Improving transmon qudit measurement on IBM Quantum hardware

Tobias Kehrer ⬤,[*] Tobias Nadolny ⬤, and Christoph Bruder
Department of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland

The Hilbert space of a physical qubit typically features more than two energy levels. Using states outside the qubit subspace can provide advantages in quantum computation. To benefit from these advantages, individual states of the d-dimensional qudit Hilbert space have to be discriminated during readout. We propose and analyze two measurement strategies that improve the distinguishability of transmon qudit states. Based on a model describing the readout of a transmon qudit coupled to a resonator, we identify the regime in hardware parameter space where each strategy is optimal. We discuss these strategies in the context of a practical implementation of the default measurement of a ququart on IBM Quantum hardware whose states are prepared by employing higher-order X gates that make use of two-photon transitions.

I. INTRODUCTION

Conventional quantum computing is based on qubits which are realized on two-level subspaces of a larger physical Hilbert space. A number of physical realizations of qubits have been proposed and implemented on various platforms. These include superconducting qubits [1], trapped ions [2], cold atoms and Rydberg atoms [3], as well as electron spins in quantum dots [4]. On all of these platforms, it is necessary to isolate the two-dimensional qubit space from the remaining states of the physical Hilbert space to avoid leakage out of the computation space. However, utilizing qudits, i.e., d-dimensional building blocks of quantum computation, can provide advantages [5–17]. Examples include implementing an ancilla qubit within the second and third excited states of a qudit [6]. Another example is the so-called shelving [9]: by transferring the population of the first excited state to the second excited state, the error of identifying the ground state decreases.

Superconducting qubits [18,19] are prominent building blocks of noisy intermediate-scale quantum systems. The most promising example is the so-called transmon that can effectively be described as a quantum anharmonic electromagnetic oscillator. In this system, the two lowest-energy levels are identified as the qubit. Taking into account higher-lying transmon levels leads to a natural realization of a superconducting qudit. The smallest extension of the qubit is the qutrit, i.e., a three-level system. Qutrits have been used to implement a Toffoli gate [10] with a significantly lower number of elementary gates compared with a realization based on two-level systems. Another interesting example is the recent experimental demonstration of a qutrit Greenberger Horne Zeilinger (GHZ) state [11].

In general, if one is interested in measuring a qudit state, a proper classification of all levels involved is needed. In [20], the qubit state is determined by a fit of the time evolution of the system. In setups which do not provide time-resolved data, such as the current IBM Quantum [21] devices, other methods of separating qudit states have to be employed [22–24]. The strategies described in both [22] and [23] involve exciting the qudit-resonator system at readout drive frequencies other than the default frequency. At the default frequency, the distinguishability of the ground state and first excited state is maximized, whereas using the adapted frequencies aims at optimizing distances between different pairs of qudit states.

In this paper, we propose and evaluate improvements of the measurement scheme of transmon qudit states by enhancing their distinguishability. To optimize the readout, we determine the measurement errors from the assignment matrix whose entries denote the probability to classify a measurement outcome to a state $|i\rangle$ even if state $|j\rangle$ was prepared. This assignment matrix is calculated using qudit-state-dependent resonator steady-state amplitudes obtained from a model describing the readout of a transmon qudit by driving a coupled resonator. The default measurement schedule of most superconducting quantum hardware consists of a single-tone drive applied to the readout resonator. The frequency of the tone is chosen to maximally separate the ground and first excited states. The strategies we propose are based on modified readout resonator drive frequencies that take into account the separation of all qudit states. These strategies include a single-frequency as well as a multifrequency readout scheme. For a ququart, viz., the four lowest states of a qudit, we compare the proposed strategies in simulation and show that depending on hardware parameters, both strategies can be beneficial. We furthermore compare the model to a measurement of the drive-frequency-dependent resonator states on a current IBM Quantum device.

The paper is organized as follows. In Sec. II, we present a mean-field model describing the readout sequence of a

---
[*]tobias.kehrer@unibas.ch
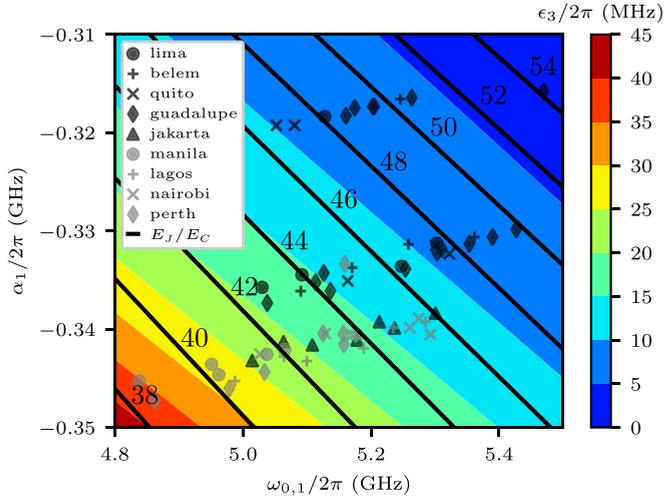
FIG. 1. Overview of qudit parameters of the IBM Quantum devices listed in the legend. The qubit resonance frequency and anharmonicity are denoted by $\omega_{0,1}$ and α_1; see Appendix A. The energy dispersion ϵ_3 of the third excited state given in Eq. (A5) follows from these device specifications. The labeled straight black lines denote constant values of E_J/E_C.

transmon qudit coupled to a harmonic readout resonator. Based on this model, we calculate the readout drive-frequency-dependent assignment errors between multiple states that in some limits can be expressed analytically. In Sec. III, we analyze both proposed readout schemes that aim to minimize these errors. We compare the data that we generated on current IBM Quantum hardware (see Sec. IV) to the readout model and strategies discussed in Secs. II and III. The data for Fig. 1 were obtained on May 23, 2023, whereas the data for Figs. 2–6 were obtained on July 7, 2023. To improve the state preparation required in this procedure, we propose to add two-photon transitions to the universal gate set of qudit gates [7] and show that this will reduce the execution time of certain qudit circuits and the duration of X-gate calibrations; see Appendixes B 2 and C. Finally, we conclude in Sec. V.

II. TRANSMON-RESONATOR SYSTEM

A. Effective Hamiltonian

The fundamental building blocks of a superconducting quantum computer are a quantum anharmonic oscillator, i.e., the transmon qudit, coupled to a harmonic oscillator, i.e., the readout resonator. Following [19], the corresponding effective Hamiltonian obtained by treating the Jaynes-Cummings interaction between the qudit and the resonator as a small perturbation (see Appendix B 1) reads

$$H_{\text{eff}} = \sum_j (\omega_j + \chi_{j-1,j} + \chi_j a^\dagger a)|j\rangle\langle j| + \omega_r a^\dagger a. \quad (1)$$

The parameter ω_j is the energy (see Appendix A) of the bare qudit state $|j\rangle$, ω_r is the energy of the readout resonator, and $a^{(\dagger)}$ is its annihilation (creation) operator. Here and in the rest of the paper, we set $\hbar = 1$. The second-order corrections to the qudit and resonator energies, $\chi_{j-1,j}$ and χ_j, are defined in

Eqs. (B11) and (B12). Additionally, we describe a coherent driving of the resonator at frequency ω_d by [1]

$$H_d = \frac{\Omega}{2}(e^{i\omega_d t - i\phi}a + e^{-i\omega_d t + i\phi}a^\dagger), \quad (2)$$

which enables the readout of qudit states.

B. Readout of states

The readout of a transmon qudit, in short, consists of driving the readout resonator while recording the response signal. We model the time evolution of a general quantum state ρ comprised of a qudit and its readout resonator by the following Lindblad master equation:

$$\frac{d}{dt}\rho = -i[H_{\text{eff}} + H_d, \rho] + \kappa \mathcal{D}[a](\rho), \quad (3)$$

$$\mathcal{D}[a](\rho) = a\rho a^\dagger - \frac{1}{2}(a^\dagger a\rho + \rho a^\dagger a), \quad (4)$$

where κ is the decay rate of the resonator. Using the effective Hamiltonian given by Eq. (1) and assuming the qudit to be in state $|j\rangle$, we arrive at the equation of motion of the mean-field amplitude $A \equiv \langle a \rangle = \text{Tr}[a\rho]$,

$$\frac{d}{dt}A = -i(\omega_r + \chi_j)A - i\frac{\Omega}{2}e^{-i\omega_d t + i\phi} - \frac{\kappa}{2}A. \quad (5)$$

The fact that A depends on the qudit state $|j\rangle$ is used to discriminate different qudit states. If the qudit is in a mixture or superposition of states, this measurement procedure projects the qudit onto one of its Fock states [1].

The general form of the complex value returned by an IBM Quantum device is

$$\bar{A} = \int_0^T dt\, k(t)A, \quad (6)$$

where $k(t)$ encodes the kernel integration instructions (see meas_kernel [25] in QISKIT [26]) and T is the total duration of the measurement. The choice $k(t) = \exp(i\omega_d t)$ corresponds to integrating the measurement signal in the rotating frame of the drive (see Sec. II B 1), whereas the choice $k(t) = \exp(i\omega_m t)$ corresponds to a frame rotating at an arbitrary modulation frequency ω_m (see Sec. II B 2).

