >0.9). We observe eigenvalues with modulus >1 if σ is sufficiently large. These results are qualitatively stable if we add a small amount of sampling noise

of leakage is the contribution of more eigenvalues to the signal $g(k)$. For low leakage, we may expect three dominant eigenvalues with relatively large amplitude $\tilde{A}_j$ and five eigenvalues with small amplitude.

For a gate on two qutrits, identical remarks apply, except that an additional basis transformation is required from the orthogonal Gell–Mann basis to the computational qubit Pauli basis in order to keep the same division of T^{GM} as in Eq. (17). If we have two qutrits, the 80-dimensional traceless subspace is spanned by the matrices $\sigma_\mu^{GM} \otimes \sigma_\nu^{GM}$ for $\mu, \nu = 0, \ldots, 8$ except $\mu = \nu = 0$. The issue is related to terms such as $\sigma_0^{GM} \otimes \sigma_{\nu\neq0}^{GM}$ since the normalization of the qutrit identity $\left(\sigma_0^{GM} = \frac{1}{\sqrt{3}}I_3\right)$ is different from the normalization of the qubit identity ($P_0 = \frac{1}{\sqrt{2}}I_2$). This suggests that for two qutrits it is better to write T^{GM} in a basis that includes the Pauli matrices in the computational subspace ($P_\mu \otimes P_\nu$ for $\mu, \nu = 0, \ldots, 3$ except $\mu = \nu = 0$) as a sub-basis. For two qutrits, the signal may then contain up to 80 eigenvalues of which all but 15 are expected to have small amplitude in case of low leakage.

Non-Markovianity: time-correlated noise. Non-Markovian behavior of a gate can be due to temporal correlations in the classical or quantum environment of the driven qubit(s). Abstractly, we can include the environment in the gate action so that the evolution for each gate application is a unitary given by some U_{total} acting on system and environment. We can expand the Pauli transfer matrix of U_{total} in a Pauli basis for system and environment and view T_{comp} as a sub-block of T_{total}, similar as in the case of leakage. Diagonalizing T_{total} and taking the trace over the computational space will result in an expression such as Eq. (18). For example, an additional spectator or environment qubit can lead to a signal $g(k)$ of a single-qubit gate having contributions from 15 eigenvalues. Choosing a sufficiently large K may allow one to resolve these eigenvalues, even those with small amplitude.

A more malicious, but physically reasonable,[30] form of classical

non-Markovian noise makes gate-parameters temporally correlated. In order to numerically study the effect of non-Markovian noise, we consider a toy example in which a perfect CZ gate is followed by a rotation around the x axis on one qubit. For a series of k repetitions of a perfect CZ gate, we assume that each one is followed by the same rotation $R_x(\phi)$ acting always on the same qubit. We assume that the angle ϕ is Gaussian-distributed with mean 0 and standard deviation σ: $\mathbb{P}_\sigma(\phi) = exp(-\phi^2/2\sigma^2)/\sqrt{2\pi}\sigma$. The time evolution for k repetitions is then given by

$$\mathcal{S}_k(\rho) = \int_{-\infty}^{+\infty} d\phi \mathbb{P}_\sigma(\phi)(R_x(\phi)\,CZ)^k \rho(CZ\,R_x(\phi)^\dagger)^k. \quad (19)$$

Note that $\mathcal{S}_k \neq (\mathcal{S}_1)^k$ since this noise is correlated across multiple repetitions of the gate. Furthermore, we assume perfect state preparation and measurement. In this case, one can represent the noisy gate by some unitary U_{total} acting on the two qubits and on a classical stochastic mixture of angles ϕ. The continuous nature of this classical environment state leads to a lack of a hard cut-off on the number of eigenvalues in $g(k)$.

