Testing complementarity on a transmon quantum processor

Pedro M. Q. Cruz [1,2,*] and J. Fernández-Rossier[1,†]

[1]*QuantaLab, International Iberian Nanotechnology Laboratory, 4715-330 Braga, Portugal*
[2]*Departamento de Ciência de Computadores, Universidade do Porto, 4169-007 Porto, Portugal*

We propose quantum circuits to test interferometric complementarity using symmetric two-way interferometers coupled to a which-path detector. First, we consider the two-qubit setup in which the controlled transfer of path information to the detector subsystem depletes interference on the probed subspace, testing the visibility-distinguishability tradeoff via minimum-error state discrimination measurements. Next, we consider the quantum eraser setup, in which reading out path information in the right basis recovers an interference pattern. These experiments are then carried out in an IBM superconducting transmon processor. A detailed analysis of the results is provided. Despite finding good agreement with theory at a coarse level, we also identify small but persistent systematic deviations preventing the observation of full particlelike and wavelike statistics. We understand them by carefully modeling two-qubit gates, showing that even small coherent errors in their implementation preclude the observation of Bohr's strong formulation of complementarity.

I. INTRODUCTION

Complementarity is one of the most intriguing aspects of quantum mechanics, as it sets limits to our capability to gather simultaneous information about different observables. It was first discussed in the context of wave-particle duality by the founders of the theory, but it soon became clear that it applies to any set of noncommuting observables. The central conundrum of complementarity is often formulated in terms of a quantum object that can take two alternatives, or paths, that we denote by $|0\rangle$ and $|1\rangle$. The wave function is prepared in an equal superposition of the paths with a relative phase added:

$$|\psi\rangle = c_0|0\rangle + c_1 e^{i\phi}|1\rangle. \qquad (1)$$

As the phase is changed and the expectation value of a complementary observable is measured, an interference pattern emerges with an amplitude proportional to Re$\{c_0^* c_1\}$. This can only occur if the object evolves through both ways simultaneously, interfering with itself (wavelike character). If a system is set up to gather information about which path the object actually realizes (particlelike character), the interference contrast, or visibility, is washed out.

This modern understanding of the interplay between which-path information and interference visibility began its formulation in 1927 with Bohr's *principle of complementarity* [1]. Building on previous achievements [2,3], he first argued complementarity to be a fundamental quantum feature whereby objects possess pairs of properties the knowledge of which is mutually exclusive, the observation of either unambiguous particle trajectories or interference fringes being just one example.

In the same year, Einstein conceptualized a recoiling double-slit interferometer hoping to falsify Bohr's principle [4]. This first which-path thought experiment was discussed and followed by other proposals over the years [5–7], but only analyzed in depth half a century later [8]. On that analysis of the momentum exchange between the photon and the slit, the tradeoff between the interference visibility and the amount of which-path information retrieved was quantified. This description of the partial knowledge obtained from complementary observables confirmed and extended Bohr's original dichotomous version in which complete knowledge of one of them implies maximal uncertainty about the other.

Soon after, other which-path experiments were proposed to test complementarity enforced by physical mechanisms not arising from the position-momentum uncertainty relation [9,10]. The information-theoretic treatment initiated in [8] was further pursued by others [11–13] and a *duality relation* for general bipartite two-dimensional systems was obtained by Jaeger *et al.* [14] and Englert [15], connecting the visibility $\mathcal{V}$ (amplitude) of the pattern produced by the interfering object to the amount of which-path information $\mathcal{D}$ [see Eq. (6)] captured by the detector:

$$\mathcal{V}^2 + \mathcal{D}^2 \leqslant 1. \qquad (2)$$

This analysis was then extended to the context of imperfect detection and quantum erasure [16], as well as more than two interferometric paths [17]. After the turn of the century, it was also shown that the *a posteriori* distinguishability of the paths, $\mathcal{D}$, contains two parts to it: the *a priori* predictability and the genuine quantum correlations, as quantified by the concurrence entanglement measure. Equation (2) was alternatively cast as a triality and generalized to bipartite systems of any dimension [18,19].

Originally, Englert argued that the duality relation in Eq. (2) was independent of Heisenberg's uncertainty principle. Since then, however, a new entropic formulation of the

---
*pedro.cruz@inl.int
†On leave from Departamento de Física Aplicada, Universidade de Alicante, 03690 San Vicente del Raspeig, Spain.

latter has been discovered, allowing its equivalence to the former to be established for multipath interferometers [20,21]. Other wave-particle duality relations based on different information quantifiers and applicable to Hilbert spaces of arbitrary dimension have also recently been introduced [22,23].

Complementarity, and the controlled depletion of interference by including which-path detectors in the experimental setup, have been observed in very ingenious experiments probing diverse interferometer-detector systems, including atom-photon [24,25], photon-photon [26–31], electron-electron [32–34], coupled atomic nuclei in molecules [35–37], and superconducting circuits [38–40]. Wave-particle duality has also been observed in optical interferometers in which the path information is not stored in a subsystem but carried by an additional degree of freedom of the photons along their path [41,42].

Today, with the advent of transmon quantum computers and IBM Quantum (IBM Q) cloud services [43], new tests of complementarity are being performed. This platform has been used to carry out which-path experiments with multipath interferometers and extract quantifiers of visibility and distinguishability using an unambiguous state discrimination detector readout in [44]. The triality relation in [19] was tested on two qubits with a tomography-based circuit in [45]. Moreover, a state tomography approach was also used in [46] to test different complementarity relations obtained from density matrix properties of random states of up to three qubits.

Here, we study two classes of which-path experiments using simple two-qubit circuits to implement a symmetric two-way interferometer coupled to a single detector. First, we design a setup in which the controlled transfer of path information to the detector subsystem depletes interference on the probed subspace, testing Eq. (2) by quantifying distinguishability via minimum-error state discrimination measurements. Then, we consider a nondelayed quantum-eraser configuration in which the detector readout basis destroys path information allowing interference to be recovered. We experiment on a superconducting quantum processor made available online by IBM Q. In this architecture, microwave pulses drive arbitrary single-qubit rotations and entanglement of fixed-frequency transmons. We pay close attention to fine details in the experimental data to accurately interpret results and draw their fundamental consequences.

The rest of this paper is organized as follows. In Sec. II, we introduce the abstract quantum circuits that implement optimal which-path detection and quantum erasure on any gate based quantum computer. In Sec. III, we present the experimental platform and the physical circuits, describe the methodology, and analyze experimental results. In Sec. IV, reassessment of the physical operations and model evaluation are performed to understand the encountered deviations from the theory. Finally, Sec. V discusses the main takeaways and presents our conclusions.

II. WHICH-PATH QUANTUM CIRCUITS

Our starting point is the quantum circuit representation of a Mach-Zehnder interferometer (MZI) [47] [see Fig. 1(a)]. The first Hadamard gate prepares an equal superposition of the two qubit states, $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, very much like the first beam
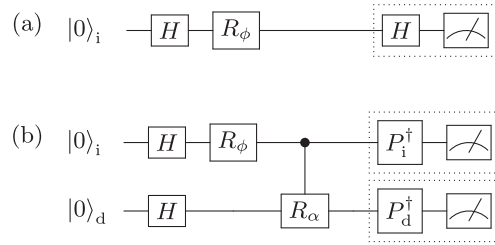

FIG. 1. (a) Single-qubit circuit equivalent to a Mach-Zehnder interferometer. (b) Same setup, coupling a second qubit, q_d, to the interferometer qubit, q_i, to act as a which-path detector. The readout basis is chosen by defining $P_i^\dagger$ and $P_d^\dagger$ before the Z basis measurement. Setting $P_i^\dagger = H$ and varying the value of α, interference on q_i can be depleted, independently of reading out q_d. Setting $P_d^\dagger = O_\alpha^\dagger$ maximizes which-path information retrieval. Making $P_d^\dagger = \mathbb{1}$ instead erases path information rendering detection outcomes useless while allowing full-visibility interference patterns on q_i to be recovered. Both measurements occur simultaneously. A delayed-choice experiment would probe the detector after the interferometer.

splitter provides two alternatives in the MZI. The phase gate R_ϕ introduces a relative phase between the two alternatives, yielding the state in Eq. (1) with $c_0 = c_1 = \frac{1}{\sqrt{2}}$. The second H gate, equivalent to the second beam splitter in the MZI, is used to measure the X operator with a readout in the Z basis. The output state $|\psi'\rangle = H|\psi\rangle$ on q_i satisfies

$$\langle\psi'|Z|\psi'\rangle = \langle\psi|H^\dagger ZH|\psi\rangle = \langle\psi|X|\psi\rangle = \cos\phi. \quad (3)$$

Thus, by measuring $|\psi'\rangle$ in the computational basis we obtain the expectation $\langle\psi|X|\psi\rangle$ and detect interference, probing thereby the linear superposition and the added relative phase. In this setup, we gain no information about which path, i.e., state $|0\rangle$ or $|1\rangle$, was taken by the qubit before the X basis measurement.

The which-path detector is implemented by adding a second qubit, q_d, to the circuit [Fig. 1(b)]. The controlled phase gate introduces a relative phase α in the joint interferometer-detector state when both qubits are in state $|1\rangle$. The wave function of the two-qubit system after this operation, in the basis $|xy\rangle = |x\rangle_i \otimes |y\rangle_d$, is

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |\delta_0\rangle + e^{i\phi}|1\rangle \otimes |\delta_1\rangle), \quad (4)$$

where $|\delta_x\rangle$, $x \in \{0, 1\}$, denotes the state of the detector qubit conditioned to the path $|x\rangle$ taken by the first one:

$$|\delta_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$

$$|\delta_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\alpha}|1\rangle). \quad (5)$$

The *distinguishability* $\mathcal{D}$ of these states is quantified by the trace distance of their density matrices, yielding

$$\mathcal{D} \equiv \frac{1}{2}\||\delta_0\rangle\langle\delta_0| - |\delta_1\rangle\langle\delta_1|\|_T \quad (6)$$

$$= \sqrt{1 - |\langle\delta_0|\delta_1\rangle|^2} = \sqrt{\sin^2\frac{\alpha}{2}}, \quad (7)$$

where we have used

$$\langle \delta_0 | \delta_1 \rangle = \tfrac{1}{2}(1 + e^{i\alpha}). \tag{8}$$

Thus, $\mathcal{D}$ is maximal for $\alpha = \pi$ and null for $\alpha = 0$.