In this paper, we mainly consider the offset charge configuration $n_g = 0$. The value of n_g influences the transmon qudit energy spectrum; see Appendix A. Note that due to the significant dependence on n_g of the third and higher excited states, their corresponding readout resonator states may be smeared out in phase space if charge noise is present. In Fig. 1, we present an overview of the energy dispersion ϵ_3 of the third excited state defined in Eq. (A5) for several IBM Quantum devices. Since ϵ_3 decreases with increasing E_J/E_C, qudits that lie in the upper-right region are preferred in general.

1. Rotating frame of drive

In this section, we will work in the rotating frame of the drive. Quantities in this frame will be denoted by the superscript (d). Since Eq. (5) is defined in the laboratory frame, we choose $k(t) = \exp(i\omega_d t)$ to transform the signal into the
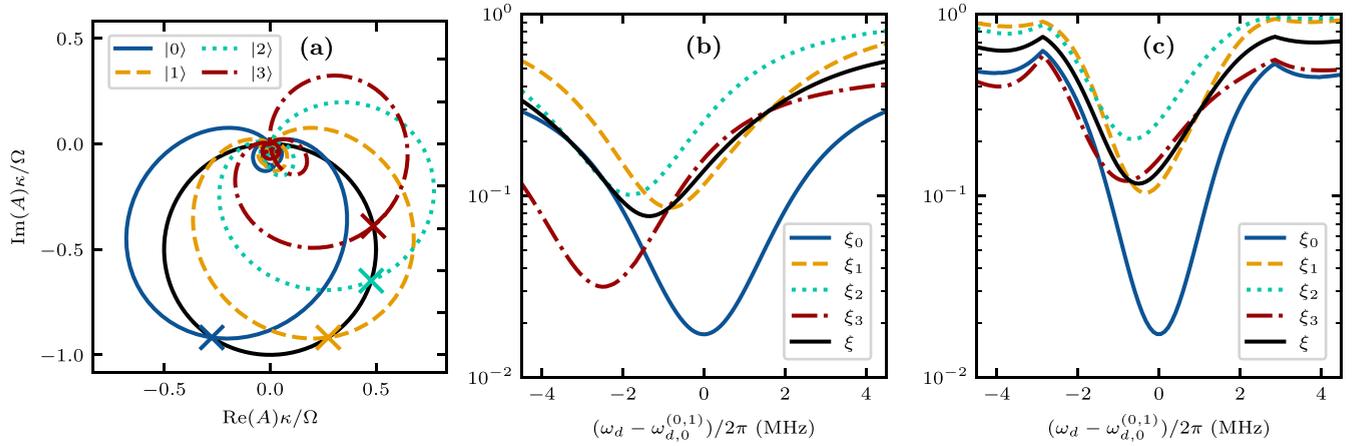
FIG. 2. (a) Drive-frequency-dependent phase-space positions $A_j^{(d)}$ and $A_j^{(m)}$ of the coherent state of the resonator given the qudit prepared in $|j\rangle$; see Eqs. (7) and (12). For $\omega_m = \omega_d$, the trajectories of all states $A_j^{(d)}$ match, denoted by the black circle. The colored lines denote $A_j^{(m)}$ for $\omega_m = \omega_{d,0}^{(0,1)}$. Crosses highlight the positions at $\omega_d = \omega_m$ where both models match, $A_j^{(d)} = A_j^{(m)}$. (b),(c) Error measures ξ_j and ξ in the frame of ω_d and ω_m respectively; see Eqs. (17) and (18). Following Appendixes A and B 1, for these plots, we determine E_J/E_C by the qubit parameters $\omega_{0,1}$ and α_1 from ibm_lagos Q4. Moreover, we choose $g/2\pi = 100$ MHz, $\Omega/2\pi = 100$ MHz, $\kappa/2\pi = 5$ MHz, $T = 0.35$ μs, $\sigma_j = 0.13\Omega/\kappa$, $\phi = 0$, and $n_g = 0$.

rotating frame of the drive and obtain

$$\frac{\bar{A}^{(d)}}{T} \xrightarrow{\kappa T \gg 1} -\frac{\Omega}{2}\frac{e^{i\phi}}{\omega_r + \chi_j - \omega_d - i\kappa/2} \equiv A_j^{(d)}. \quad (7)$$

Here, $A_j^{(d)}$ is the steady-state resonator amplitude when the qudit is in state $|j\rangle$ and defines a coherent state $|A_j^{(d)}\rangle$. Its dependence on the resonator drive frequency is presented in Fig. 2(a). Varying ω_d, the steady-state amplitudes $A_j^{(d)}$ of the readout resonator move on a circle centered at $A_c = -ie^{i\phi}\Omega/2\kappa$ with diameter Ω/κ. At resonance $\omega_{d,0}^{(j)} = \omega_r + \chi_j$, the states reach the maximum amplitude $2A_c$. For qudit readout, it is important that the distance $d_{i,j} = |A_i^{(d)} - A_j^{(d)}|$ between two qudit-state-dependent resonator states is large. We can identify two regimes of how the positions of the states in phase space depend on the readout drive frequency. For a large resonator decay rate $\kappa > |\chi_i - \chi_j|$, all states are close to the position of maximum amplitude within the same frequency range. In this case, $d_{i,j}$ exhibits only one maximum at

$$\omega_{d,0}^{(i,j)} = \omega_r + \frac{\chi_i + \chi_j}{2}, \quad (8)$$

$$d_{i,j}\left(\omega_{d,0}^{(i,j)}\right) = \frac{2\Omega|\chi_i - \chi_j|}{(\chi_i - \chi_j)^2 + \kappa^2}. \quad (9)$$

In contrast, for a small resonator decay rate $\kappa < |\chi_i - \chi_j|$, the frequency ranges where the state amplitudes $A_j^{(d)}$ are close to the maximum amplitude do not match. Here, two drive frequencies $\omega_d = \omega_{d,\pm}^{(i,j)}$ maximize the distance $d_{i,j}$,

$$\omega_{d,\pm}^{(i,j)} = \omega_{d,0}^{(i,j)} \pm \frac{1}{2}\sqrt{(\chi_i - \chi_j)^2 - \kappa^2}, \quad (10)$$

$$d_{i,j}\left(\omega_{d,\pm}^{(i,j)}\right) = \frac{\Omega}{\kappa} \equiv d_c. \quad (11)$$

where d_c denotes the diameter of the circle on which the states move. Thus, at $\omega_{d,\pm}^{(i,j)}$, both states are located on opposite

sides of the circle, which is the maximum separation they can obtain.

If we set the drive frequency to $\omega_{d,0}^{(i,j)}$ (or $\omega_{d,\pm}^{(i,j)}$), i.e., maximizing the distance between state $|i\rangle$ and $|j\rangle$, the distance between other pairs of states is in general reduced and hence not optimal for discrimination of these states. Therefore, in Sec. III, we present two measurement strategies to mitigate this issue.

2. General rotating frame

In a frame of a general rotation frequency ω_m, i.e., choosing $k(t) = \exp(i\omega_m t)$, the state reached in the long-time limit $\kappa T \gg 1$ is time dependent,

$$\frac{\bar{A}^{(m)}}{T} \xrightarrow{\kappa T \gg 1} e^{i[(\omega_m - \omega_d)T/2]}\operatorname{sinc}[(\omega_d - \omega_m)T/2]A_j^{(d)} \equiv A_j^{(m)}, \quad (12)$$

where $\operatorname{sinc}(x) = \sin(x)/x$ and the superscript (m) is used to denote quantities in this frame. The difference between Eqs. (7) and (12) is an additional factor of sinc peaking at $\omega_d = \omega_m$. These resonator state amplitudes and their dependence on the drive frequency ω_d are also visualized in Fig. 2(a). The states $A_j^{(d)}$ move on the black circle with diameter Ω/κ, whereas the motion of the states $A_j^{(m)}$ follows a distorted circle (colored lines).

III. MEASUREMENT STRATEGIES

In the previous section, we presented a model describing the readout on superconducting quantum hardware. The centers of the Wigner functions of the coherent readout resonator states when the qudit is in state $|j\rangle$ are given by A_j. Due to intrinsic quantum noise and hardware limitations, the possible readout resonator states for each qudit state overlap. This leads to potential misclassification and thus measurement errors when reading out the qudit states.

In the following, we propose two strategies for improving qudit readout compared to the default measurement scheme that utilizes a single resonator drive frequency that optimizes the classification of $|0\rangle$ and $|1\rangle$. The first strategy consists of finding a single frequency that maximizes the distinguishability between all d qudit Fock states. In the second strategy, we allow for multiple different drive frequencies.