We apply the matrix-pencil method to the corresponding signal $g^{NO\ SPAM}(k)$ and we use an F-test to determine the optimal number of eigenvalues for each σ (Fig. 4). For $\sigma = 22.9°$ and $K = 50$, we find eigenvalues with modulus clearly >1. Those are unphysical but not excluded by the matrix-pencil method. We expect that such $|\lambda^{est}| > 1$ disappear when considering a longer signal, since $g(k)$ does not increase exponentially in k. In other words, this is a sign that the signal contains more spectral content than can be resolved from the time scale set by K. Indeed, for $\sigma = 22.9°$ we have made the same analysis for larger Ks up to $K = 200$ and we find that those eigenvalues get closer and closer to 1. If instead we fix $K = 50$ and consider different σs, we find that for a low σ (e.g., 5.7°) unphysical eigenvalues are not present (Fig. 4), whereas for $\sigma > 22.9°$ (e.g. 40.1°) they get again closer and closer to 1. This latter fact can be understood by noting that increasing σ is analogous to enlarging the time scale set by K, as the characteristic time scale of dephasing gets shorter for a fixed K. Based on these observations, we conclude that there is a certain intermediate time scale at which eigenvalues >1 are extrapolated from the data in the presence of sufficiently strong non-Markovian noise of the kind described in this section. Section "Frame mismatch accumulation" discusses a model with a different kind of time correlation leading to a spectral footprint which is incompatible with that of a TPCP map.

Non-Markovianity: coherent revivals. In order to better understand the occurrence of eigenvalue estimates $|\lambda^{est}| > 1$, we apply the matrix-pencil method on a signal (of a somewhat different physical origin), which has a revival over the time period set by K.

It is well known that, in the exchange of energy between a two-level atom with a bosonic mode, the Rabi oscillations of the two-level atom are subject to temporal revivals. These revivals are due to the fact that the bosonic driving field is not purely classical but rather gets entangled with the state of the qubit via the Jaynes–Cummings interaction. In particular, for a coherent driving field with coherent amplitude a with average photon number $\bar{n} = |a|^2$, the probability for the atom to be excited equals (see section 3.4.3 in ref. [20]):

$$P_e(t) = \frac{1}{2} + \frac{1}{2}\sum_{n=0}^{\infty} p_a(n)\cos(\Omega t\sqrt{n+1}), \quad (20)$$

with $p_a(n) = \exp(-|a|^2)\frac{|a|^{2n}}{n!}$. We consider $\bar{n} = 5$ and sample the damped oscillatory function $P_e(t) - \frac{1}{2}$ at regular intervals $k\Omega\delta t$ with $\Omega\delta t = 0.05$ and $k = 0, \ldots, K = 900$. The signal function $g(t) = P_e(t) - \frac{1}{2}$ contains eigenvalues equal to $\lambda_n = \exp(\pm i 0.05\sqrt{n+1})$ with amplitudes according to the Poisson distribution $p_a(n)$ with
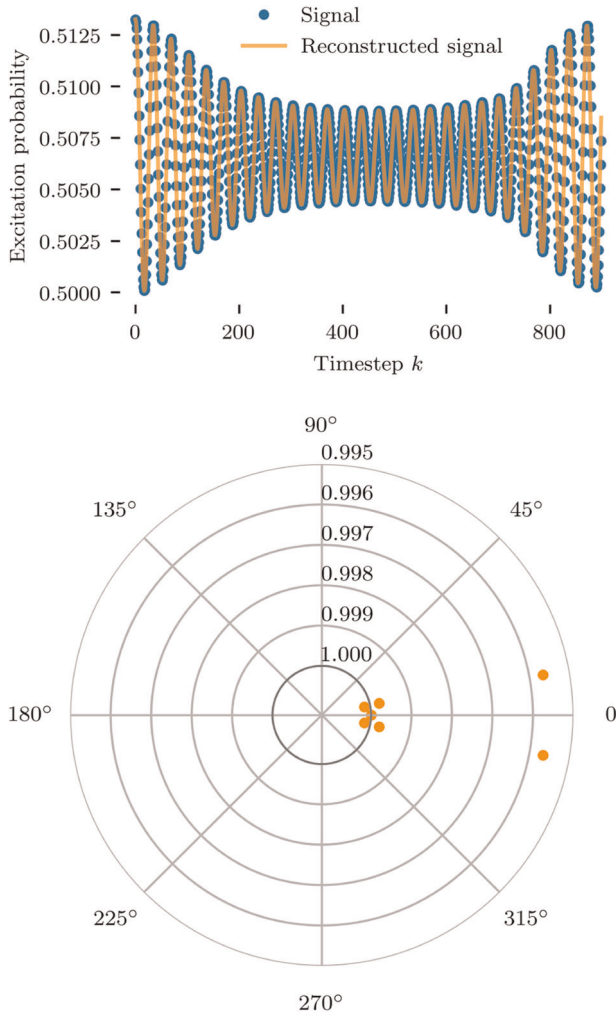
Fig. 5 Study of the reviving signal given in Eq. (20) for $k \cdot \Omega \delta t = k \cdot 0.05$, $\bar{n} = 5$, and $K = 900$. We find that the reviving signal is well reconstructed by a fit with 15 eigenvalues, some of which are distinctly separated as can be seen in the spectral footprint. Some of the eigenvalues are estimated to be >1. This is another example in which the matrix-pencil method gives unphysical eigenvalues in the presence of non-Markovian behavior (revivals here, time-correlated parameters in Fig. 4)