Since we want to measure the detector to correctly identify the path taken by the first qubit, we should maximize the probability of success. This is accomplished by obtaining the maximum which-path information leaked with a measurement in the optimal basis [48]. For that, we use the eigenvectors of $|\delta_0\rangle \langle \delta_0| - |\delta_1\rangle \langle \delta_1|$ to build an operator O_α, which we conjugate transpose to obtain

$$O_\alpha^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & ie^{-\frac{i\alpha}{2}} \\ 1 & -ie^{-\frac{i\alpha}{2}} \end{pmatrix}. \tag{9}$$

Making $P_d^\dagger = O_\alpha^\dagger$ in the circuit of Fig. 1(b) transforms the reduced density matrix over q_d as $\hat{\rho}_x = O_\alpha |\delta_x\rangle \langle \delta_x| O_\alpha^\dagger$, which measured in the computational basis yields the outcome y on the detector with maximal probability of successfully identifying the correct path $|x\rangle$ on q_i:

$$p_{\text{succ}} = \tfrac{1}{2}[p(0_d|\hat{\rho}_0) + p(1_d|\hat{\rho}_1)] = \tfrac{1}{2}(1 + \mathcal{D}). \tag{10}$$

The significance of the distinguishability is therefore that it bounds the maximum possible success probability. For $\alpha = \pi$, $\hat{\rho}_0 = \mathbb{1} - \hat{\rho}_1 = |0\rangle \langle 0|$, allowing for fully determined path information. If we run the circuit in Fig. 1(b) with $P_i^\dagger = \mathbb{1}$ and $P_d^\dagger = O_\alpha^\dagger$ we can obtain $\mathcal{D}$ empirically from

$$p_{\text{succ}} = \tfrac{1}{2} p(0_d|0_i) + \tfrac{1}{2} p(1_d|1_i), \tag{11}$$

where $p(y|x) = p(xy)/\sum_y p(xy)$ stands for the conditional probability and $p(xy)$ is the measurement probability of the $|xy\rangle$ output state.

A. Which-path: Interference vs detection

The interference pattern produced by the first qubit in the circuit of Fig. 1(b) with $P_i^\dagger = H$ is obtained from

$$\langle X \rangle \equiv \text{Tr}\, \hat{\sigma} X, \tag{12}$$

where $\hat{\sigma}$ is the reduced density matrix on q_i before the basis change operation, associated to the two-qubit state from Eq. (4) tracing over the states of q_d:

$$\hat{\sigma} \equiv \text{Tr}_{q_d} |\Psi\rangle\langle\Psi|. \tag{13}$$

After straightforward algebra, we obtain

$$\langle X \rangle = \text{Re}\,\{e^{i\phi}\langle \delta_0|\delta_1\rangle\} = \frac{\cos\phi + \cos(\phi+\alpha)}{2}. \tag{14}$$

Thus, it is apparent that the contrast of the interference pattern along ϕ is controlled by the overlap of the possible states of the detector qubit. When this overlap, and thereby contrast, is maximal ($\alpha = 0$), which-path information is null. On the other hand, for $\alpha = \pi$ there is no interference and maximal which-path information is stored on the detector, allowing for its readout to unambiguously identify the path taken by the first qubit. In other words, it is by maximally entangling the two-qubit system ($\alpha = \pi$) that we get q_i to display classical particlelike statistics with no interference visibility, regardless of reading out the detector.

Experimentally, this interference pattern is obtained from $\langle X \rangle = p(0_i) - p(1_i)$, approximating the output probabilities
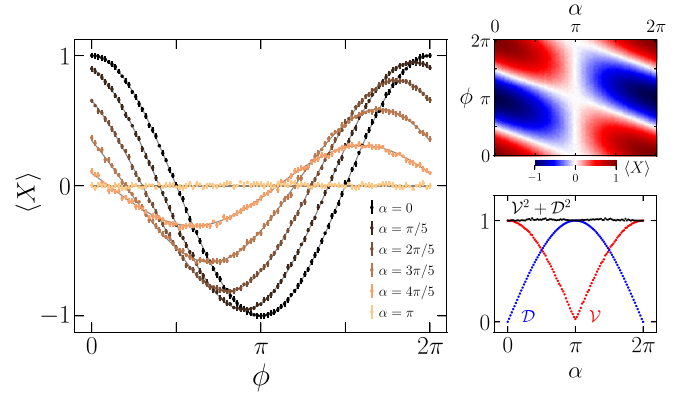

FIG. 2. In-silico pure-state simulations obtained with two families of circuits of the type of Fig. 1(b), executed with 8192 shots for each of 10 201 pairs of values $\alpha, \phi \in [0, 2\pi]$. The first family consists of the circuit with $P_i^\dagger = H$ and $P_d^\dagger = O_\alpha^\dagger$. The top right panel shows the measured $\langle X \rangle$ on q_i for all these circuits. The left side panel illustrates a subset of the data for six different values of α: actual results (dots with $\pm 2\sigma_{\langle X \rangle}$ errorbar) are compared against the theoretical expectation given by Eq. (14) (solid lines). The third panel shows the results for visibility and distinguishability; data in blue are obtained from the second family of circuits, where the circuit in Fig. 1(b) is executed with $P_i^\dagger = \mathbb{1}$ and $P_d^\dagger = O_\alpha^\dagger$, and Eqs. (10) and (11) are employed to extract $\mathcal{D}$.

with the relative frequencies of the readouts. This process is subject to shot noise, the standard deviation of which scales with $\frac{1}{\sqrt{S}}$, where S is the number of readouts, or shots. In Fig. 2 we show the results of an in-silico simulation generated by ramping the phase ϕ in the first qubit, for several values of the control parameter α. With 8192 shots per circuit, the curves traced by the data are in perfect agreement with Eq. (14), and the 95.4% confidence interval contained by $\langle X \rangle \pm 2\sigma_{\langle X \rangle}$ is almost imperceptible. As we vary α from zero to π, the amplitude of the interference pattern gradually vanishes.

To summarize these observations we define the interference visibility and compute it with Eq. (14):

$$\mathcal{V} \equiv \frac{\langle X \rangle_\alpha^{\max}(\phi) - \langle X \rangle_\alpha^{\min}(\phi)}{2 + \langle X \rangle_\alpha^{\max}(\phi) + \langle X \rangle_\alpha^{\min}(\phi)} = \sqrt{\cos^2\left(\frac{\alpha}{2}\right)}. \tag{15}$$

Using Eq. (8), visibility is related to the overlap

$$\mathcal{V} = |\langle \delta_0|\delta_1\rangle|, \tag{16}$$

and therefore obeys the well-known [14,15] result

$$\mathcal{V}^2 + \mathcal{D}^2 = 1. \tag{17}$$

This provides a mathematical description of the complementarity relation between distinguishability and visibility: imprinting which-path information of the first qubit into the second one gradually destroys interference. In the next section, we address the question of whether the interference is actually destroyed or if we could recover it by interrogating the detector differently.

B. Quantum erasure

As we have seen in the previous section, the interference pattern of the first qubit is compromised if the which-path

information is stored in the detector, even if this second qubit is not measured. We now consider the setup of Fig. 1(b) when the state of the detector is instead measured with $P_d^\dagger = \mathbb{1}$, keeping $P_i^\dagger = H$.

This detector readout destroys the which-path information by collapsing the state directly into the Z basis. Given that both $|0\rangle$ and $|1\rangle$ have equal weight on that projection, path information is lost regardless of the value of α. To see how the irreversible destruction of the which-path information can restore interference, we write the quantum state just before readout as

$$|\Psi_{\text{out}}\rangle = \frac{1 + e^{i\phi}}{2\sqrt{2}} |00\rangle + \frac{1 + e^{i\phi + i\alpha}}{2\sqrt{2}} |01\rangle$$
$$+ \frac{1 - e^{i\phi}}{2\sqrt{2}} |10\rangle + \frac{1 - e^{i\phi + i\alpha}}{2\sqrt{2}} |11\rangle. \quad (18)$$

With that, we have the following probabilities for all the possible measurement outcomes:

$$\begin{aligned} p(00) &= \tfrac{1}{4}(1 + \cos\phi), \\ p(01) &= \tfrac{1}{4}[1 + \cos(\phi + \alpha)], \\ p(10) &= \tfrac{1}{4}(1 - \cos\phi), \\ p(11) &= \tfrac{1}{4}[1 - \cos(\phi + \alpha)], \end{aligned} \quad (19)$$

where $p(xy) = |\langle\Psi_{\text{out}}|xy\rangle|^2$. Let us then look at the expectation value of the interferometer qubit in the X basis, broken down according to the outcome of the second qubit, which will be $|0\rangle$ or $|1\rangle$ with equal probability. Conditioning on these events we get

$$\langle X_0 \rangle \equiv p(0_i|0_d) - p(1_i|0_d) = \cos\phi, \quad (20)$$

$$\langle X_1 \rangle \equiv p(0_i|1_d) - p(1_i|1_d) = \cos(\phi + \alpha). \quad (21)$$

It is clear that by discriminating q_i readouts following the guidance of the detector, two $\mathcal{V} = 1$ interference patterns are obtained, shifted by α. It is only when these measurements are aggregated, disregarding the detector output and protecting thereby path information, that we return to the situation where the interference visibility is compromised by α, recovering the result in Eq. (14):

$$\tfrac{1}{2}\langle X_0 \rangle + \tfrac{1}{2}\langle X_1 \rangle = \langle X \rangle. \quad (22)$$

Note that if we perform the same break-down procedure with $P_d^\dagger = O_\alpha^\dagger$ instead we cannot recover full-visibility interference patterns like these. Nature seems to know that our measurement projection has already obtained some which-path information, therefore it does not let us observe the complete wavelike behavior of the system. Similar approaches to quantum erasure have been described in [25,49,50].