A. Assignment matrix

To arrive at a measure of the distinguishability of states, we introduce the measurement assignment matrix M [27]. The qudit-state-dependent resonator states are defined by their steady-state amplitude A_j. We assume their Wigner functions to follow a two-dimensional Gaussian distribution,

$$G(z, A_j, \sigma_j) = \frac{1}{2\pi\sigma_j^2} \exp\left(-\frac{|z - A_j|^2}{2\sigma_j^2}\right), \qquad (13)$$

centered at A_j with standard deviation σ_j larger than the intrinsic quantum noise. The elements of M are given by

$$M_{i,j} = \int d^2z \, G(z, A_j, \sigma_j) \prod_{k \neq i} \Theta_{i,k} \qquad (14)$$

and define the probability to classify a measurement as state $|i\rangle$ even if state $|j\rangle$ was prepared. The region corresponding to each state $|i\rangle$ is defined by the maximum likelihood estimator (MLE) leading to

$$\Theta_{i,k} = \Theta[G(z, A_i, \sigma_i) - G(z, A_k, \sigma_k)], \qquad (15)$$

where Θ denotes the Heaviside function. For $\sigma_j = \sigma$ (valid assumption for this hardware setup; see the discussion about the distribution of σ of Gaussian fits in the second paragraph of Sec. IV), the MLE is equivalent to the minimum distance estimator (MDE) that implies

$$\Theta_{i,k} = \Theta(|z - A_k| - |z - A_i|). \qquad (16)$$

Using the MDE, a data point z is assigned to the region of state A_i if its Euclidean distance to all of the other states A_k is larger. In contrast, using the MLE, a data point z is assigned to the region of the state A_i that has the largest probability density. For simplicity and since in our measurements all σ_j are comparable, we choose the MDE throughout this paper.

Ideally, $M_{i,j} = \delta_{i,j}$, meaning perfect measurement. We define two measures ξ_j and ξ,

$$\xi_j = 1 - M_{j,j}, \qquad (17)$$

$$\xi = \frac{1}{d} \sum_{j=0}^{d-1} \xi_j, \qquad (18)$$

where ξ_j is the probability of misclassifying the qudit state $|j\rangle$. The dependence of ξ_j and ξ on the readout resonator drive frequency is shown in Figs. 2(b) and 2(c). The measurement errors ξ_j achieve their minima at different readout resonator drive frequencies. If $\omega_m \neq \omega_d$, the locations of the minima cannot be distinguished as well as for $\omega_m = \omega_d$. In the current setup of IBM Quantum hardware, the frequency ω_m of the rotating frame cannot be changed. Therefore, the difference between the frequency dependencies of all ξ_j is less pronounced.

Note that for setups where all qudit states lie on a circle (e.g., always for qutrits) and $\sigma_j = \sigma$, $M_{i,j}$ can be expressed in terms of Owen's T function; see Appendix E. This allows for fast numerical calculation of Eq. (14).

B. Finite sampling

In experiments, measuring an unknown state $|\psi\rangle = \sum_j c_j |j\rangle$ in the Z basis is equivalent to estimating its populations $p_j \equiv |c_j|^2$ based on a set of N data points $\{z_j\}$, also called shots. For each shot, the total state is projected onto one of the d qudit states $|j\rangle$ with probability p_j. Therefore, the total probability distribution of measuring one shot at z given $\vec{p}$ is a sum of all d Gaussians defined in Eq. (13) weighted by p_j. The measurement task can be understood as learning the parameters $\vec{p} = (p_j)$ of this multimodal probability distribution,

$$P(z|\vec{p}) = \sum_{j=0}^{d-1} p_j G(z, A_j, \sigma_j), \qquad (19)$$

where A_j and σ_j are obtained from a separate measurement. The space of possible $\vec{p}$,

$$p_j \in [0, 1], \quad p_{d-1} = 1 - \sum_{j=0}^{d-2} p_j, \qquad (20)$$

can be mapped to a $(d-1)$-simplex, using the normalization condition of $\vec{p}$.

We describe the measurement analysis as follows. Each shot is labeled by a state $|j\rangle$ depending on its phase-space distance (MDE) to the d qudit Fock states. The resulting list of counts $\vec{N}$ can be used to obtain information about $\vec{p}$. In Appendix D, we show that by using Bayes' formalism, the probability distribution of $\vec{p}$ given $\vec{N}$ is the Dirichlet distribution,

$$P(\vec{p}|\vec{N}) = \frac{1}{\mathcal{N}} \text{Dir}(M\vec{p}, \vec{N}). \qquad (21)$$

The component N_j of $\vec{N}$ equals the number of shots classified as state $|j\rangle$. The assignment matrix M reflects the fact that some shots are classified incorrectly. Given $\vec{N}$, the location of the maximum (also called mode) can be computed analytically,

$$\vec{p}_{\text{mode}} = \frac{1}{N} M^{-1} \vec{N}. \qquad (22)$$

This result is similar to a common procedure known in QISKIT as "measurement error mitigation." Note that applying the inverse of M to $\vec{N}$ can lead to negative components of $\vec{p}_{\text{mode}}$. In QISKIT, this problem is circumvented by approximating $\vec{p}_{\text{mode}}$ by the valid $\vec{p}\,'$ that is closest (in two-norm) to $\vec{N}$ (see method least_squares in qiskit.utils.mitigation._filters.py [26]),

$$\vec{p}\,' = \underset{\vec{p}}{\text{argmin}}(|\vec{N}/N - M\vec{p}|^2). \qquad (23)$$

Equation (23) is the estimate of the state populations $\vec{p}$ after measuring $\vec{N}$ shots. We will use the uncertainty of these estimates, viz., the numerically calculated standard deviations SD$[p_j]$, to decide which of the proposed

strategies performs best, i.e., exhibits the smallest standard deviation.

C. Comparison of strategies

We consider two strategies that make use of either one or multiple drive frequencies. In the default readout scheme of superconducting quantum hardware, measurement pulses with a single drive frequency that maximizes the distinguishability between the qubit states $|0\rangle$ and $|1\rangle$ are applied.

The first strategy we propose replaces the default frequency by the one that optimally separates all qudit states in phase space simultaneously. Since, in general, the state $|\psi\rangle$ that we want to measure is unknown, we suggest to optimize ξ [see Eq. (18)], which is the average of the individual measurement errors ξ_j.

The second strategy uses N/d shots for each of the d different frequencies at which individual states are most isolated, i.e., ξ_j are minimal. We will show that this strategy is advantageous in cases when there is no single frequency at which all states are separated well enough. Hardware parameters and the state to be measured determine which of the two strategies outperforms the other.

To compare both strategies, we draw $N = 1000$ samples from the probability distribution given by Eq. (19) for $\sigma_j = \sigma$ and an equal-superposition state $p_j = 1/d$. The drive frequencies we use are the locations of the minimum of ξ for the single-drive strategy and the minima of ξ_j for the multifrequency strategy. Each sample is classified using the MDE [see Eq. (16)], i.e., by its Euclidean distance to the nearest state $|j\rangle$. The final probability distribution for the p_j of the single-frequency strategy is given in Eq. (21). The final probability distribution for the multifrequency strategy is the normalized product of the term in Eq. (21) for each measurement frequency ω_k,

$$P(\vec{p}\,|\{\vec{n}_k\}) \propto \prod_{k=0}^{d-1} \mathrm{Dir}(M(\omega_k)\vec{p}\,|\vec{n}_k), \qquad (24)$$

where $\vec{n}_k$ is the list of counts of classified shots for the kth measurement frequency. The standard deviation $\mathrm{SD}[p_j]$ is computed numerically from this distribution.

Figure 3 shows the dependence of the ratio of both averaged standard deviations,

$$\mathrm{SD}_{s/m} = \frac{1}{d} \sum_{j=0}^{d-1} \mathrm{SD}_{s/m}[p_j], \qquad (25)$$

of p_j on hardware parameters $\sigma_j = \sigma$ and κ. The blue region corresponds to setups for which the standard deviation SD_s of p_j using a single-drive-frequency scheme is smaller. In contrast, the red region corresponds to hardware configurations where it is beneficial to measure at multiple frequencies, i.e., SD_m of the multidrive frequency scheme is smaller. The gray region indicates parameter values for which both standard deviations exceed $\mathrm{SD}_{s/m} \geqslant 0.1$. Since the expected values p_j lie in [0,1], this threshold corresponds to an uncertainty of at least 10%.
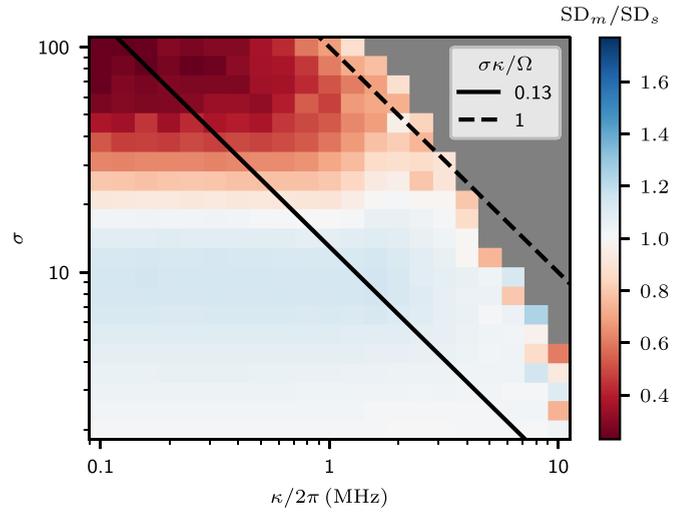
FIG. 3. Ratio of the standard deviation SD_m for the multifrequency strategy and the standard deviation SD_s of the single-frequency strategy applied to an equal-superposition state taking $N = 1000$ shots. The gray region indicates where both standard deviations exceed $\mathrm{SD}_{s/m} \geqslant 0.1$. The straight lines denote constant values of $\sigma\kappa/\Omega$. We take the same qudit parameters as in Fig. 2 and choose $g/2\pi = 100$ MHz and $\Omega/2\pi = 100$ MHz.