mean photon number $\bar{n}$.

We observe that the matrix-pencil method finds eigenvalues >1, see Fig. 5, which contribute significantly ($p < 0.01$ via F-test) to the reconstructed signal. We can understand this feature of eigenvalues >1 as a way in which the matrix-pencil method handles revivals: the signal has more spectral content than what can be resolved from the window of time given by K, in particular there is no hard cut-off on the number of eigenvalues that contribute. We have observed that an analysis of the signal over a longer period of time, that is, a larger K up to $K = 5000$, gives eigenvalues whose norm converges to at most 1.

DISCUSSION

We have introduced spectral quantum tomography, a simple method that uses tomographic data of the repeated application of a noisy quantum gate to reconstruct the spectrum of this quantum gate in a manner resistant to SPAM errors. We have experimentally validated our method on one- and two-qubit gates and have also numerically investigated its behavior in the presence of temporally correlated non-trivial error models.

The effective upshot of leakage and non-Markovian noise is that the signal will have more spectral content than what can be resolved given a chosen sequence length K, leading to unphysical features in the spectrum such as an eigenvalue estimate >1, or the absence of a real eigenvalue. Even though we have seen in our examples that a physical spectrum can be regained by going to larger K, depending on the noise model, this convergence may be very slow requiring much data-taking time. Hence, these unphysical features are useful markers for deviations from our model of repeated TPCP qubit maps $\mathcal{S}^k$. We view it as an open question how well one can reliably distinguish different sources of deviations.

An interesting application of the spectral tomography method could be the assessment of logical gates on encoded quantum information in a SPAM-resistant fashion. In this logical scenario (for, say, a single logical qubit), one first prepares the eigenstates of the logical Pauli operators $\bar{X}$, $\bar{Y}$, and $\bar{Z}$. One then applies a unit of error-correction $k = 0, \ldots, K$ times: a single unit could be, say, the repeated error correction for L rounds of a distance-L surface code. Or a unit is the application of a fault-tolerant logical gate, e.g., by means of code-deformed error correction or a transversal logical gate followed by a unit of error correction. After k units, one measures the logical Pauli operators fault-tolerantly and repeats experiments to obtain the logical signal $\bar{g}(k)$. Studying the spectral features of such logical channel will give information about the efficacy of the quantum error correction unit and/or the applied logical gate while departures from the code space or a need to time-correlate syndrome data beyond the given QEC unit can show up as leakage and non-Markovian errors.

METHODS

Single-qubit case with non-diagonalizable matrix T

In general, a matrix T can be brought to Jordan normal form by a similarity transformation, i.e., $T = VJV^{-1}$ with $J = \oplus_i J_i$ where each Jordan block J_i is of the form

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}, \tag{21}$$

see, e.g., Theorem 3.1.11 in ref. [31]. T is diagonalizable when each Jordan block is fully diagonal.