III. EXPERIMENTAL RESULTS

We now discuss the experimental testing of the circuits in the previous section. For that matter, we took advantage of the IBM Q platform [43], which provides remote access to digital quantum processors based on superconducting fixed-frequency transmon qubits coupled via the cross-resonance gate. In particular, we used the Qiskit open-source software package [51] to prepare several experiments for execution in *ibmq_toronto*, a 27-qubit *Falcon r4* processor with a heavy-hexagon qubit lattice [52] achieving *quantum volume* 32.

The core of a transmon is formed by two superconducting islands coupled through two Josephson junctions (JJs). The relevant dynamical quantities of the JJs are the electron imbalance across the junction and the difference of the phases of the superconducting order parameter, which for all practical purposes can be considered conjugate variables. The qubit states $|0\rangle$ and $|1\rangle$ are encoded in the two lowest-energy states of the Hamiltonian describing this system. The control and readout of the JJs is implemented by means of their capacitive coupling to the electromagnetic modes confined in a transmission line resonator. For details on the working principles of the device we refer the reader to the original papers [53–55] and recent reviews [56,57].

IBM Q processors have been previously used to test other fundamental aspects of quantum mechanics, such as Bell and Mermin inequalities [58,59]. Quantum computing systems based on transmons are also being developed by others, such as Rigetti [60], Google Quantum AI [61], and Quantum Inspire [62].

A. Methodology

The methodology used to test any of the circuits in Fig. 1 involved two steps: first, the estimation of the relevant observables from a sample of S shots for each of C values for the variables of the circuit being tested. At a repetition time of $500\,\mu$s per shot, the total processing time of an experiment adds up to about $500\,CS\,\mu$s. The set of circuits comprising a full experiment was executed in batches of 896. Four additional measurement error mitigation [63–65] circuits were included in each batch, except for the experiment with the circuit in Fig. 1(a). Within each batch, one shot of each circuit was executed sequentially before moving on to each of the remaining shots. Batches ran one after another until completion of all the executions required to carry out an experiment. Each circuit in a circuit family is defined by its ϕ and α values. Hence, the order in which circuits are executed sets the order in which this space is probed. Because of the second step of the methodology, this space was probed uniformly at random.

In the second step, the previous C estimates were employed to find a global experimental model $f(\phi, \alpha)$ through a curve-fitting procedure minimizing

$$\chi^2 = \sum_{j=1}^{C} \left(\frac{z_j - \hat{z}_j}{\sigma_j} \right)^2, \quad (23)$$

where $\hat{z}_j \equiv f[(\phi, \alpha)_j]$, and the standard error σ_j quantifies the dispersion associated with each estimate $\hat{z}_j$. Although the measurement error mitigation procedure is used to infer sources of error during data analysis, all models below are fitted to raw, unmitigated data.

In order to set up the model $f(\phi, \alpha)$ to fit the data, the theoretical expression $g(\phi, \alpha)$ obtained for the case of ideal unitary execution of a given quantum circuit is modified with two additional features. First, one fit parameter is introduced
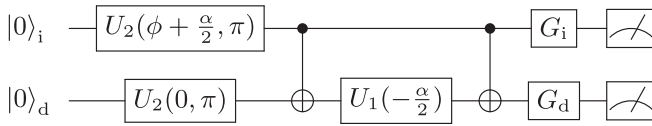
FIG. 3. Actual physical circuit implemented to reproduce Fig. 1(b) in terms of the elementary gateset of the device. Final gates are defined for each readout basis according to the correspondence: $P_i^\dagger = H \Rightarrow G_i = U_2(0, \pi)$, $P_i^\dagger = \mathbb{1} \Rightarrow G_i = \mathbb{1}$, $P_d^\dagger = O_\alpha^\dagger \Rightarrow G_d = U_2(0, \frac{3\pi}{2})$, and $P_d^\dagger = \mathbb{1} \Rightarrow G_d = U_1(\alpha/2)$.

per circuit variable, replacing

$$\phi \rightarrow \phi + \Theta_1,$$
$$\alpha \rightarrow \alpha + \Theta_2, \quad (24)$$

in the $g(\phi, \alpha)$ model to account for calibration errors in the physical gates. These two additional parameters improve the modeling of the experimental data and their seemingly *ad hoc* introduction will be justified in Sec. IV. Second, since the quantum processor is not perfectly isolated from the environment, the dynamics over this subspace is nonunitary. To accommodate for the existence of a classical mixture with identically distributed noise on all the output distributions from an experiment, the pure state expression is scaled by η and a bias parameter ε is introduced. The final experimental model carries at most four parameters and is written as

$$f(\phi, \alpha) = \eta\, g(\phi, \alpha) + \varepsilon. \quad (25)$$

After completion of the curve-fitting procedure, the autocorrelation of the residuals is checked in the order of the execution of the circuits. If observed, it is associated with drifts in the calibration and coherence properties of the device during the experiment. The attribution is made possible by the execution protocol whereby the (ϕ, α) space is sampled at random, since it removes the correlation component coming from the possible use of an incorrect model to fit the data. In order to detect autocorrelation without making any strong assumptions about the generating error process, the residuals are divided into positive or negative at their average value and the nonparametric single-sample Wald-Wolfowitz runs test [66] is performed on the sequence. Evidence for autocorrelation is found if the test rejects the null hypothesis for errors to be independent and identically distributed.

All required circuits in an experiment must be executed on the same two-qubit pair and within a single calibration round of the system to maintain calibration error consistency across the tested variable space. For this reason, the random sampling protocol brings an additional benefit. If the calibration of the device changes before sampling all the intended variable values, it is still possible to fit the model with the already obtained values uniformly spread over the variable space. Furthermore, if mild environmental drifts in the coherence properties do occur, randomly sampling (ϕ, α) allows filtering outliers uniformly in this space.

At the time of the experiment, IBM Q hardware uses the controlled-NOT (CNOT) gate together with a fixed set of parametrized single-qubit gates to implement any quantum circuit. For this reason, the circuits described in Fig. 1 were converted to the native operations as represented in Fig. 3.
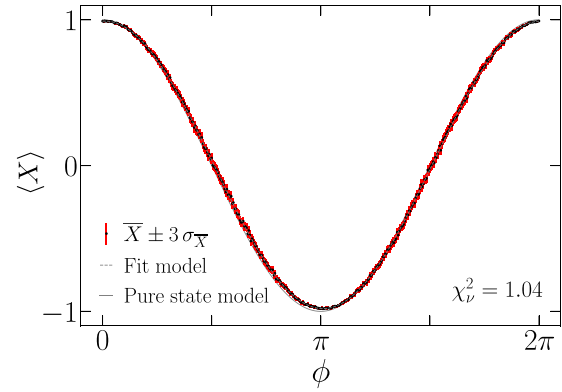

FIG. 4. Experimental results for the circuit in Fig. 1(a). The $\overline{X}$ estimations are represented with almost imperceptible $\pm 3\sigma_{\overline{X}}$ confidence intervals and adjust very well to the properly constructed Eq. (25) model, as measured by χ_ν^2. This experimental model is also very close to the pure state model, Eq. (3).

Only two single-qubit physical gates were used:

$$U_1(y) = \begin{pmatrix} 1 & 0 \\ 0 & e^{iy} \end{pmatrix},$$

$$U_2(x, y) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -e^{iy} \\ e^{ix} & e^{i(x+y)} \end{pmatrix}. \quad (26)$$

One might notice this circuit can be simplified for some values of ϕ and α. However, such simplifications should not be done experimentally since we must ensure all the results are obtained from the same number of operations and execution time. This guarantees adherence of the whole dataset to the same experimental model.

Finally, the distribution of residuals is also inspected in the (ϕ, α) space to assess any pattern indicating the use of an incorrect model to fit the data. Model adequacy issues are discussed further in Sec. IV.

B. Two-way interferometer

Our first experimental results were obtained for the setup in Fig. 1(a), taking $S = 8192$ shots for each of $C = 401$ evenly spaced values of $\phi \in [0, 2\pi]$ (see Fig. 4). This one-qubit circuit is used as a control experiment to confirm that an interference pattern is observed. For each ϕ, we registered the readout X_j, collecting then all readouts to compute the relative frequencies $(\overline{p})$ of the two possible outcomes and estimate $\langle X \rangle$ from the mean $\overline{X} = \overline{p}(0) - \overline{p}(1)$ with standard error $\sigma_{\overline{X}} = s/\sqrt{S}$, where $s^2 = \sum_{j=1}^{S} \frac{(X_j - \overline{X})^2}{S-1}$ is the sample variance.

Based on Eq. (3), these data were modeled with an experimental model $\overline{X} = \eta \cos(\phi + \Theta_1) + \varepsilon$. For the chosen S and C, the minimization procedure finds very good agreement between the data and the model, scoring $\chi_\nu^2 \equiv \chi^2/\mathrm{dof} \approx 1.04$, where $\mathrm{dof} = 398$. The fit parameters, and their respective one standard deviation errors, are determined with very good precision as

$$\eta = 0.985\,81 \pm 0.000\,24,$$
$$\varepsilon = 0.005\,19 \pm 0.000\,22,$$
$$\Theta_1 = -0.032\,19 \pm 0.000\,60. \quad (27)$$

The residuals are randomly distributed across the full range of the independent variable, as evidenced by the runs test. Yielding a p value of 0.61, it fails to reject the null hypothesis at a significance level of 0.05. Therefore, we attest robustness of the system to the quantum fluctuations that could have had an effect on the circuit outputs during the ≈ 30 min taken by the experiment.