The overall trend is that for small σ, i.e., strongly located Gaussians, the single-frequency strategy performs at a similar, slightly better level than the multifrequency strategy. The multifrequency strategy is preferable for large σ, when the overlap of the Gaussians would be too large using a single drive frequency. Intuitively, this is expected since, for small κ and large σ, only one state is isolated from the others which group together at the origin in phase space; see discussion of regimes $\kappa \lesssim |\chi_j - \chi_{j+1}|$ in Sec. II B 1. We also added lines of constant relative uncertainty $\sigma\kappa/\Omega$. Along these lines, the Gaussian widths σ are fixed in units of the diameter Ω/κ of the circle on which the states move in the rotating frame of the drive. The solid line corresponds to $\sigma = 0.13\Omega/\kappa$ chosen in Figs. 2(b) and 2(c), whereas the dashed line approximately matches the threshold of $\mathrm{SD}_{s/m} \geqslant 0.1$. Following the solid black line, the standard deviation of the single-frequency strategy appears to exhibit a minimum around $\kappa/2\pi = 1 - 2$ MHz. For fixed $\sigma\kappa/\Omega$ and small resonator decay rates κ, the states move around the circle rather individually, whereas for large κ, the states move as a group; cf. Sec. II B 1.

IV. MEASUREMENT OF A QUQUART

In this section, we will compare the model described in Sec. II to data obtained from ibm_lagos Q4. We prepare the four lowest Fock states of the transmon qudit and measure them for various readout drive frequencies. In Appendix C, the preparation of the individual qudit Fock states is described. Sequences of X gates defined as simple Gaussian pulses are used to prepare the four lowest Fock states of the transmon qudit. As mentioned earlier, we make use of higher-order X gates; see Appendix B 2. These provide a reduction of the
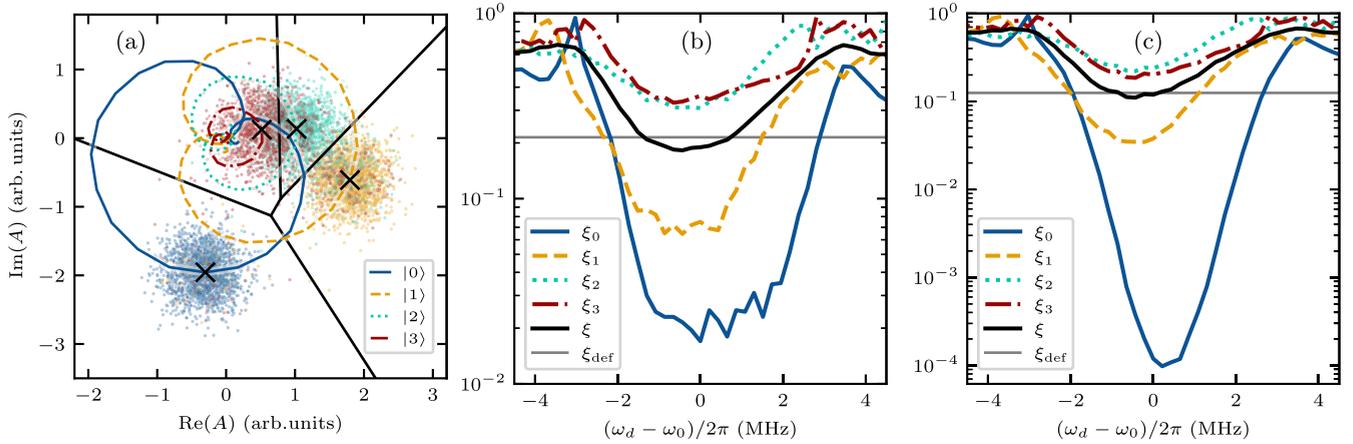
FIG. 4. (a) Drive-frequency-dependent ququart measurements on ibm_lagos Q4, the experimental equivalent of Fig. 2(a). The colored curves (with white shadows) correspond to the centers of Gaussian fits to the Fock states $|j\rangle$. Black straight lines indicate the boundaries of regions assigned to individual Fock states. These boundaries are constructed using the minimum distance estimator; see Eq. (16). For the drive frequency $\omega_0 = 7.2463$ GHz, we show the measurement results of all $N = 2000$ shots for each prepared Fock state. Black crosses mark the centers of their Gaussian fits. (b) Measurement errors $\xi_j = 1 - M_{j,j}$ based on all measured shots of the data presented in (a). Here, the elements $M_{i,j}$ of the assignment matrix equal the relative number of shots, N_i/N, that are classified as $|i\rangle$ even if $|j\rangle$ is prepared. The horizontal gray line denotes the measurement error ξ_{def} obtained using the default measurement pulse. (c) Measurement errors ξ_j based on Gaussian fits to the data presented in (a) and the assignment matrix M defined in Eq. (14). Using the centers and average σ of Gaussian fits for each readout resonator drive frequency, we calculate M numerically. The ξ_j shown in (b) are larger than those in (c) since they do not only represent assignment errors, but also include additional errors such as qudit decay, leakage, and imperfect state preparation. In both (b) and (c), the minimum of the average assignment error ξ will be smaller than ξ_{def} obtained by the default pulse if the modulation frequency is chosen such that $\omega_m = \omega_d$; cf. Figs. 2(b) and 2(c).

execution time of certain quantum circuits and a reduction of the duration of X-gate calibrations.

In Fig. 4(a), we show the measurements of the four lowest Fock states for various readout resonator drive frequencies. This plot is the experimental equivalent of Fig. 2(a). For each Fock state and for each readout resonator drive frequency, we take $N = 2000$ shots while keeping the other drive parameters fixed at the default values. For $\omega_0 = 7.2463$ GHz, Fig. 4(a) shows all shots in the color of the prepared Fock state. This value of ω_0 is -5.5 MHz off the default frequency reported by the IBM Quantum device. Black crosses highlight the centers of the Gaussian fits. For other drive frequencies, we only plot the centers of the Gaussian fits as colored lines (with a white shadow). The straight black lines denote the boundaries of regions [defined via MDE; see Eq. (16)] that are assigned to one Fock state.

We analyze the measurement errors in two ways. First, we define the elements $M_{i,j}$ by the relative number of shots, N_i/N, classified as $|i\rangle$ even if $|j\rangle$ is prepared. In this way, M incorporates misclassification errors but also additional errors such as imperfect qudit state preparation. From this matrix, we obtain the errors ξ_j, displayed in Fig. 4(b). Second, we use the centers of the Gaussian fits for each qudit state and for each value of the resonator drive frequency and a fixed value of σ to compute the assignment matrix defined in Eq. (14). By examining these Gaussian fits, we find a narrow distribution of the σ values: $\sigma = (0.302 \pm 0.017)$ [same arbitrary units as in Fig. 4(a)]. The resulting errors ξ_j are shown in Fig. 4(c). Here, the ξ_j only represent errors that arise from misassignment of shots drawn from the multi-Gaussian distribution; see Eqs. (13) and (14). Since real devices feature other sources of error, e.g., qubit decay, leakage, and imperfect state prepara-

tion, the values of ξ_j presented in Fig. 4(b) are larger than in Fig. 4(c).

Our model, visualized by the theory plots in Figs. 2(a) and 2(c), shows qualitative agreement with the data presented in Figs. 4(a) and 4(c). In both Figs. 4(b) and 4(c), the horizontal gray line ξ_{def} denotes the average assignment error of the four lowest Fock states using the default readout pulse and should be compared with the solid black line ξ. The corresponding data were taken from Rabi calibration measurements, similar to Fig. 6(b), at the drive amplitude that is closest to the fitted optimum.

We find a dependence of the measurement errors ξ_j on the readout resonator frequency as expected. The data presented in Figs. 4(b) and 4(c) suggest that the default measurement frequency is not ideal to separate all four qudit states. However, the minima appear at only slightly different positions. Note that the difference in positions is only small due to IBM Quantum software/hardware limitations: ω_m cannot be set to its ideal value $\omega_m = \omega_d$; see Sec. II B 2. We expect the impact of varying the readout resonator drive frequency to be much higher if it is possible to analyze all data in the rotating frame of the drive; compare Figs. 2(b) and 2(c).