An example of a non-diagonalizable Lindblad superoperator on a single qubit has been constructed in ref. [32]. Using this, one can easily get a single-qubit superoperator $\mathcal{S}$ for which the traceless block of the Pauli transfer matrix is a non-diagonalizable matrix T as follows. Let $\mathcal{S}(\rho) = \exp(\mathcal{L}\epsilon)(\rho) \approx \rho + \epsilon\mathcal{L}(\rho) + O(\epsilon^2)$ with $\mathcal{L}(\rho) = -i[\frac{YZ}{2}, \rho] + \mathcal{D}[(2x)^{1/2}\sigma_-](\rho) + \mathcal{D}[y^{1/2}X](\rho)$ with $\mathcal{D}[A](\rho) = A\rho A^\dagger - \frac{1}{2}\{A^\dagger A, \rho\}$ and real parameters $x, y \geq 0$. This implies that $\mathcal{S}$ has the 4×4 Pauli transfer matrix

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \epsilon x & -\epsilon y & 0 \\ 0 & \epsilon y & 1 - \epsilon(x + 2y) & 0 \\ 2\epsilon x & 0 & 0 & 1 - 2\epsilon(x + y) \end{pmatrix} + O(\epsilon^2).$$

Taking some small ϵ and $x \neq 0$, one can check that the submatrix T does not have 3 eigenvectors and it has a pair of degenerate eigenvalues, so T is not diagonalizable. When we take $x = 0$, $\mathcal{S}$ is unital, that is, $\mathcal{S}(I) = I$, and the submatrix T is not diagonalizable either.

Even though a matrix T is not always diagonalizable, there still exists the so-called Schur triangular form for any matrix T.[31] This form says that $T = W(D + E)W^\dagger$, with W a unitary matrix, D a diagonal matrix with the eigenvalues of T, and E a strictly upper-triangular "nilpotent" matrix with non-zero entries only above the diagonal. Since the $N \times N$ matrix E is strictly upper-triangular, one has $\text{Tr}[D^j E^j] = 0$ for all $j \neq 0$. If we use this form in Eq. (12), one obtains for any k

$$g^{\text{NO SPAM}}(k) = \text{Tr}[T^k] = \text{Tr}[(D + E)^k] = \text{Tr}[D^k], \tag{22}$$

since any product of the form $D^{l_1}E^{l_2}D^{l_3}\ldots E^{l_m}$ with some non-zero $l_i > 0$ is a matrix with zeros on the diagonal. In case of SPAM errors and non-diagonalizable T, we consider

$$g(k) = \mathrm{Tr}[W^\dagger T_{\mathrm{prep}} T_{\mathrm{meas}} W (D + E)^k], \tag{23}$$

where $W^\dagger T_{\mathrm{prep}} T_{\mathrm{meas}} W$ is not the identity matrix due SPAM errors, implying that $g(k)$ can depend on E and have a non-exponential dependence on k. Thus, in the special case of a non-diagonalizable matrix T, the signal $g(k)$ would not have the dependence on the eigenvalues as in Eq. (13).

In particular, we can examine the physically interesting non-diagonalizable Case 3 in section "Relation to relaxation and dephasing times" in this light, taking $h_y = h_z = 0$ and a critical $h_x^{\mathrm{cr}} = \frac{\Gamma_1 - \Gamma_2}{2}$. The dynamics of the Lindblad equation after time t induces a superoperator S_t, which will have the following action on the Pauli operators:

$$\mathcal{S}_t(X) = \exp(-\Gamma_2 t)X,$$
$$\mathcal{S}_t(Y) = \exp(-(\Gamma_1 + \Gamma_2)t/2)\left[(1 + th_x^{\mathrm{cr}})Y - h_x^{\mathrm{cr}}Z\right],$$
$$\mathcal{S}_t(Z) = \exp(-(\Gamma_1 + \Gamma_2)t/2)\left[h_x^{\mathrm{cr}} tY + (1 - h_x^{\mathrm{cr}} t)Z\right].$$

Here we can note the linear dependence on t due to the system being critically damped. If we consider the signal $g(t) = \sum_\mu \mathrm{Tr}[P_\mu \mathcal{S}_t(P_\mu)]$, we see that this linear dependence on t drops out in accordance with Eq. (22), i.e., this trace only depends on the eigenvalues and has an exponential dependence on t. In the presence of SPAM errors, some of the linear dependence could still be observable for such critically damped system. In addition, coefficients such as $c_{\mu\nu}(t) = \mathrm{Tr}[P_\mu \mathcal{S}_t(P_\nu)]$ can depend linearly on t, making such tomographic data less suitable to extract eigenvalue information.