In Fig. 4 we plot experimental results against theory. Comparing the fitted parameters with Eq. (3), η shows an amplitude drop most probably due to readout errors but in principle also to the open nature of the system's evolution, while the value for Θ_1 represents systematic calibration errors. The asymmetry introduced by the ε bias is consistent with both energy dissipation by which the qubit decays from state $|1\rangle$ to $|0\rangle$ and different readout error rates for each of these states.

C. Optimal which-path detection

This experiment requires the computation of two families of circuits. One applies $P_i^\dagger = H$ to estimate $\langle X \rangle$ on q_i and obtain $\mathcal{V}$, whereas the other is used to find $\mathcal{D}$ by making $P_i^\dagger = \mathbb{1}$. Both types of circuits carry $P_d^\dagger = O_\alpha^\dagger$ and were executed with $S = 8192$ shots on $C = 10\,201$ pairs of (ϕ, α) values sampled on a regular 101×101 square grid in $\phi, \alpha \in [0, 2\pi]$. The full experiment therefore included $20\,402$ circuits in 23 batches and took ≈ 23 h to complete. The batches were prepared with both families of circuits being executed at the same values of the circuit variables in alternating order, shot by shot. The (ϕ, α) space was sampled uniformly at random.

The results are analyzed separately for each circuit family. For the first one, we obtained $\overline{X}$ and $\sigma_{\overline{X}}$ on the first qubit as in the previous section. The experimental model approximating Eq. (14) has four parameters, and the fitting procedure converges with $\chi_\nu^2 \approx 2.19$. The fit parameters, with one standard deviation error, are

$$\eta = 0.930\,39 \pm 0.000\,17, \quad \Theta_1 = -0.037\,04 \pm 0.000\,26,$$
$$\varepsilon = 0.037\,00 \pm 0.000\,09, \quad \Theta_2 = 0.005\,12 \pm 0.000\,37. \tag{28}$$

Once again, we see no large shifts in calibration with $\Theta_{1,2}$ close to zero. This time, however, we observe a slightly larger drop in η as well as a larger ε bias in comparison with the one-qubit circuit. Assuming the model to be correct, causes might include larger measurement error in the physical qubit used to encode q_i, as well as execution errors introduced by the two-qubit gate, not present in the previous circuit.

From the results obtained with the second family of circuits, distinguishability $\overline{\mathcal{D}}$ is estimated using the relative frequencies of successfully determining the path on q_i. We follow Eqs. (10) and (11) to obtain

$$\overline{\mathcal{D}} = \frac{\overline{p}(00)}{\overline{p}(00) + \overline{p}(01)} + \frac{\overline{p}(11)}{\overline{p}(10) + \overline{p}(11)} - 1, \tag{29}$$

and the associated standard error

$$\sigma_{\overline{\mathcal{D}}} = \sqrt{\frac{\overline{p}(00)\,\overline{p}(01)}{S[\overline{p}(00) + \overline{p}(01)]^3} + \frac{\overline{p}(10)\,\overline{p}(11)}{S[\overline{p}(10) + \overline{p}(11)]^3}}. \tag{30}$$
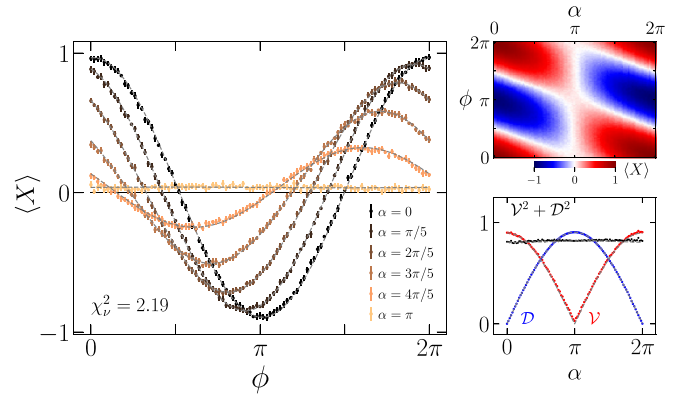

FIG. 5. Experimental results for the two families of circuits in Fig. 1(b) in which the detector qubit is measured in the optimal basis and the interferometer qubit is measured both in the X and the Z basis. Experimental data for q_i on the left side panel correspond to $\overline{X} \pm 2\sigma_{\overline{X}}$. Each of these panels is equivalent to those described in Fig. 2.

We proceed with the two-dimensional fit of all the estimations with a three-parameter model obtained from Eq. (7). The procedure yields $\chi_\nu^2 \approx 1.34$, providing

$$\eta = 0.902\,16 \pm 0.000\,27, \quad \Theta_2 = -0.002\,24 \pm 0.000\,29,$$
$$\varepsilon = 0.002\,53 \pm 0.000\,23. \tag{31}$$

Thus, we see no relevant shifts in calibration but a noticeable drop in η attributable to the same causes as those for the first family of circuits, with the addition of measurement errors on q_d. The cause for the ε bias is not as straightforward to single out as before.

We compute the visibility and distinguishability curves from the two fitted models, obtaining then $\mathcal{V}^2 + \mathcal{D}^2$. These are represented by the dashed lines in the bottom right panel of Fig. 5 and essentially match the values obtained from the experimental data (in color). However, while the dashed visibility curve obtained from the fitted model reaches $\mathcal{V} = 0$ by construction, when we look directly at the experimental visibility values (red dots) we find $\mathcal{V}_{\min} = 0.030 \pm 0.004$ at $\alpha = \pi$ and $\mathcal{V}_{\max} = 0.917 \pm 0.004$ at $\alpha = 1.94\pi$, where the standard deviation error is obtained with $10\,000$ bootstrap samples. As for the blue experimental data points, for $\overline{\mathcal{D}}$, they represent the average of the values for $\overline{\mathcal{D}}$ obtained with all the different circuits for that α value and the different ϕ.

Comparing this panel with that of Fig. 2, the loss of information is apparent for both the interference visibility $\mathcal{V}$ and the path distinguishability $\mathcal{D}$, lowering the amplitudes of Eqs. (7) and (15), and modifying Eq. (17) to $\mathcal{V}^2 + \mathcal{D}^2 < 1$. This comes about due to the mixing of the final pure state distributions at the output with the noise distribution. In this case, we conclude the noise distribution mostly arises from the errors occurring during the readout process by verifying that the measurement error mitigation protocol closely restores $\eta \approx 1$ in both models, and thus $\mathcal{V}^2 + \mathcal{D}^2 \approx 1$. Since the circuits are so shallow, decoherence does not really play a role in the loss of quantum information here.

The runs tests for the residual errors in the order of execution of the first and the second family of circuits yield p
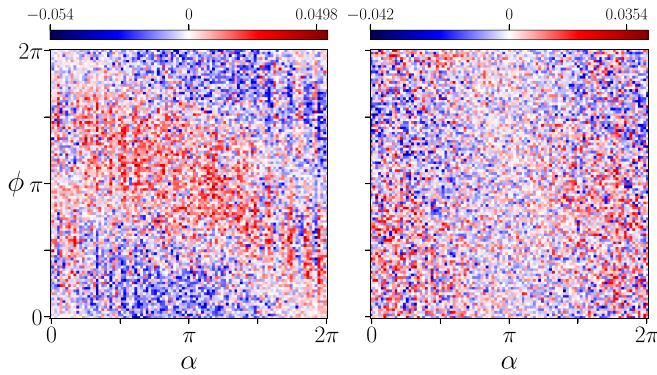
FIG. 6. Residual error maps $z_j - \hat{z}_j$ for the second experiment. On the left, the residuals of the fit to the $\overline{X}$ model of the first family of circuits are shown. On the right, those of the fit to the $\overline{\mathcal{D}}$ model, obtained with the second family of circuits, are shown. The colorbars are marked with the lowest and highest recorded values.

values numerically close to zero. Therefore, we reject the null hypothesis at the 0.05 significance level and have evidence for autocorrelation in the data. This signals some slight instability in the physical properties of the system, which we examined only to find no significant impact in the presented analysis. Both experimental models are still able to closely fit the data, with residual standard errors, RSE $\equiv \sqrt{\frac{1}{\text{dof}} \Sigma_{j=1}^{C} (z_j - \hat{z}_j)^2}$, of 0.012 74 in the first case and 0.009 46 in the second case. Scores $R^2 \approx 0.999\,26$ and $0.998\,88$ are obtained for each model, respectively.

More importantly, we find a small but systematic deviation from the theory, mostly in the results from the first family of circuits, but also on those from the second. Regarding the former, the anomaly is already suggested by the fact that $\chi_\nu^2 \approx 2.19$, but we confirm it by visual inspection of the plot of the fit residuals as a function of ϕ and α. While an adequate fit would result in a random distribution of residual errors in this space, the left panel of Fig. 6 clearly shows a pattern. As for the latter model, the pattern in the right panel of Fig. 6 is less visible, in agreement with the lower $\chi_\nu^2 \approx 1.34$.

Even if of small magnitude, it becomes clear with the resolution of the data that a functional dependence in ϕ and α is missing in the experimental models to be able to fully explain the data. In Sec. IV we reanalyze these results in light of an improved theoretical model of the entangling operation.

D. Quantum eraser

Let us now discuss the third experiment in which we ran the circuit in Fig. 1(b) with $P_i^\dagger = H$ and $P_d^\dagger = \mathbb{1}$. As in the previous section, we choose $S = 8192$ and $C = 10\,201$ pairs of (ϕ, α) values sampled on a regular 101×101 square grid in $\phi, \alpha \in [0, 2\pi]$. The experiment's 10 201 circuits in 12 batches took ≈ 12 h to complete.