In this paper, we focused on the analysis of only four qudit states since the readout of higher excited states beyond $|3\rangle$ becomes difficult for several reasons. Higher excited states are more sensitive to charge noise; see Fig. 5. Since χ_j depends on the qudit spectrum ω_j, charge noise leads to ambiguous steady-state amplitudes. In addition, finding a single drive frequency that properly separates all qudit states becomes more difficult with an increasing number of qudit states. For example, for the IBM Quantum device that we utilized in

this paper, we estimate $\chi_1 < \chi_4 < \chi_2$ which indicates that the steady-state amplitude corresponding to $|4\rangle$ lies between $|1\rangle$ and $|2\rangle$. We expect that the more states are involved, the better the performance of a multifrequency strategy in comparison to a single-frequency strategy given a small $\kappa < |\chi_i - \chi_j|$; cf. Fig. 3.

V. CONCLUSION

We have presented a model that describes phase-space measurement data of qudit states on superconducting quantum hardware. Our model qualitatively matches the data that we generated on a current IBM Quantum device. For qudit-state preparation, we employ higher-order X gates between $|j\rangle$ and $|j + 2\rangle$. This scheme leads to a reduction of the execution time of qudit quantum circuits as well as of the duration of X-gate calibrations. Based on our model, we have compared the performance of two measurement strategies, a single-frequency and a multifrequency scheme, in simulations. For each strategy, we have identified the regime in hardware parameter space where it is optimal. The multifrequency strategy is superior when the qudit-state-dependent resonator states overlap significantly.

To use the full potential of both strategies, it is necessary to adjust the modulation frequency ω_m of the device. This is currently not possible on IBM Quantum hardware. Despite these software/hardware restrictions, we still find differences in the frequency locations of the minima of the individual measurement errors ξ_j and an improvement over the measurement error ξ_{def} using the default measurement pulse. We expect a better performance of the strategies for setups that operate in the rotating frame of the drive $\omega_m = \omega_d$.

In the future, adaptive measurement schemes that change the drive frequency from shot to shot or between bunches of shots may be possible. This can lead to a further improvement of transmon qudit measurements.

APPENDIX A: TRANSMON ENERGY LEVELS

To estimate resonance frequencies for various transitions, we numerically compute the energy levels of the transmon qudit Hamiltonian [19,28],

$$H_{\text{T}} = 4E_C \sum_n (n - n_g)^2 |n\rangle\langle n|$$
$$- \frac{E_J}{2} \sum_n (|n\rangle\langle n + 1| + |n + 1\rangle\langle n|), \quad \text{(A1)}$$
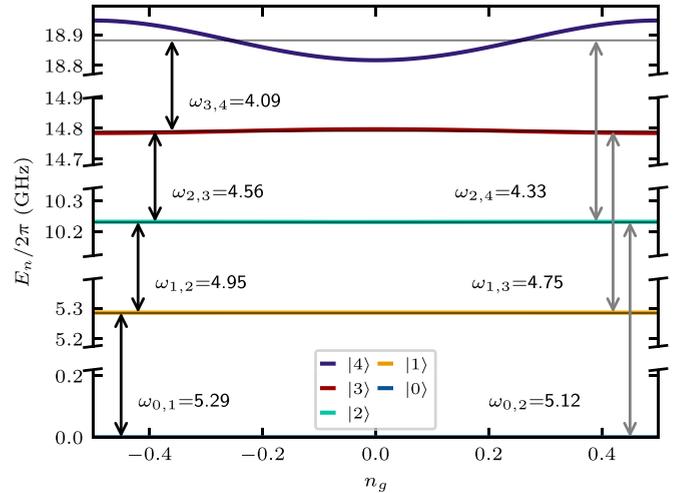
FIG. 5. Numerical prediction of the energy levels E_n of ibm_lagos Q4 based on Eq. (A1) and $E_J/E_C \approx 45.6$. The transition frequencies $\omega_{i,j}$ are displayed in units of (2π) GHz.

depending on the offset charge n_g and the ratio E_J/E_C. The sorted eigenvalues $E_n(n_g)$ are shifted such that $E_0(0) = 0$. We define the average transition frequency $\omega_{i,j}$ between $|i\rangle$ and $|j\rangle$ of both configurations $n_g = 0, 1/2$ as

$$\omega_{i,j} = \frac{E_j(0) + E_j(1/2) - E_i(0) - E_i(1/2)}{2(j - i)}, \quad \text{(A2)}$$

and the frequency difference $\Delta\omega_{i,j}$ as

$$\Delta\omega_{i,j} = \frac{E_j(0) - E_i(0) - E_j(1/2) + E_i(1/2)}{j - i}. \quad \text{(A3)}$$

The anharmonicity α_j and the energy dispersion ϵ_j of the transmon qudit are defined by

$$\alpha_j = \omega_{j,j+1} - \omega_{j-1,j}, \quad \text{(A4)}$$

$$\epsilon_j = E_j(0) - E_j(1/2). \quad \text{(A5)}$$

We numerically obtained the fundamental parameter E_J/E_C of a specific IBM Quantum backend by demanding that the qubit frequency $\omega_{0,1}$ and anharmonicity α_1 match the values reported by this device. In Fig. 5, we plot the dependence of E_n on n_g for the five lowest states. The values of the frequency difference vary from $\Delta\omega_{0,1}/2\pi = 25.1$ kHz to $\Delta\omega_{3,4}/2\pi = -142$ MHz. In Fig. 1, the values ϵ_3 for a number of IBM Quantum devices are displayed. For large ϵ_j compared to $\omega_{j,j+1}$, the Wigner function of the state $|j\rangle$ effectively is smeared out in phase space since both configurations $n_g = 0, 1/2$ exhibit different resonance frequencies.

APPENDIX B: SCHRIEFFER-WOLFF TRANSFORMATION

In this appendix, we will use Schrieffer-Wolff transformations [29] to obtain perturbative approximations of, first, the Jaynes-Cummings interaction between a qudit and a readout resonator and, second, the two-photon drive of the qudit.

In general, the system of interest can be defined by the Hamiltonian

$$H = H_0 + \lambda H_1 + \lambda V. \quad \text{(B1)}$$

Here, H_0 and H_1 are block diagonal in the subsystems, whereas V is block off diagonal. To find an effective block-diagonal Hamiltonian, i.e., eliminate the block off-diagonal part V, the unitary transformation $U = e^S$ is applied to H,

$$H_{\text{eff}} = UHU^{\dagger} = e^S H e^{-S}. \quad \text{(B2)}$$

Expanding the anti-Hermitian operator $S = -S^{\dagger}$ as $S = \sum_{n=1}^{\infty} \lambda^n S^{(n)}$, H_{eff} can be expressed as

$$H_{\text{eff}} = H + [S, H] + \frac{1}{2}[S, [S, H]] + \cdots = \sum_{n=0}^{\infty} \lambda^n H_{\text{eff}}^{(n)}. \quad \text{(B3)}$$

The first-order contribution reads $H^{(1)} = H_1 + V + [S^{(1)}, H_0]$. To eliminate the block off-diagonal V in this expression, we impose $V = -[S^{(1)}, H_0]$. Since H_0 is block diagonal, $S^{(1)}$ has to be block off diagonal. As a consequence, $[S^{(1)}, V]$ is block diagonal. The second-order contribution reads $H^{(2)} = \frac{1}{2}[S^{(1)}, V] + [S^{(1)}, H_1] + [S^{(2)}, H_0]$, and imposing $[S^{(1)}, H_1] = -[S^{(2)}, H_0]$ guarantees the second order to be block diagonal. The second-order contribution to the effective block-diagonal Hamiltonian is then given by

$$H_{\text{eff}}^{(2)} = \frac{1}{2}[S^{(1)}, V]. \quad \text{(B4)}$$

We choose a superposition of all operators appearing in V as an ansatz for $S^{(1)}$.

1. Jaynes-Cummings interaction

Following the notation of [19], the Hamiltonian describing a transmon qudit coupled to a readout resonator reads

$$
\begin{aligned}
H_q &+ H_r + H_{\text{int}} \\
&= \sum_j \omega_j |j\rangle\langle j| + \omega_r a^{\dagger} a \\
&+ \sum_j g_{j,j+1}(a^{\dagger}|j\rangle\langle j+1| + a|j+1\rangle\langle j|), \quad \text{(B5)}
\end{aligned}
$$

where ω_j is the energy (see Appendix A) of the bare qudit state $|j\rangle$, ω_r is the energy of the readout resonator, and $a^{(\dagger)}$ is its annihilation (creation) operator. The parameters $g_{j,j+1}$ denote generalized Jaynes-Cummings coupling strengths between the qudit and the resonator. Using the approximation $g_{j,j+1} = g\sqrt{j+1}$, the interaction Hamiltonian reduces to $g(a^{\dagger}b + ab^{\dagger})$ [18], where $b^{(\dagger)}$ is the annihilation (creation) operator of the transmon qudit. The qudit and resonator Hamiltonians H_q and H_r denote two blocks of commuting operators $|i\rangle\langle j|$ and $a^{(\dagger)}$. The generalized Jaynes-Cummings interaction couples both blocks.