Upper bound on the entanglement fidelity with the targeted gate

In this section, we show how to relate the eigenvalues of the Pauli transfer matrix of a TPCP map $\mathcal{S}$ to an upper bound on the entanglement fidelity (and hence the average gate fidelity) with the targeted unitary gate U. Naturally, one can only expect to obtain an upper bound on the gate fidelity, since the eigenvalues do not provide information about the eigenvectors of $\mathcal{S}$. If the actual eigenvectors deviate a lot from ideal, the actual gate fidelity could be very low, so one can certainly not derive a lower bound on the fidelity based on the eigenvalues.

Lemma 1. Let the eigenvalues of the $N \times N$ matrix $T^\mathcal{S}$ be $\{\lambda_i\}_{i=1}^N$ with $N = d^2 - 1$ for a d-dimensional system. Let U be the targeted gate with eigenvalues $\{\lambda_i^{\mathrm{ideal}}\}_{i=1}^N$ and let there be permutation π of ith eigenvalue λ_i, which maximizes $|\sum_i \lambda_{\pi(i)}^* \lambda_i^{\mathrm{ideal}}|$ so that $0 \leq \xi_{\max} = \max_\pi \frac{1}{N}|\sum_i \lambda_{\pi(i)}^* \lambda_i^{\mathrm{ideal}}| \leq 1$. The entanglement fidelity $\mathcal{F}_{\mathrm{ent}}(U, \mathcal{S}) = \frac{1}{N+1}\left(1 + \mathrm{Tr}\left[T^{U\dagger}T^\mathcal{S}\right]\right)$ is upper bounded as

$$\mathcal{F}_{\mathrm{ent}}(U, \mathcal{S}) \leq \frac{1}{N+1}\left(1 + N\sqrt{u(\mathcal{S}) - \frac{\sum_j |\lambda_j|^2}{N} + N\xi_{\max}}\right), \tag{24}$$

where $u(\mathcal{S})$ is the unitarity of $\mathcal{S}$.

Proof. We write $T^\mathcal{S}$ in Schur triangular form as $T^\mathcal{S} = W(D^\mathcal{S} + E)W^\dagger$ with W a unitary matrix, $D^\mathcal{S}$ a diagonal matrix with the eigenvalues of $T^\mathcal{S}$, and E a strictly upper-triangular "error" matrix with non-zero entries only above the diagonal.[31] Using the Cauchy-Schwartz inequality, one has

$$\mathrm{Tr}[T^{U\dagger}T^\mathcal{S}] \leq \mathrm{Tr}[T^{U\dagger}WD^\mathcal{S}W^\dagger] + (\mathrm{Tr}[E^\dagger E])^{1/2}(\mathrm{Tr}[T^{U\dagger}T^U])^{1/2}. \tag{25}$$

Note that for a unitary gate U, $T^{U\dagger} = (T^U)^T = (T^U)^\dagger$ and $T^{U\dagger}T^U = I$ implying that T is an orthogonal matrix with unit singular values. We thus have $(\mathrm{Tr}[T^{U\dagger}T^U])^{1/2} = \sqrt{N}$. One has $\mathrm{Tr}[T^{\mathcal{S}\dagger}T^\mathcal{S}] = \mathrm{Tr}[(D^\mathcal{S} + E)^\dagger(D^\mathcal{S} + E)] = \mathrm{Tr}[(D^{\mathcal{S}\dagger}D^\mathcal{S} + E^\dagger E)]$, using the strict upper-triangularity of E. In other words, $\mathrm{Tr}[E^\dagger E] = \mathrm{Tr}[T^{\mathcal{S}\dagger}T^\mathcal{S}] - \sum_i |\lambda_i|^2$ where λ_i are the eigenvalues of $T^\mathcal{S}$. Recognizing that $\frac{1}{N}\mathrm{Tr}[T^{\mathcal{S}\dagger}T^\mathcal{S}] = u(\mathcal{S})$, we obtain an upper bound on the second term in Eq. (25).