As discussed above, our goal is to demonstrate how measuring q_d in a projection that cannot access the gathered information renders which-path detection useless while recovering full wavelike behavior on q_i. To do so, we separate the readouts on the latter qubit as conditioned by the output $y \in \{0, 1\}$ on the detector, to estimate $\langle X_y \rangle$ in Eqs. (20) and
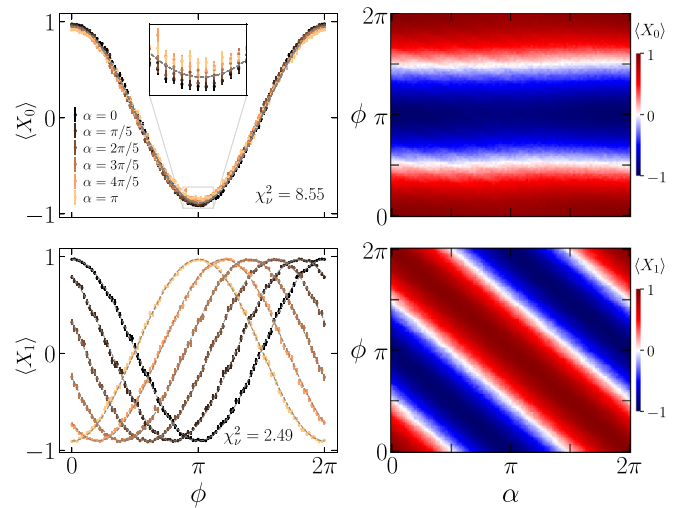

FIG. 7. Experimental results for the quantum eraser circuit. The expectation values of the X operator on the first qubit broken down by the outcome of the detector are shown, with $\overline{X_0} \pm 2\sigma_{\overline{X_0}}$ ($\overline{X_1} \pm 2\sigma_{\overline{X_1}}$) in the top (bottom) row. Right-side panels plot the full set of results for the tested (ϕ, α) values, while the left-side panels show a selection for six different α.

(21) from

$$\overline{X_y} = \frac{\overline{p}(0y) - \overline{p}(1y)}{\overline{p}(0y) + \overline{p}(1y)}.\tag{32}$$

The associated standard error is obtained with

$$\sigma_{\overline{X_y}} = \sqrt{\frac{4\,\overline{p}(0y)\,\overline{p}(1y)}{S[\overline{p}(0y) + \overline{p}(1y)]^3}}.\tag{33}$$

The results are shown in Fig. 7, for $\overline{X_0}$ in the top row, and $\overline{X_1}$ in the bottom row. Qualitatively, the full-visibility interference patterns expected in theory are recognizable for both $\overline{X_0}$ and $\overline{X_1}$. To make the analysis quantitative, we first fitted $\overline{X_0}$ data to the expected experimental model obtained from Eq. (20). We find

$$\eta = 0.917\,40 \pm 0.000\,11, \quad \Theta_1 = -0.031\,18 \pm 0.000\,20,$$
$$\varepsilon = 0.033\,61 \pm 0.000\,09,\tag{34}$$

with $\chi_\nu^2 \approx 8.55$ and RSE $= 0.030\,83$. The other subset of the data, for $\overline{X_1}$, was targeted by an experimental model based on Eq. (21), achieving $\chi_\nu^2 \approx 2.49$, RSE $= 0.018\,85$, and parameters

$$\eta = 0.937\,65 \pm 0.000\,11, \quad \Theta_1 = -0.005\,65 \pm 0.000\,10,$$
$$\varepsilon = 0.033\,79 \pm 0.000\,09, \quad \Theta_2 = -0.005\,65 \pm 0.000\,10.\tag{35}$$

As can be seen, the values for the fit parameters are very precise in both cases. Furthermore, the runs tests, computed on the chronologically ordered sequence of residuals for both fits separately, fail to reject the null hypothesis for randomness at a significance level of 0.05. Thus, during the execution of this two-qubit circuit, the device had very stable performance.

However, χ_ν^2 scores show larger deviations from 1, indicating some cause for concern. In fact, the plot of residuals

as a function of (ϕ, α) shows once again a pattern for both the fits with $\overline{X_0}$ and $\overline{X_1}$ data. The slight disagreement with theory is especially marked for the $\overline{X_0}$ results and can even be seen in the inset plot on the top-left panel of Fig. 7. A weak dependence in α is apparent, and the visibility of the interference slightly decreases as α approaches π. This suggests that an imperfection in the implementation of the controlled-R_α gate leaks a small amount of information to the detector qubit which the readout in the Z basis does not destroy, hence slightly reducing visibility on the top qubit. We postpone the analysis of this problem to Sec. IV.

Finally, to understand how effective the detector qubit is at discriminating the paths over the interferometer, we measured distinguishability in this readout basis. To that end, we carried out yet another independent experiment. This fourth one translates $P_i^\dagger = P_d^\dagger = \mathbb{1}$ in the language of Fig. 1(b) to $G_i = \mathbb{1}$ and $G_d = U_1(\frac{\alpha}{2})$ in Fig. 3. The same grid with $C = 10\,201$ pairs of (ϕ, α) values is probed, as well as the same number of shots $S = 8192$.

Distinguishability is again measured from Eqs. (29) and (30). Since p_{succ}, as computed with Eq. (11), is expected to be 0.5 with the detector readout directly on the states in Eq. (5), theory predicts the *measured* distinguishability [16] to be

$$\mathcal{D}_m = 0, \tag{36}$$

independently of ϕ and α. Note that for the previous case of the detector measurement in the optimal basis, the measured distinguishability, as obtained with Eqs. (10) and (11), matches the actual distinguishability $\mathcal{D}$ in Eq. (6) precisely because of the optimal measurement.

Given $\mathcal{D}_m = 0$, we used a simple model to fit the results in which only the ε parameter from Eq. (25) is present. The procedure returns $\varepsilon = -0.003\,70 \pm 0.000\,11$ with $\chi_\nu^2 \approx 2.06$ and RSE = 0.015 81. Inspection of the residuals as a function of ϕ and α once again reveals a smooth pattern. Therefore, measured distinguishability is neither zero nor constant across the (ϕ, α) space, confirming the conclusion from the previous experiment of a failure to completely erase the quantum information and produce entirely wavelike behavior. This experiment also had no significant environmental perturbations during execution, as evidenced by the p value of 0.96 for the runs test on the ordered sequence of residuals, failing to reject the null at a significance level of 0.05.

IV. REASSESSING THE CIRCUIT MODELS

As we have seen in the previous two-qubit experiments, results were frequently slightly off the theoretical predictions put forward in Sec. II. These anomalies could not be removed with measurement error mitigation, indicating they do not stem from the readout process. We now take into account the systematic deviations between the ideal gates in the presented circuits and the actual hardware-level operations to rederive the theoretical models. We analyze how this allows for a better fit of the experimental results and thus accounts for the moderate, but systematic, deviations from the theory.

A. Accounting for coherent errors

Take again the operations in the physical circuit of Fig. 3. With the exception of the gates parametrized by ϕ and α, all other single-qubit gates are applied with fixed input values. Depending on the choices for G_i and G_d, any of the circuits will only have U_2 gates on the top qubit, and both U_1 and U_2 gates on the bottom qubit. Let us first model single-qubit gate errors (SQGEs) by adding a bias parameter to each rotation angle of the U_2 gate on the interferometer qubit,

$$U_2^{(q_i)}(x, y) \equiv U_2(x + \theta_1, y + \theta_2), \tag{37}$$

and other independent bias parameters to both types of gates on the detector qubit,

$$U_2^{(q_d)}(x, y) \equiv U_2(x + \theta_3, y + \theta_4), \tag{38}$$

$$U_1^{(q_d)}(y) \equiv U_1(y + \theta_5). \tag{39}$$

Considering these three types of gates, with the five additional parameters, to be the actual operations in Fig. 3, we can recalculate output expressions for all the circuits presented before. Namely, Eq. (7), as computed from Eqs. (10) and (11), and Eqs. (14), (20), (21), and (36). We obtain, respectively,

$$\mathcal{D} = \sqrt{\cos^2(\theta_3 + \theta_4) \sin^2\left(\frac{\alpha}{2} - \theta_5\right)}, \tag{40}$$

$$\langle X \rangle = \frac{1}{2}[\cos(\alpha + \phi + \theta_1 + \theta_2 - \theta_5) + \cos(\phi + \theta_1 + \theta_2 + \theta_5)], \tag{41}$$

$$\langle X_0 \rangle = \cos(\phi + \theta_1 + \theta_2 + \theta_5), \tag{42}$$

$$\langle X_1 \rangle = \cos(\alpha + \phi + \theta_1 + \theta_2 - \theta_5), \tag{43}$$

$$\mathcal{D}_m = 0. \tag{44}$$

These models provide a microscopic justification for the $\Theta_{1,2}$ parameters introduced with Eq. (24). As can be seen, θ_3 and θ_4 errors might reduce the maximum attainable $\mathcal{D}$, while θ_1, θ_2, and θ_5 can only shift the interference pattern for $\langle X \rangle$ without compromising $\mathcal{V}$. However, none of these SQGEs explains the slight dependence in α of the interference contrast for $\langle X_0 \rangle$, nor the varying $\mathcal{D}_m$ for different ϕ and α, both of which have been observed.

To understand what might be missing, let us note that state-of-the-art hardware achieves error rates reaching 10^{-4} on single-qubit gates [43]. The daunting challenge, though, is to apply the CNOT gate with the same level of accuracy. The current implementation of the CNOT gate in IBM Q devices uses the cross-resonance (CR) gate together with single-qubit rotations to reach errors between 10^{-2} and 10^{-3}. Significant work is being devoted to improve the fidelity of the CR gate with carefully designed pulse sequences. These would ideally generate only a ZX interaction term. An echoed CR sequence with target rotary pulses has recently shown improved performance in reducing the undesired error terms in the final pulse Hamiltonian [67]. The CNOT used in the experiments we report is based on this sequence.