We start with identifying the block-diagonal and block off-diagonal parts,

$$H_0 = H_q + H_r, \quad H_1 = H_d, \quad \text{(B6)}$$

$$V = H_{\text{int}}. \quad \text{(B7)}$$

Using a superposition of all operators appearing in V as an ansatz for $S^{(1)}$,

$$S^{(1)} = \sum_j (C_j a^{\dagger}|j\rangle\langle j+1| - C_j^* a|j+1\rangle\langle j|), \quad \text{(B8)}$$

the coefficients C_j are obtained as

$$C_j = \frac{g_{j,j+1}}{\omega_j - \omega_{j+1} + \omega_r}. \quad \text{(B9)}$$

Replacing C_j in Eq. (B8) by this expression and using the definition $g_{-1,0} = 0$ as well as the sign convention of [1,19] leads to

$$\frac{1}{2}[S^{(1)}, V] = \sum_j \chi_{j-1,j}|j\rangle\langle j| + \sum_j \chi_j a^{\dagger}a|j\rangle\langle j|. \quad \text{(B10)}$$

Here,

$$\chi_{j,j+1} = \frac{g_{j,j+1}^2}{\omega_{j+1} - \omega_j - \omega_r}, \quad \text{(B11)}$$

$$
\begin{aligned}
\chi_j &\equiv \chi_{j-1,j} - \chi_{j,j+1} \\
&= \frac{g_{j-1,j}^2}{\omega_j - \omega_{j-1} - \omega_r} - \frac{g_{j,j+1}^2}{\omega_{j+1} - \omega_j - \omega_r}, \quad \text{(B12)}
\end{aligned}
$$

and we have neglected terms proportional to $(a^2|j+2\rangle\langle j| +$ H.c.). This is justified by the possibility to interpret these terms as perturbations that are eliminated by a second Schrieffer-Wolff transformation. This will lead to terms proportional to $|j\rangle\langle j|$, $a^{\dagger}a|j\rangle\langle j|$, $(a^{\dagger}a)^2|j\rangle\langle j|$, and also $(a^4|j+4\rangle\langle j| + $ H.c.). Importantly, for typical values of $g_{j,j+1}$, ω_j, and ω_r, the coefficients of all these terms are a factor of 10^4 smaller than the previous second-order contributions and can therefore safely be neglected. In conclusion, in Eq. (B10), we arrived at corrections to the Hamiltonian that are diagonal in the qudit and resonator states. The shifts of the qudit and resonator energies are

$$
\begin{aligned}
\tilde{\omega}_j &= \omega_j + \chi_{j-1,j}, \\
\tilde{\omega}_{r,j} &= \omega_r + \chi_j. \quad \text{(B13)}
\end{aligned}
$$

The resonance frequencies of the qudit transitions $|i\rangle \leftrightarrow |j\rangle$ can be estimated to

$$\tilde{\omega}_{i,j} = \frac{\tilde{\omega}_j - \tilde{\omega}_i}{j - i}. \quad \text{(B14)}$$

2. Qudit drive

In analogy to Appendix B 1, we perform a Schrieffer-Wolff transformation in a system consisting of a qudit and its drive only. We will use this method to predict the Rabi oscillation frequencies of second-order transitions $|j\rangle \leftrightarrow |j+2\rangle$. For previous work on multiphoton transitions, see, e.g., [30,31]. Starting with the Hamiltonian [1],

$$
\begin{aligned}
H = H_0 + \lambda V &= \sum_j (\tilde{\omega}_j - \omega_d)|j\rangle\langle j| \\
&+ \lambda \frac{\Omega_q}{2} \sum_j \sqrt{j+1}(e^{-i\phi}|j\rangle\langle j+1| + e^{i\phi}|j+1\rangle\langle j|),
\end{aligned}
$$
$$\text{(B15)}$$
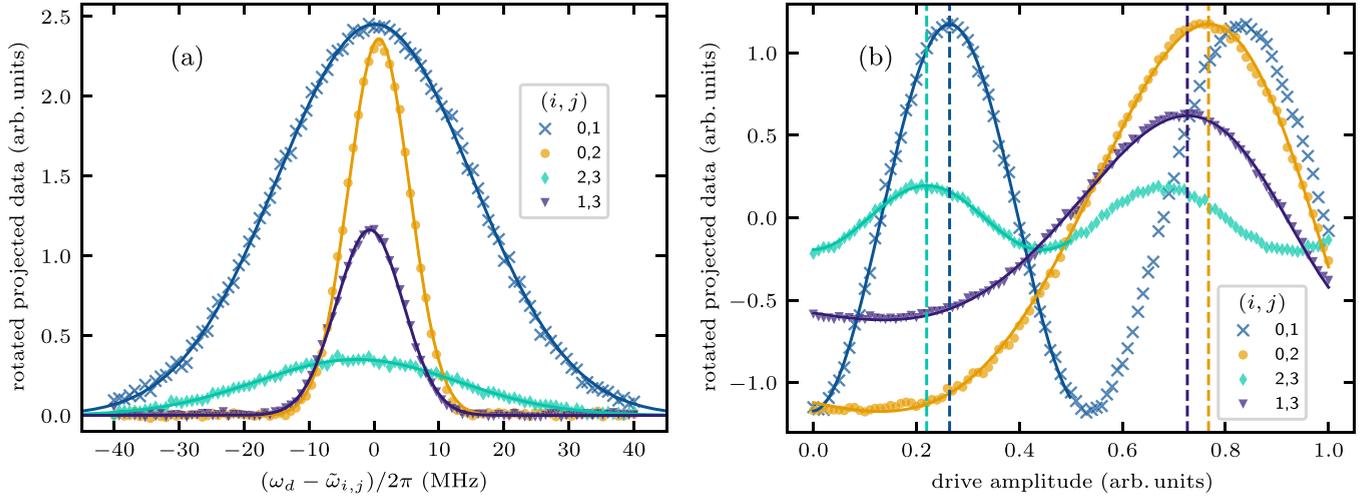
FIG. 6. (a) Qudit resonator spectroscopy of transitions $|i\rangle \leftrightarrow |j\rangle$. Colored markers denote measured data and solid curves correspond to Gaussian fits. We plot each resonance spectrum centered around the predicted transition frequency $\tilde{\omega}_{i,j}$ [see Eq. (B14)] using $g/2\pi = 65$ MHz. (b) Rabi oscillations $|i\rangle \leftrightarrow |j\rangle$ depending on the drive amplitude for fixed pulse duration. We sweep the readout resonator drive amplitude while keeping all other parameters of the drive fixed. For first-order qudit state transitions, we fit a sinusoidal dependence on a linear function of the drive amplitude in the interval [0,0.5]; see the lines connecting crosses and, respectively, diamonds. For second-order qudit state transitions, we fit a sinusoidal dependence on a quadratic function of the drive amplitude in the interval [0,1]; see the lines connecting dots and, respectively, triangles. The vertical dashed lines indicate the locations of the first maxima obtained from the fits.

an expansion in λ leads to

$$H_{\text{eff}}^{(2)} = -\frac{\Omega_q^2}{8} \sum_j f_j (e^{-2i\phi}|j\rangle\langle j+2| + e^{2i\phi}|j+2\rangle\langle j|),$$

(B16)

where

$$f_j = \frac{\sqrt{(j+1)(j+2)}(\tilde{\omega}_{j+2} - 2\tilde{\omega}_{j+1} + \tilde{\omega}_j)}{(\tilde{\omega}_{j+2} - \tilde{\omega}_{j+1} - \omega_d)(\tilde{\omega}_{j+1} - \tilde{\omega}_j - \omega_d)}.$$

(B17)

Note that these expressions only hold for $\omega_d \neq \tilde{\omega}_{j+1} - \tilde{\omega}_j$, i.e., drives that are not resonant with transitions between neighboring qudit levels $|j\rangle \leftrightarrow |j+1\rangle$.

APPENDIX C: PREPARATION OF QUDIT STATES

We prepare the four lowest Fock states of an IBM Quantum transmon qudit, described in Sec. IV, by applying sequences of calibrated X gates to the qudit ground state $|0\rangle$. For simplicity, we implement these X gates via Gaussian pulses. For each pulse, we first calibrate its drive frequency ω_d and second, its drive amplitude Ω_q.

The optimal drive frequency is obtained from a Gaussian fit to resonance measurement data shown in Fig. 6(a), where we fix the pulse amplitude to an initial estimate. First, the measured $N = 2000$ shots per qudit drive frequency are averaged. Second, these averages are rotated such that their major principal axis is oriented along the X axis. And third, the means are projected onto the X axis which justifies the axis label "rotated projected data." For the spectroscopy measurements, in addition, we define the origin of the y axis of Fig. 6(a) to correspond to the initial state of the analyzed transition and the maximum to the final state. Our estimated frequency $\tilde{\omega}_{i,j}$ is calculated by Eq. (B14) using $g/2\pi = 65$ MHz.