Now let us upper bound the first term in Eq. (25) for unknown unitary W. Assume w.l.o.g. that T^U and $D^\mathcal{S}$ are diagonal in the same basis (the additional rotation between these eigenbases can be absorbed into W). Let $T^U = \sum_i \lambda_i^{\mathrm{ideal}} P_i$ and $D^\mathcal{S} = \sum_i \lambda_i P_i$ with orthogonal projectors P_i and $\sum_i P_i = I$. Define the matrix M with entries $M_{ij} = \mathrm{Tr}[P_i WP_j W^\dagger]$. The matrix M is doubly stochastic, since $\sum_i M_{ij} = 1 = \sum_j M_{ij}$, which implies that $M =$
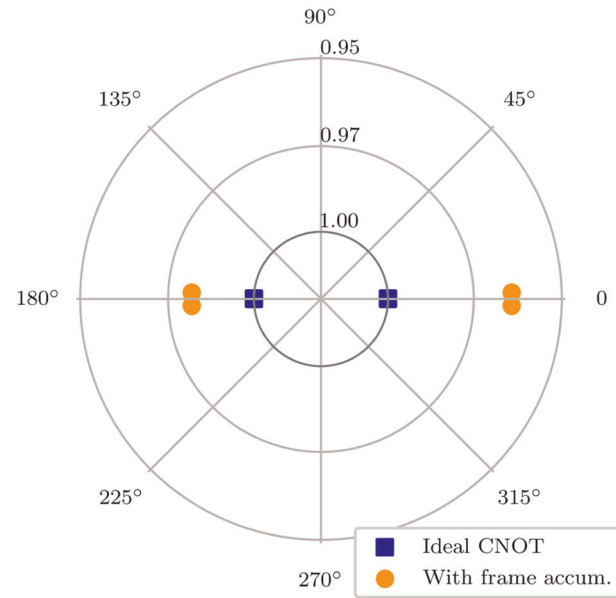
Fig. 6 Spectral footprint of a simulated CNOT gate affected by frame mismatch accumulation, for $K = 50$. The shown eigenvalues are $\{0.9636 + 0.03276i, 0.9636 - 0.0327i, -0.9804 + 0.0495i, -0.9804 - 0.0495i\}$, qualitatively matching the experimentally measured eigenvalues shown in Fig. 3 and critically matching the lack of real eigenvalues observed in Fig. 3

$\sum_m q_m \pi_m$ with $q_m \geq 0, \sum_m q_m = 1$ (Birkhoff–von Neumann theorem[31]) with permutation matrix π_m. With these facts and the convention $\langle i|\lambda^\mathcal{S}\rangle = \lambda_i$, we can bound

$$|\mathrm{Tr}[T^{U\dagger}WD^\mathcal{S}W^\dagger]| \leq \sum_m q_m |\lambda^{\mathrm{ideal}}\pi_m \lambda^\mathcal{S}| \leq N\xi_{\max}.$$

These bounds together then lead to Eq. (24).

An immediate corollary of Lemma 1 is

$$\mathcal{F}_{\mathrm{ent}}(U, \mathcal{S}) \leq \frac{1}{N+1}\left(1 + N\sqrt{1 - \frac{\sum_j |\lambda_j|^2}{N} + N\xi_{\max}}\right), \tag{26}$$

since $u(\mathcal{S}) \leq 1$ for TPCP maps. However, this is in general not a very strong upper bound on the fidelity.

We can do better in the single-qubit case by realizing that there are strong relations between the singular values σ_i of $T^\mathcal{S}$ and the absolute values of the eigenvalues $|\lambda_i|$ of $T^\mathcal{S}$. Ordering both the singular values and the eigenvalue magnitudes in descending order, we have the following (weak majorization) inequalities for arbitrary matrices

$$\prod_{i=1}^N \sigma_i = \prod_{i=1}^N |\lambda_i|, \tag{27}$$

$$\sum_{i=1}^F \sigma_i \geq \sum_{i=1}^F |\lambda_i|, \quad F \in \{1, \ldots, N-1\}. \tag{28}$$