To model the unitary errors in the CNOT implementation derived from the aforementioned CR sequence, we introduce

biased CNOT gate models written up in terms of different adimensional *bias ratios*, β_j, between the coupling strength of the extra error terms and the desired ZX interaction term. We take into account only the two-qubit subspace of the effective CR pulse Hamiltonian, and the error terms comprising entanglement with spectator qubits are not considered. The simplest

introduced gate model, BCNOT$_2$, contains only two bias ratios, β_1 and β_2, connected to IY and IZ errors. If classical crosstalk from the CR drive accidentally acting on the target qubit is included, error terms for IX, ZY, and ZZ are also considered, extending the previous model with β_3, β_4, and β_5 to obtain BCNOT$_5$:

$$
\begin{aligned}
\langle c't'| \text{ BCNOT}_5 |c\,t\rangle = {} & \frac{i(c'-1) - cc'(1+i) + ic}{\sqrt{2}}\Bigg[\cos\left(\gamma_c \pi/4\right)[t(1+i)(2t'-1) - (1+i)t' + i] \\
& + \frac{\sin\left(\gamma_c \pi/4\right)}{\gamma_c}(\beta_1 + \beta_2 + i\beta_3 + (-1)^c(\beta_4 + \beta_5 + i) + t(1+i)\{-\beta_1 \\
& + i[\beta_2 + \beta_3(2t'-1) + (-1)^c(i\beta_4 + \beta_5 + 2t'-1)]\} \\
& + t'(i-1)[\beta_1 + i\beta_2 - \beta_3 + (-1)^c(\beta_4 + i\beta_5 - 1)])\Bigg],
\end{aligned}
\tag{45}
$$

with a total of five bias parameters, where $c, t, c', t' \in \{0, 1\}$ and

$$
\gamma_c = \sqrt{[\beta_1 + (-1)^c\beta_4]^2 + [\beta_2 + (-1)^c\beta_5]^2 + [\beta_3 + (-1)^c]^2},
\tag{46}
$$

so that Eq. (45) encodes a $U(4)$ operator. In the no bias limit $\beta_{1,...,5} \to 0$, the BCNOT gates converge to the ideal CNOT, a perfect entangler [68]. When bias is present, the ability to maximally entangle some acted upon product state is severely hampered or entirely absent. The details of this derivation are provided in Appendix A.

Our goal now would be to obtain further generalizations of Eqs. (40)–(44) by replacing the ideal CNOT in the circuit of Fig. 3 first with the BCNOT$_2$ alternative and then with BCNOT$_5$, in addition to the single-qubit gates in Eqs. (37)–(39). We would then use those ten analytical models, supplemented by η and ε as in Eq. (25), to fit the experimental data, drawing conclusions about the actual physics taking place in the device.

However, such a program results in very long and cumbersome expressions which are often difficult to compute and simplify to an interpretable form. We could alleviate the problem by alternatively using a truncated series expansion of the BCNOT gates in their $\beta_{1,...,5} \approx 0$ parameters. Nevertheless, to make sure higher-order contributions are not missed when modeling the experimental results, we consider the exact BCNOT gates and fit the data to the readout probabilities obtained by numerically computing the circuits instead.

B. Model evaluation

Let us now fit the experimental results for $\mathcal{D}$, $\langle X \rangle$, $\langle X_0 \rangle$, $\langle X_1 \rangle$, and $\mathcal{D}_m$ to the theoretical predictions obtained by numerically calculating the output probabilities of the respective circuits with the different models for single- and two-qubit gates. The obtained $g(\phi, \alpha)$ models are again allowed to scale and shift with η and ϵ, following Eq. (25). We aim at identifying the best model to explain the data from each experiment, and thus understand the origin of the systematic deviations encountered in Sec. III.

For that, we evaluate the χ_ν^2 and RSE scores of each fit in order to compare them. However, we do not fit and test

a model on the same data, as this could erroneously signal better performance of more complex and parametrized models due to overfitting. We prevent it by evaluating performance on out-of-sample data, using part of the experimental results to fit the model and a different subset to test it. Specifically, we employ tenfold cross-validation with data folds selected uniformly at random in the (ϕ, α) space and score all models on equal splits. This method allows us to identify the models that maximize accuracy and minimize the structure of the residuals while preventing overfitting. We can confidently do so by comparing the assessed measures, reported in Table I, to rank model adequacy.

Looking at this table, we begin by noting the overall agreement between both scores: The models with the ten single- and two-qubit gate biases, in the last column, more accurately fit the experimental results. The magnitude of improvement is not the same for all of them, though. This is understood

TABLE I. Average performance measures for the fits obtained with different single- and two-qubit gate models, as evaluated with tenfold cross-validation. Boldface highlights the best-performing models. Table II presents the parameter values.

		No SQGEs	SQGEs		
		CNOT	CNOT	BCNOT$_2$	BCNOT$_5$
$\mathcal{D}$	χ_ν^2	1.34692	1.34766	1.34993	1.35376
	RSE	0.00950	0.00950	0.00950	0.00952
$\langle X \rangle$	χ_ν^2	5.78430	2.20479	**1.81458**	**1.81619**
	RSE	0.02062	0.01278	**0.01161**	**0.01161**
$\langle X_0 \rangle$	χ_ν^2	11.07644	8.61654	7.50310	**2.71485**
	RSE	0.03670	0.03094	0.02775	**0.02016**
$\langle X_1 \rangle$	χ_ν^2	2.83404	2.51185	2.48586	**2.42447**
	RSE	0.01997	0.01891	0.01873	**0.01865**
$\mathcal{D}_m$	χ_ν^2	2.05986	2.07046	**2.02330**	2.02358
	RSE	0.01582	0.01586	**0.01568**	**0.01568**

on the basis that the measured observables can be more or less resilient to the various coherent errors depending on the circuit, similar to what happens in Eqs. (40)–(44). At times, considering either BCNOT$_2$ or BCNOT$_5$ gates yields similar performance. However, as seen in the full table of parameters in Appendix B, using BCNOT$_5$ yields more stable and reasonable parameter values. Let us take a closer look at each row in Table I to discuss the details.

First, consider the scores for the $\langle X \rangle$ models. Table I clearly shows that taking into account SQGEs and both the BCNOT$_2$ and BCNOT$_5$ gates maximizes performance. We also find that the obvious pattern in the left panel of Fig. 6 mostly disappears to the human eye already using BCNOT$_2$ by plotting the residuals from any of the ten splits of the cross-validation procedure. However, simply on the basis of the goodness of fit measures, there is no clear choice between BCNOT$_2$ and BCNOT$_5$, indicating that there may be some overparametrization of the model obtained with the latter. The ambiguity is reduced by Table II, in which less problematic and more realistic fit parameters support the adoption of BCNOT$_5$ as the most reliable model for the origin of the data.

With respect to the $\langle X_0 \rangle$ and $\langle X_1 \rangle$ conditional interference patterns, Table I unambiguously validates SQGEs and the BCNOT$_5$ gate as the best performing choice. The goodness of fit improvement is especially significant for $\langle X_0 \rangle$. In fact, plotting again the region represented in the inset of the top-left panel of Fig. 7 with the fitted equations, we observe that the single ideal sinusoidal curve is replaced by multiple ones adjusting the points from each value of α. Even so, in the residuals of both $\langle X_0 \rangle$ and $\langle X_1 \rangle$ fits, patterns are still visually clear, albeit of diminished magnitude. In particular, these patterns show high frequency features with striking clarity, which would require further understanding. If we look closely at the bottom-left panel of Fig. 7 for instance, we can actually see these very small oscillations of the data around the sinusoidal curves of that fit.

This observation that some small additional effect is not being taken into account by the models leads us well into the discussion of $\mathcal{D}$ and $\mathcal{D}_m$. As can be seen, there is little difference in using any of the models to fit the experimental results of both these quantities, although considering SQGEs and both BCNOT$_2$ and BCNOT$_5$ slightly improves goodness of fit for the case of $\mathcal{D}_m$. It is important to note, however, that little further improvement is possible to the scores of the simplest $\mathcal{D}$ model, since these are already quite low to begin with. In fact, these scores are the best ones in the table. Nevertheless, we are also able to identify an odd feature in the plots of the experimental data for both $\mathcal{D}$ and $\mathcal{D}_m$: a smooth dependence on ϕ. This is unexpected because there is no ϕ dependence in the long and exact closed-form expressions of the models for these two quantities which integrate SQGEs and the BCNOT$_5$ gate. Once again, the data contain evidence of a small additional effect that is not included in our more complex models.

Therefore, while definitely improving goodness of fit as compared to the results in Sec. II, these new and more parametrized models still lack some of the physics that produces the results. This does not change if we consider a more complete model for SQGEs in which, on top of the biases introduced in Eqs. (37)–(39), fit parameters factored with ϕ and α are also added. This would be an attempt to capture an eventual linear dependence of the errors on those variables appearing in the nonconstant gates of Fig. 3. However, it provides no significant improvement to the performance of the models in the last three columns of Table I.

As such, we conjecture two possibilities for additional effects that could improve the leftover lack of fit and are not taken into account by the BCNOT gates based on the CR gate analysis in [67]. These are the potentially significant *spectator errors*, consisting of couplings to nearest-neighbor qubits, and the non-negligible errors due to the crosstalk coming back to the control transmon from the rotary tone of the CR gate.

Regardless of that, cross-validation shows that the more complete description of the physical operations introduced above greatly improves the modeling of the data, capturing therefore most of the physics in an interpretable manner. Moreover, the experimental observation of the inability to achieve the expected situations of full particlelike and full wavelike statistics can be explained as an important consequence of the biased CNOT operation, as shown promptly below.

C. Limitations on extreme-case statistics imposed by the biased CNOT

As we can see from Eqs. (40)–(44), SQGEs cannot prevent $\langle X \rangle$ from achieving $\mathcal{V} = 0$ and 1, since Eq. (15) would only suffer a phase shift. Similarly, SQGEs alone cannot compromise $\mathcal{V} = 1$ in $\langle X_0 \rangle$ and $\langle X_1 \rangle$ interference patterns, which would also only experience a phase shift, leaving contrast unchanged. Moreover, $\mathcal{D}_m$ is also unaffected by $\theta_{1,...,5}$ errors. The only exception is $\mathcal{D}$, which can actually fall short of achieving $\mathcal{D} = 1$ because of SQGEs. However, in this last case, we cannot tell from the experiment whether this is due to SQGEs or incoherent noise accounted for by η.