To obtain the initial estimate of the qudit drive amplitude, we define the rotation angle θ of a resonant Rabi oscillation between states $|j\rangle$ and $|k\rangle$. Comparing both sides of

$$\exp(-iHt) = \exp\left[-\frac{i}{2}\theta(|j\rangle\langle k| + |k\rangle\langle j|)\right],$$

(C1)

where the left-hand side is the time evolution of the effective drive Hamiltonian,

$$H = \Lambda(j,k)(|j\rangle\langle k| + |k\rangle\langle j|).$$

(C2)

Therefore, the rotation angle θ depends on the effective Rabi frequency $\Lambda(j,k)$ and the pulse duration t. Following Eq. (B15) for $\lambda = 1$, the rotation angle θ for Rabi oscillations between $|j\rangle$ and $|j+1\rangle$ is given by

$$\theta = t_{j,j+1}\Omega^{(j,j+1)}\sqrt{j+1}.$$

(C3)

The rotation angle for Rabi oscillations between non-neighboring states $|j\rangle$ and $|j+2\rangle$ can be computed using Eq. (B16),

$$\theta = t_{j,j+2}\frac{(\Omega^{(j,j+2)})^2}{4}f_j.$$

(C4)

The Rabi frequency of the $|0\rangle \leftrightarrow |2\rangle$ transition scales quadratically with $\Omega^{(0,2)}$. Using Eqs. (C3) and (C4), the initial estimate for the π-pulse amplitudes $\Omega_\pi^{(j,k)}$ can be related to the default X-gate amplitude $\Omega_\pi^{(0,1)}$ reported by the IBM Quantum backend,

$$\Omega_\pi^{(j,j+1)} = \frac{\Omega_\pi^{(0,1)}}{\sqrt{j+1}},$$

(C5)

$$\Omega_\pi^{(j,j+2)} = 2\sqrt{\frac{\Omega_\pi^{(0,1)}t_{0,1}}{f_j t_{j,j+2}}}.$$

(C6)

Here we used that all single-qudit operations are implemented within the same duration $t_{j,j+1} = t_{0,1}$.

After evaluating the resonance measurement, we calibrate the X-gate drive amplitude via Rabi oscillations; see Fig. 6(b). The data are rotated and projected onto the major principal axis as described before for the spectroscopy measurements. As shown in Eq. (B16), the Rabi frequency for transitions $|j\rangle \leftrightarrow |j+2\rangle$ depends nonlinearly on the drive amplitude. Since these transitions are suppressed by the small factor $\Omega_q f_j$, we choose $t_{j,j+2} = 2t_{0,1}$ such that $\Omega_\pi^{(0,2)}$ does not exceed the limits of IBM Quantum software/hardware restrictions. The π amplitude of an X-gate pulse is identified with the location of the first maximum in Fig. 6(b), indicated by dashed lines. For transitions between neighboring states, we fit a sine dependence on a linear function of Ω_q and, for second-order transitions, we fit a sine dependence on a second-order polynomial of Ω_q. Using those fits, any desired rotation angle, e.g., π for an X gate or $\pi/2$ for a Hadamard gate, can be mapped back to a corresponding pulse amplitude.

The sequence of calibrating drive frequency and amplitude described above can be iterated several times to improve gate fidelity. For simplicity, we consider only one round of calibrations. To increase fidelity, we chose the initial value for $\Omega_{1,3}$ based on prior test measurements.

Implementing gates in the $|j\rangle \leftrightarrow |j+2\rangle$ subspace results in two advantages. First, our implementation of an X-gate $X_{j,j+2}$ between $|j\rangle$ and $|j+2\rangle$ takes only twice the single-qudit gate duration $t_{0,1}$. In contrast, using single-qudit gates, $X_{j,j+2}$ consists of three single-qudit operations $X_{j,j+1}X_{j+1,j+2}X_{j,j+1}$ with a total duration of $3t_{0,1}$. Second, the calibration of the drive frequencies (amplitudes) for $|0\rangle \leftrightarrow |1\rangle$ and $|0\rangle \leftrightarrow |2\rangle$ are independent of each other and can therefore be combined into a single QISKIT job (set of measurements submitted to an IBM Quantum device). In contrast, the frequency calibration for the transition $|1\rangle \leftrightarrow |2\rangle$ depends on the Rabi measurement for the transition $|0\rangle \leftrightarrow |1\rangle$. In total, we can perform our calibration in four QISKIT jobs: (i) drive frequency of X_{01} and X_{02}, (ii) drive amplitude of X_{01} and X_{02}, (iii) drive frequency of X_{12}, X_{23}, and X_{13}, and (iv) drive amplitude of X_{12}, X_{23}, and X_{13}. In contrast, the standard sequential calibration of single-qudit X gates would take six jobs: two for each of the three single-qudit X gates between neighboring states.

APPENDIX D: DIRICHLET DISTRIBUTION

We describe the probability distribution that estimates the populations $p_j \equiv |c_j|^2$ of the state we want to analyze using Bayesian inference. To this end, we define a recursion relation

$$P^{(j+1)}(\vec{p}) = \frac{P(z_j|\vec{p})}{P(z_j)} P^{(j)}(\vec{p}), \tag{D1}$$

$$P(z_j) = \int d^{d-1}p \, P(z_j|\vec{p}) P^{(j)}(\vec{p}), \tag{D2}$$

between the estimated probability distribution (prior) $P^{(j)}(\vec{p})$ of the Gaussian amplitudes $\vec{p}$ before and after receiving the jth data point z_j, also called shot. Each shot follows the probability distribution $P(z_j|\vec{p})$ defined in Eq. (19). After obtaining N data points, this leads to the probability distribution

$P^{(N)}(\vec{p})$,

$$P^{(N)}(\vec{p}) = \frac{\prod_{j=0}^{N-1} P(z_j|\vec{p})}{\prod_{j=0}^{N-1} P(z_j)} P^{(0)}(\vec{p}). \tag{D3}$$

The initial prior $P^{(0)}(\vec{p})$ is chosen to be a uniform distribution.

If the width of the Gaussians is small compared to their distances, this method is equivalent to the description given in the following; see, also, Sec. III B. For simplicity, instead of using this Bayesian ansatz of the probability distribution, each of the N shots is classified as one of the d qudit states. The components N_j of $\vec{N}$ equal the number of shots classified as $|j\rangle$. This procedure corresponds to neglecting the position z_j of this shot, i.e., the Gaussian weight $G(z, A_j, \sigma_j)$ of the shot. Given N_j, the posterior probability distribution $P^{(N)}(\vec{p})$ for the qudit populations p_j for a perfect measurement is equal to

$$P(\vec{p}|\vec{N}) = \text{Dir}(\vec{p}, \vec{N}) = \frac{(N+d-1)!}{\prod_{k=0}^{d-1} N_k!} \prod_{k=0}^{d-1} p_k^{N_k}, \tag{D4}$$

with

$$\sum_{j=0}^{d-1} p_j = 1, \quad \sum_{j=0}^{d-1} N_j = N, \tag{D5}$$

introducing the Dirichlet distribution Dir [32]. The location of its maximum (also called mode) with respect to p_j is given by N_j/N, and its variances follow as

$$\text{Var}[p_j] = \frac{(N_j+1)/(N+d)[1-(N_j+1)/(N+d)]}{N+d+1}. \tag{D6}$$

Defining $n_j = N_j/N$, for large N the variance of p_j scales like $n_j(1-n_j)/N$.

We now consider the assignment matrix M [see Eq. (14)], which describes misclassification errors. Using a Bayesian posterior ansatz, we find that the probability distribution has to be modified to

$$P(\vec{p}|\vec{N}) = \frac{1}{\mathcal{N}} \text{Dir}(M\vec{p}, \vec{N}), \tag{D7}$$

where

$$\mathcal{N} = \int_{V_{\vec{p}}} d^d p \, \text{Dir}(M\vec{p}, \vec{N}). \tag{D8}$$

The assignment matrix maps proper states $\vec{p}$ from $V_{\vec{p}}$ [related to a $(d-1)$-simplex] to a subspace $V_{M\vec{p}} \subseteq V_{\vec{p}}$. Therefore, if $\vec{N}/N \in V_{M\vec{p}}$, the mode of Eq. (D7) is naturally $M^{-1}\vec{N}/N$.