For single-qubit channels, we can also impose TPCP constraints to the singular values of the channel. In particular, we have [ref. [33], Eq. (4)]

$$\sigma_i \leq 1, \quad \forall i \in \{0, 1, 2, 3\}, \tag{29}$$

$$\sigma_1 + \sigma_2 \leq 1 + \sigma_3. \tag{30}$$

Using these relations, we can upper bound the unitarity of a single-qubit

channel $\mathcal{S}$, given its eigenvalues, using the optimization:

$$\underset{\sigma_1,\sigma_2,\sigma_3}{\text{minimize}}\ u(\mathcal{S}) = \tfrac{1}{3}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2)$$

$$\text{subject to } \sigma_1\sigma_2\sigma_3 = |\lambda_1||\lambda_2||\lambda_3|,$$
$$1 \geq \sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0,$$
$$\sigma_1 + \sigma_2 \leq 1 + \sigma_3,$$
$$\sigma_1 + \sigma_2 \geq |\lambda_1| + |\lambda_2|,$$
$$\sigma_1 + \sigma_2 + \sigma_3 \geq |\lambda_1| + |\lambda_2| + |\lambda_3|.$$

This is a non-convex optimization problem in three variables, for which a global minimum can be numerically computed given $\lambda_1, \lambda_2, \lambda_3$. This gives an upper bound on the unitarity of $\mathcal{S}$ and hence on the entanglement fidelity of $\mathcal{S}$ to the target unitary U. In the main text, we use this optimization to give non-trivial upper bounds on the fidelities of single-qubit gates realized on superconducting chips and analyzed using the spectral tomographic method.

Frame mismatch accumulation

In section "Spectral tomography on two superconducting chips," we noted that the data gathered for the CNOT gate cannot be explained by a model of a noisy TPCP map $\mathcal{S}$ repeated k times. Here we propose a simple coherent model that qualitatively reproduces the features observed in Fig. 3 and we call this the frame mismatch accumulation model. Let $\mathcal{S}_0$ be a TPCP map that is a good approximation of the targeted gate applied exactly once (in the main text, this was the CNOT) and let V be some unitary. In the frame mismatch accumulation model, we assume that k consecutive applications of the gate are equal to:

$$\mathcal{S}_k = \prod_{i=0}^{k} (V^\dagger)^i \mathcal{S}_0 V^i = (V^\dagger)^{k+1}(V\mathcal{S}_0)^k. \qquad (31)$$

Intuitively, this can be interpreted as an increasing mismatch between the frame in which $\mathcal{S}_0$ was defined and the frame in which the gate was implemented at the ith repetition, up to $i = k$.

We apply this model to a CNOT gate, choosing $\mathcal{S}_0$ to be an ideal CNOT gate and choosing $V = \exp(-i\frac{\theta}{2}I \otimes Y)$ with $\theta = 0.05$ deg. In the case of the cross-resonance CNOT gate performed on *ibmqx4*, this may correspond to an imperfect cancellation of the $I \otimes Y$ term.[34] In Fig. 6, we see that this example closely reproduces the eigenvalues shown in Fig. 3. At the same time, we note that the qualitative features observed in Fig. 6 do not depend on the choice of the rotation axis of V (for either qubit), as long as the rotation does not commute with $\mathcal{S}_0$ (which would leave the gate unaffected by the frame mismatch).

AUTHOR CONTRIBUTIONS
All authors contributed to the development of the theoretical concepts presented. The experiments on the IBM QE and the QuTech QI were performed and analyzed by J.H. and the simulations on non-Markovianity were performed by J.H. and F.B., under supervision of B.M.T. All authors contributed to the writing of the manuscript.