Here, we quickly show that the biased CNOT gate introduced in Eq. (45) can have this limiting effect, thus preventing the observation of full particlelike and wavelike statistics even with ideal single-qubit gate execution. To demonstrate that, we again recompute the output probabilities of all the circuits studied before, this time by taking the full expression for BCNOT$_5$ as replacement of the CNOT in Fig. 3 while preserving single-qubit gates as in Eq. (26). From those, all other quantities follow.

Before attempting an exact solution, it is instructive to examine the series expansion of the models truncated to first order in all $\beta_j \approx 0$, with $j = 1, \ldots, 5$. These five models derived from the four different circuits are

$$\mathcal{D} = \sin\left(\frac{\alpha}{2}\right) - (\beta_1 + \beta_2 + \beta_4 + \beta_5)\cos\left(\frac{\alpha}{2}\right), \tag{47}$$

$$\langle X \rangle = \frac{(1 + \beta_2 - \beta_4)\cos\phi + (1 - \beta_2 + \beta_4)\cos(\alpha + \phi)}{2} - \frac{\beta_1 + \beta_2 + \beta_4 + \beta_5}{2}[\sin\phi - \sin(\alpha + \phi)], \tag{48}$$

$$\langle X_0 \rangle = \cos(\phi) + \sin(\phi)\left[(\beta_1 - \beta_5)\sin\left(\frac{\alpha}{2}\right) - \beta_1 - \beta_2 - \beta_4 - \beta_5\right], \tag{49}$$

$$\langle X_1 \rangle = \cos(\alpha + \phi) + \sin(\alpha + \phi)\left[(\beta_1 - \beta_5)\sin\left(\frac{\alpha}{2}\right) + \beta_1 + \beta_2 + \beta_4 + \beta_5\right], \tag{50}$$

$$\mathcal{D}_m = -\beta_1 + \beta_5 + (\beta_2 - \beta_4)\cos\left(\frac{\alpha}{2}\right) - \frac{\pi}{2}\beta_3\sin\left(\frac{\alpha}{2}\right). \tag{51}$$

The first thing to note is that there is a lot of over-parametrization in these expressions. In fact, the first four models could be fitted by considering only nonzero β_1 and β_2 parameters. Moreover, in some of these equations, the limitations on the measurement statistics can already be confirmed. For instance, Eq. (48) cannot exhibit $\mathcal{V} = 0$. However, understanding the extreme values for $\mathcal{D}$ and $\mathcal{V}$ is mostly inconclusive using these approximations and requires an exact approach.

As previously stated, calculating complete closed-form models for these circuits results in exceedingly long and cumbersome expressions. Instead, we employ numerics to compute all circuits and derived expressions exactly. The limitations to full particlelike and wavelike statistics imposed by BCNOT₅ become apparent for all the circuits we tested. The severity of the effect varies with the magnitude of the different bias ratios, but it starts continuously as soon as there is some nonzero β_j. The only exception occurs for $\mathcal{V}_{\langle X \rangle}$, in which β_3 alone does not deform the function or prevent us from reaching the extremes. We illustrate this with an example in Fig. 8, which plots visibility and distinguishability obtained with the different circuits by specifying some arbitrarily chosen values for the biases.

On the top row, we put together the visibility obtained for $\langle X \rangle$ with the associated distinguishability, as obtained with the optimal measurement of the detector. As shown, even with only a little bias present, the interferometer qubit cannot display $\mathcal{V} = 0$ or 1, and the detector measurement cannot achieve $\mathcal{D} = 1$. This effect becomes more perceptible as the bias is in-creased. Note that $\mathcal{D}$ is no longer the actual distinguishability but what we would think the distinguishability is by using the $O_\alpha^\dagger$ basis measurement. It would be more accurate to name it $\mathcal{D}_m'$ but we keep $\mathcal{D}$ to avoid confusion with notation.

On the bottom row of Fig. 8, the visibility of the conditional interference patterns $\langle X_0 \rangle$ and $\langle X_1 \rangle$, which should ideally be 1, is plotted. In this case, $\mathcal{V} \approx 1$ for two α values within the 4π period, but all other interference patterns clearly have $\mathcal{V} < 1$. The decrease in visibility as α approaches π from zero is consistent with the experimental observations in Fig. 7. In addition, we also plot $\mathcal{D}_m$, which confirms via a different route that the readout in the Z basis cannot completely erase two-qubit correlations and produce full wavelike behavior. Note that the 4π period comes from the range of the input variables to the gates in Fig. 3, which is 2π.

These consequences to extreme-case statistics are thus in alignment with our experimental results. Recall that in the optimal which-path detection experiments we could not confirm $\mathcal{V} = 0$ for any value of α, corresponding to full particlelike dynamics, and in the quantum eraser experiments we could not obtain full wavelike statistics since we did not observe constant high $\mathcal{V}$ or $\mathcal{D}_m = 0$ for all α. Therefore, biased two-qubit operations may be a fundamental limitation in realizing Bohr's strong formulation of complementarity, and experimental tests should not overlook eventual small systematic deviations.

V. DISCUSSION AND CONCLUSIONS

In this paper, we set out to test interferometric complementarity using a superconducting transmon system. To do so, we proposed a general two-qubit quantum circuit to implement which-path detection with tunable sensitivity and adjustable readout basis. This circuit can be executed in any gate-based quantum computer with a universal gate set, regardless of the underlying hardware.

We considered four specific choices of the readout basis and thus introduced four different parametrized circuits, which were theoretically analyzed. First, we looked at two circuits in which the detector qubit is measured in the optimal guess basis. This allows capturing the maximum amount of the stored which-path information, optimizing the probability of guessing the correct path through the interferometer independently of its readout basis. We measured the interferometer qubit both in the X and the Z basis. By combining the results from the two circuits, the duality relation was tested for different visibility and distinguishability values, controlled by the detector sensitivity parameter α.

The other two circuits we studied measure the detector qubit in the computational basis directly, implementing the so-called quantum eraser setup. In this case, one expects

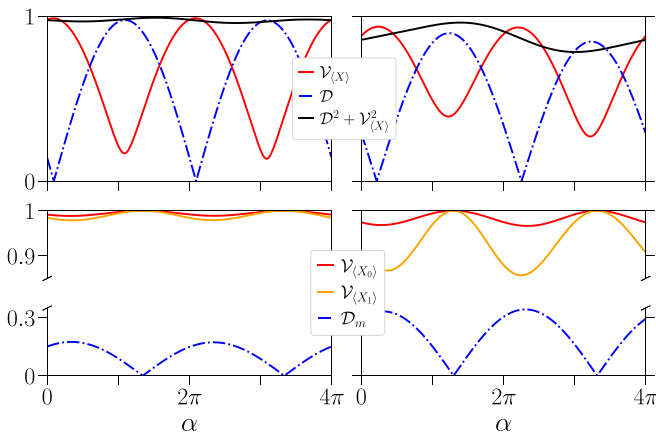
FIG. 8. Exact visibility and distinguishability plots for the different circuits encoded in Fig. 3, with the CNOT gate replaced by a BCNOT₅ with example bias values. In all of the panels, $\beta_1 = 0.06$, $\beta_3 = -0.05$, $\beta_4 = -0.07$, and $\beta_5 = 0.06$. While $\beta_2 = 0.09$ on the left-side panels, it is drastically increased to 0.3 on the right-side ones to emphasize the effect.

the which-path information to be destroyed at readout of the detector for any value of α, obtaining a constant and null measured distinguishability. We have shown that by using the guidance of the detector to sort the measurement results on the interferometer, two full-visibility interference patterns should be recovered, something that is not possible with the optimal readout.

These circuits were then implemented in an IBM Q quantum system controlled remotely. At a coarse level, we have found good agreement between the experiments and the theory, which is expected given the low depth of the quantum circuits. Interestingly, however, some small but persistent systematic deviations from the theoretical analysis have also been identified at a finer resolution.

Specifically, there was always some degree of waviness in the X basis measurements of the interferometer qubit for all α. This has prevented the observation of complete particlelike statistics, contrary to what was expected. Similarly, a complete wavelike behavior of the interferometer was not reached for all α with the quantum eraser circuits. We have found the visibility of the conditional interference patterns to depend on α, as well as a residual amount of measured distinguishability across the ϕ and α space, both characters of particlelike dynamics. This means that we could devise a way to win the which-path guessing game with a probability greater than 50% even using the quantum eraser setup, provided we can experiment with the circuit to tune our strategy before we bet.

We were able to relate these deviations between theory and experiments to coherent errors in single-qubit and CNOT gates. By properly modeling both types of operations in the physical circuits, the latter of which is based on the cross-resonance gate implementation, we have upgraded the theoretical models and showed these improve data explainability. In doing so, we have demonstrated that it is very difficult to actually generate full wavelike and particlelike statistics, since it requires bias-free two-qubit gates in our circuits.

This understanding raises the question of whether this limitation is exclusive to our implementation and how can nonbiased perfect two-qubit entanglers be operated to overcome it, a pertinent observation given how difficult it appears to be to precisely obtain maximal entanglement. Confidently realizing this prediction of quantum theory would allow us to address Bohr's strong principle of complementarity and not only the weaker duality relation in Eq. (2). Of course, in order to resolve the issue, one must be able to test the circuits with a level of precision in the measured observables and resolution in the variable space that can pierce through very small systematic deviations.

One option to correct the bias without attempting to improve gate calibration would be to customize circuits [69–71]. However, adding circuit depth makes the system more susceptible to incoherent noise, leading to results farther from the pure state limit. Moreover, our purposes were not to estimate observables by mixing measurements obtained with different circuits or qubit pairs [72], which also trades coherent for incoherent errors. Therefore, our findings also show that which-path experiments can be a way to benchmark quantum hardware.