APPENDIX E: ANALYTICAL ASSIGNMENT MATRIX

In this appendix, we present an analytical expression for the assignment matrix in specific setups that allows for a fast numerical implementation. The following expressions hold for systems with $\sigma_j = \sigma$. Furthermore, all states have to lie on a circle centered at $A_c = x_c + iy_c$. Examples are qudit systems with $\omega_m = \omega_d$ (cf. discussion in Sec. II B 1) or qutrit systems with arbitrary ω_m. Equation (14) can be

written as

$$M_{i,j} = \frac{1}{4}\left[1 - \text{erf}\left(\frac{x_j - x_c}{\sqrt{2}\sigma}\right)\right]$$
$$+ T\left(\frac{x_j - x_c}{\sigma}, -a_i, \frac{y_c - y_j + a_i(x_j - x_c)}{\sigma}\right), \quad (E1)$$

where the slope of the bisecting line between A_i and A_{i-1} is

$$a_i = -\frac{x_i - x_{i-1}}{y_i - y_{i-1}}. \quad (E2)$$

Here, the variables x_j and y_j are the real and imaginary part of A_j. The coordinate system is rotated such that A_i and A_{i+1} are aligned along the real axis in phase space.

Finally,

$$T(h, a, b) = \frac{1}{2\sqrt{2\pi}} \int_h^\infty dx \, \exp\left(-\frac{x^2}{2}\right) \text{erf}\left(\frac{ax + b}{\sqrt{2}}\right) = \frac{1}{4}\text{erf}\left(\frac{b}{\sqrt{2(1+a^2)}}\right)\left[1 - \text{erf}\left(\frac{h}{\sqrt{2}}\right)\right]$$
$$+ T\left(\frac{b}{\sqrt{1+a^2}}, a + \frac{h(1+a^2)}{b}\right) + T\left(h, a + \frac{b}{h}\right) - T\left(\frac{b}{\sqrt{1+a^2}}, \frac{h\sqrt{1+a^2}}{b}\right) - T\left(h, \frac{b}{h\sqrt{1+a^2}}\right) \quad (E3)$$

is a generalized version of Owen's T function, $T(h, a) = T(h, a, 0)$ [33,34].

[1] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, Rev. Mod. Phys. **93**, 025005 (2021).

[2] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, Trapped-ion quantum computing: Progress and challenges, Appl. Phys. Rev. **6**, 021314 (2019).

[3] M. Saffman, Quantum computing with atomic qubits and Rydberg interactions: Progress and challenges, J. Phys. B: At. Mol. Opt. Phys. **49**, 202001 (2016).

[4] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, Semiconductor spin qubits, Rev. Mod. Phys. **95**, 025003 (2023).

[5] Y. Wang, Z. Hu, B. C. Sanders, and S. Kais, Qudits and high-dimensional quantum computing, Front. Phys. **8**, 589504 (2020).

[6] L. E. Fischer, D. Miller, F. Tacchino, P. K. Barkoutsos, D. J. Egger, and I. Tavernelli, Ancilla-free implementation of generalized measurements for qubits embedded in a qudit space, Phys. Rev. Res. **4**, 033027 (2022).

[7] L. E. Fischer, A. Chiesa, F. Tacchino, D. J. Egger, S. Carretta, and I. Tavernelli, Universal qudit gate synthesis for transmons, PRX Quantum **4**, 030327 (2023).

[8] F. Tacchino, A. Chiesa, R. Sessoli, I. Tavernelli, and S. Carretta, A proposal for using molecular spin qudits as quantum simulators of light-matter interactions, J. Mater. Chem. C **9**, 10266 (2021).

[9] S. S. Elder, C. S. Wang, P. Reinhold, C. T. Hann, K. S. Chou, B. J. Lester, S. Rosenblum, L. Frunzio, L. Jiang, and R. J. Schoelkopf, High-fidelity measurement of qubits encoded in multilevel superconducting circuits, Phys. Rev. X **10**, 011001 (2020).

[10] A. Fedorov, L. Steffen, M. Baur, M. da Silva, and A. Wallraff, Implementation of a Toffoli gate with superconducting circuits, Nature (London) **481**, 170 (2012).

[11] A. Cervera-Lierta, M. Krenn, A. Aspuru-Guzik, and A. Galda, Experimental high-dimensional Greenberger-Horne-Zeilinger entanglement with superconducting transmon qutrits, Phys. Rev. Appl. **17**, 024062 (2022).

[12] M. A. Yurtalan, J. Shi, M. Kononenko, A. Lupascu, and S. Ashhab, Implementation of a Walsh-Hadamard gate in a superconducting qutrit, Phys. Rev. Lett. **125**, 180504 (2020).

[13] D. J. Egger, M. Werninghaus, M. Ganzhorn, G. Salis, A. Fuhrer, P. Müller, and S. Filipp, Pulsed reset protocol for fixed-frequency superconducting qubits, Phys. Rev. Appl. **10**, 044030 (2018).

[14] S. Zeytinoğlu, M. Pechal, S. Berger, A. A. Abdumalikov, A. Wallraff, and S. Filipp, Microwave-induced amplitude- and phase-tunable qubit-resonator coupling in circuit quantum electrodynamics, Phys. Rev. A **91**, 043846 (2015).

[15] A. Ciavarella, N. Klco, and M. J. Savage, Trailhead for quantum simulation of SU(3) Yang-Mills lattice gauge theory in the local multiplet basis, Phys. Rev. D **103**, 094501 (2021).

[16] M. Jerger, P. Macha, A. R. Hamann, Y. Reshitnyk, K. Juliusson, and A. Fedorov, Realization of a binary-outcome projection measurement of a three-level superconducting quantum system, Phys. Rev. Appl. **6**, 014014 (2016).

[17] S. Li, B.-J. Liu, Z. Ni, L. Zhang, Z.-Y. Xue, J. Li, F. Yan, Y. Chen, S. Liu, M.-H. Yung, Y. Xu, and D. Yu, Super-robust geometric control of a superconducting circuit, Phys. Rev. Appl. **16**, 064003 (2021).

[18] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, Phys. Rev. A **69**, 062320 (2004).

[19] 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, Charge-insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[20] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff, Control and tomography of a three level superconducting artificial atom, Phys. Rev. Lett. **105**, 223601 (2010).

[21] IBM Quantum, https://quantum.ibm.com/, 2023.

[22] L. Chen, H.-X. Li, Y. Lu, C. W. Warren, C. J. Križan, S. Kosen, M. Rommel, S. Ahmed, A. Osman, J. Biznárová, A. Fadavi Roudsari, B. Lienhard, M. Caputo, K. Grigoras, L. Grönberg, J. Govenius, A. F. Kockum, P. Delsing, J. Bylander, and G. Tancredi, Transmon qubit readout fidelity at the threshold for quantum error correction without a quantum-limited amplifier, npj Quantum Inf. **9**, 26 (2023).

[23] C. Wang, M.-C. Chen, C.-Y. Lu, and J.-W. Pan, Optimal readout of superconducting qubits exploiting high-level states, Fundament. Res. **1**, 16 (2021).

[24] K. C. Miao, M. McEwen, J. Atalaya, D. Kafri, L. P. Pryadko, A. Bengtsson, A. Opremcak, K. J. Satzinger, Z. Chen, P. V. Klimov, C. Quintana, R. Acharya, K. Anderson, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. C. Bardin, A. Bourassa, J. Bovaird *et al.*, Overcoming leakage in scalable quantum error correction, arXiv:2211.04728.

[25] D. C. McKay, T. Alexander, L. Bello, M. J. Biercuk, L. Bishop, J. Chen, J. M. Chow, A. D. Córcoles, D. J. Egger, S. Filipp, J. Gomez, M. Hush, A. Javadi-Abhari, D. Moreda, P. Nation, B. Paulovicks, E. Winston, C. J. Wood, J. Wootton, and J. M. Gambetta, Qiskit backend specifications for OpenQASM and OpenPulse experiments, arXiv:1809.03452.

[26] Qiskit contributors, Qiskit: An open-source framework for quantum computing (unpublished).

[27] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, Mitigating measurement errors in multiqubit experiments, Phys. Rev. A **103**, 042605 (2021).

[28] U. Vool and M. Devoret, Introduction to quantum electromagnetic circuits, Intl. J. Circuit Theory Appl. **45**, 897 (2017).

[29] J. R. Schrieffer and P. A. Wolff, Relation between the Anderson and Kondo Hamiltonians, Phys. Rev. **149**, 491 (1966).

[30] F. W. Strauch, S. K. Dutta, H. Paik, T. A. Palomaki, K. Mitra, B. K. Cooper, R. M. Lewis, J. R. Anderson, A. J. Dragt, C. J. Lobb, and F. C. Wellstood, Strong-field effects in the Rabi oscillations of the superconducting phase qubit, IEEE Trans. Appl. Supercond. **17**, 105 (2007).

[31] S. Danilin, A. Vepsäläinen, and G. S. Paraoanu, Experimental state control by fast non-Abelian holonomic gates with a superconducting qutrit, Phys. Scr. **93**, 055101 (2018).

[32] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Dirichlet and Inverted Dirichlet Distributions*, in Probability and Statistics, edited by W. A. Shewhart and S. S. Wilks (John Wiley & Sons, Ltd, Hoboken, New Jersey, 2000), Chap. 49.

[33] D. B. Owen, Tables for computing bivariate normal probabilities, Ann. Math. Stat. **27**, 1075 (1956).

[34] D. B. Owen, A table of normal integrals, Commun. Stat. Simul. Comput. **9**, 389 (1980).