ADDITIONAL INFORMATION

REFERENCES
1. Knill, E. et al. Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
2. Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106**, 180504 (2011).
3. Magesan, E. et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.* **109**, 080505 (2012).
4. Wood, C. J. & Gambetta, J. M. Quantification and characterization of leakage errors. *Phys. Rev. A* **97**, 032306 (2018).
5. Wallman, J., Granade, C., Harper, R. & Flammia, S. T. Estimating the coherence of noise. *New J. Phys.* **17**, 113020 (2015).
6. Dirkse, B., Helsen, J. & Wehner, S. Efficient unitarity randomized benchmarking of few-qubit clifford gates. E-prints at https://arxiv.org/abs/1808.00850 (2018).
7. Erhard, A. et al. Characterizing large-scale quantum computers via cycle benchmarking. E-prints at https://arxiv.org/abs/1902.08543 (2019).
8. Onorati, E., Werner, A. H. & Eisert, J. Randomized benchmarking for individual quantum gates. E-prints at https://arxiv.org/abs/1811.11775 (2018).
9. Kimmel, S., da Silva, M. P., Ryan, C. A., Johnson, B. R. & Ohki, T. Robust extraction of tomographic information via randomized benchmarking. *Phys. Rev. X* **4**, 011050 (2014).
10. Blume-Kohout, R. et al. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nat. Commun.* **8**, 14485 (2017).
11. Greenbaum, D. Introduction to quantum gate set tomography. E-prints at https://arxiv.org/abs/1509.02921 (2015).
12. O'Brien, T. E., Tarasinski, B. & Terhal, B. Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments. *New J. Phys.* **21**, 023022 (2019).
13. Bennink, R. S. & Lougovski, P. Quantum process identification: a method for characterizing non-markovian quantum dynamics. E-prints at https://arxiv.org/abs/1803.02438 (2018).
14. Wolf, M. M. & Perez-Garcia, D. Assessing quantum dimensionality from observable dynamics. *Phys. Rev. Lett.* **102**, 190504 (2009).
15. Strikis, A., Datta, A. & Knee, G. C. Quantum leakage detection using a model-independent dimension witness. *Phys. Rev. A* **99**, 032328 (2019).
16. Rudnicki, L., Puchała, Z. & Życzkowski, K. Gauge invariant information concerning quantum channels. *Quantum* **2**, 60 (2018).
17. Ruskai, M.-B., Szarek, S. & Werner, E. An analysis of completely-positive trace-preserving maps on m2. *Linear Algebra Appl.* **347**, 159–187 (2002).
18. Wolf, M. Quantum channels and operations guided tour. https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf (2012).
19. Horodecki, M., Horodecki, P. & Horodecki, R. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A* https://doi.org/10.1103/PhysRevA.60.1888 (1998).
20. Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford University Press, Oxford, 2006).
21. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
22. Sarkar, T. K. & Pereira, O. Using the matrix pencil method to estimate the parameters of a sum of complex exponentials. *IEEE Antennas Propag. Mag.* **37**, 48–55 (1995).
23. Potts, D. & Tasche, M. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra Appl.* **439**, 1024–1039 (2013).
24. Helsen, J., Wallman, J. J., Flammia, S. T. & Wehner, S. Multi-qubit randomized benchmarking using few samples. Preprint at https://arxiv.org/pdf/1701.04299.pdf (2017).
25. Barends, R. et al. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500 (2014).
26. Xue, X. et al. Benchmarking gate fidelities in a si/sige two-qubit device. *Phys. Rev. X* **9**, 021011 (2019).
27. Ballance, C., Harty, T., Linke, N., Sepiol, M. & Lucas, D. High-fidelity quantum logic gates using trapped-ion hyperfine qubits. *Phys. Rev. Lett.* **117**, 060504 (2016).
28. Wu, C.-F. J. et al. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Stat.* **14**, 1261–1295 (1986).
29. Seber, G. & Wild, C. *Nonlinear Regression* (John Wiley & Sons, Hoboken, NJ, 2003).
30. Rol, M. A. et al. A fast, low-leakage, high-fidelity two-qubit gate for a programmable superconducting quantum computer. E-prints at https://arxiv.org/abs/1903.02492 (2019).
31. Horn, R. A. & Johnson, C. R. *Matrix Analysis* (Cambridge University Press, 1985).

32. Sarandy, M. S. & Lidar, D. A. Adiabatic approximation in open quantum systems. *Phys. Rev. A* **71**, 012331 (2005).

33. Wolf, M. M. & Perez-Garcia, D. The inverse eigenvalue problem for quantum channels. Preprints at https://arxiv.org/abs/1005.4545 (2010).

34. Sheldon, S., Magesan, E., Chow, J. M. & Gambetta, J. M. Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Phys. Rev. A* **93**, 060302 (2016).