The which-path experiments discussed here were initially proposed as *gedanken experiments* to highlight counter-

intuitive concepts in quantum theory. Their experimental verification required direct access to sophisticated equipment, available only to a few groups worldwide. With the advent of cloud-based quantum computers, our paper provides a way to remotely test this fundamental principle of quantum mechanics in the circuit model language of quantum computing. The simplicity of the procedure makes it even suitable for undergraduate quantum-mechanics courses.

APPENDIX A: THE BIASED-CNOT GATE MODEL

In the following, we derive the biased-CNOT gate used to model the physical CNOT in the device. Let us start by considering a ZX interaction Hamiltonian, with propagator $U_{ZX}(\theta) = \exp(-i\frac{\theta}{2}ZX)$. For $\theta = \pi/2$, it yields

$$U_{ZX}\left(\frac{\pi}{2}\right) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -i & 0 & 0 \\ -i & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & i & 1 \end{bmatrix} \equiv U_{ZX;\frac{\pi}{2}}, \quad \text{(A1)}$$

which is locally equivalent to a CNOT and capable of maximally entangling the product state $(|00\rangle + |10\rangle)/\sqrt{2}$. Even if we do not know the exact single-qubit rotations that together with $U_{ZX;\frac{\pi}{2}}$ give the CNOT, we can capture them by solving CNOT $= U_{ZX;\frac{\pi}{2}}M$ for M to obtain

$$M = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i & 0 & 0 \\ i & 1 & 0 & 0 \\ 0 & 0 & -i & 1 \\ 0 & 0 & 1 & -i \end{bmatrix}, \quad \text{(A2)}$$

obeying $[M, U_{ZX;\frac{\pi}{2}}] = 0$. One way of decomposing M would be with the $S^\dagger$ and $R_x(-\pi/2)$ single-qubit gates, applied on the control and target qubits, respectively:

$$M = S^\dagger \otimes R_x\left(-\frac{\pi}{2}\right). \quad \text{(A3)}$$

Ideally, the cross-resonance gate can be used to simulate an effective $U_{ZX}(\theta)$ coupling. It consists in a microwave drive of the control transmon at the frequency of the target, rotating the state of the latter depending on the state of the former. The CR tone has been studied and modeled by an effective Hamiltonian in the Pauli basis [73,74], and multipulse echo sequences and cancellation tones were designed to isolate as much as possible the ZX term. On IBM Q processors, the echoed CR with a target rotary pulse sequence achieves the current state-of-the-art performance in reducing the undesired error terms in the final interaction Hamiltonian [67], given by

$$H_{\text{eff}} = \tilde{\nu}_{ZX}\frac{ZX}{2} + H_{\text{err}_1} + H_{\text{err}_2}. \quad \text{(A4)}$$

Here,

$$H_{\text{err}_1} = \tilde{\nu}_{IY} \frac{IY}{2} + \tilde{\nu}_{IZ} \frac{IZ}{2} \qquad (A5)$$

contains the error terms produced by the standard CR two-pulse echo sequence in the two-qubit subspace, and

$$H_{\text{err}_2} = \tilde{\nu}_{IX} \frac{IX}{2} + \tilde{\nu}_{ZY} \frac{ZY}{2} + \tilde{\nu}_{ZZ} \frac{ZZ}{2} \qquad (A6)$$

contains those arising from classical crosstalk from the CR drive accidentally acting on the target qubit. The error terms comprising entanglement with spectator qubits are not being considered.

If H_{err_1} and H_{err_2} could be made vanishingly small, we would implement the perfect entangler in Eq. (A1) with $\tau = \frac{\pi}{2\tilde{\nu}_{ZX}}$ in $\exp(-i\tau H_{\text{eff}})$. However, when considering the full Hamiltonian, a systematic deviation arises. We can write the resulting block-diagonal operator, $U_{\text{eff};\frac{\pi}{2}}(\mathbf{B})$, in terms of five bias parameters $\beta_1 \equiv \frac{\tilde{\nu}_{IY}}{\tilde{\nu}_{ZX}}$, $\beta_2 \equiv \frac{\tilde{\nu}_{IZ}}{\tilde{\nu}_{ZX}}$, $\beta_3 \equiv \frac{\tilde{\nu}_{IX}}{\tilde{\nu}_{ZX}}$, $\beta_4 \equiv \frac{\tilde{\nu}_{ZY}}{\tilde{\nu}_{ZX}}$, and $\beta_5 \equiv \frac{\tilde{\nu}_{ZZ}}{\tilde{\nu}_{ZX}}$, taken together as $\mathbf{B} \equiv \{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5\}$. The entries of this 4×4 matrix can be compactly encoded by

$$\langle c't' | U_{\text{eff};\frac{\pi}{2}}(\mathbf{B}) | c\, t \rangle = (c + c' - 1)^2 \Bigg[\cos\left(\gamma_c \frac{\pi}{4}\right)(t + t' - 1)^2$$

$$+ \frac{\sin(\gamma_c \pi/4)}{\gamma_c}(i(t+t'-1)[\beta_2 + (-1)^c \beta_5]$$

$$- (t + t' - 2tt')\{i\beta_3 + i(-1)^c$$

$$+ (-1)^{t'} \beta_1 + (-1)^{t'+c} \beta_4\}) \Bigg], \qquad (A7)$$

where $c, t, c', t' \in \{0, 1\}$ and γ_c is written as in Eq. (46). One can check that $\lim_{\mathbf{B} \to \mathbf{0}} U_{\text{eff};\frac{\pi}{2}}(\mathbf{B}) = U_{ZX;\frac{\pi}{2}}$.

We now apply the same local operation, M, to this result to obtain the physical approximation to the CNOT. However, because $[M, U_{\text{eff};\frac{\pi}{2}}(\mathbf{B})] \neq 0$, an assumption is required on the ordering of the operators. If we define $\text{BCNOT}_5 \equiv U_{\text{eff};\frac{\pi}{2}}(\mathbf{B})M$, the result is in Eq. (45).

Now suppose we had chosen the definition $\text{BCNOT}'_5 \equiv M U_{\text{eff};\frac{\pi}{2}}(\mathbf{B})$ instead. We would have obtained the same operator we get by mapping $\{\beta_1 \to \beta_2, \beta_2 \to -\beta_1, \beta_3 \to \beta_3, \beta_4 \to \beta_5, \beta_5 \to -\beta_4\}$ in BCNOT_5. Therefore, while different, the parametrizations of BCNOT'_5 and BCNOT_5 are connected. This implies that if we use any of these definitions to obtain models to fit experimental data, both options will be able to adjust it and only the interpretation of each fit parameter changes. This suffices for our goal, which is not to fully characterize each bias individually but only to confirm the whole set can model the data.

The same applies to the only two other ways of obtaining the CNOT by combining the target entangler with phase and X rotation gates:

$$\text{BCNOT}''_5 \equiv (S^\dagger \otimes \mathbb{1}) \cdot U_{\text{eff};\frac{\pi}{2}}(\mathbf{B}) \cdot \left[\mathbb{1} \otimes R_x\left(-\frac{\pi}{2}\right) \right], \qquad (A8)$$

$$\text{BCNOT}'''_5 \equiv \left[\mathbb{1} \otimes R_x\left(-\frac{\pi}{2}\right) \right] \cdot U_{\text{eff};\frac{\pi}{2}}(\mathbf{B}) \cdot (S^\dagger \otimes \mathbb{1}). \qquad (A9)$$

For the first option, we have $\text{BCNOT}''_5 = \text{BCNOT}_5$, while for the second one we again obtain the same $\text{BCNOT}_5 \to \text{BCNOT}'''_5$ mapping we got for BCNOT'_5. Therefore, it is also important to note that these maps from BCNOT_5 to BCNOT'_5 and BCNOT'''_5 do not mix β_j parameters from the different error Hamiltonians in Eqs. (A5) and (A6). This is helpful if we want to ignore one of these in the analysis, as it guarantees these four sequences to arrive at a biased CNOT gate still connected among each other. Our BCNOT_2 model takes $H_{\text{err}_2} = 0$.

APPENDIX B: FURTHER DETAILS ON MODEL EVALUATION

As described in Sec. IV, we take different combinations of ideal and faulty models for the physical gates in Fig. 3 to compute four models of the output probabilities of each circuit. We then fit the relevant quantities by allowing parameter optimization within the fixed bounds $|\beta_j| \leqslant 0.5$, $|\theta_k| \leqslant \pi/10$, $\eta \in [0.7, 1]$, $|\varepsilon| \leqslant 0.5$. We employ tenfold cross-validation to assess goodness of fit as well as to find the optimal parameters by averaging them over the different splits used by the algorithm. Precision errors are given by the standard deviation of these ten values.

Consider the optimal parameters in the first column of Table II concerning the models using the ideal CNOT and no SQGEs. The precision margins in this case are among the best in the table. This is simply because, being less parametrized, these models quickly lock to the data and do not allow much freedom in optimizing the function. However, as seen in Table I, they yield the worst goodness of fit measures.

In the next three columns of Table II, with the more complex models, the optimization returned some final parameter values right at the predefined bounds. Consider the second column. As explained, these fits are performed using numerical calculations as reference, but in this particular case the closed-form expressions are simple enough, being given by Eqs. (40)–(44). Looking at the results in the table, we see that the parameters with unstable values (underlined) are actually not present in those expressions. Thus, it is unsurprising the fit could not optimize them, and these can be ignored.

Now for the third and fourth columns, there are two overlapping reasons contributing to some unstable parameter values. The first is the same as before: Some of these error parameters, though possibly existing in the gates in which they were introduced, might not reflect in the observables being measured by different circuits. However, by calculating the circuit numerically we cannot tell which ones, since all parameters are included with each gate to obtain the most general possible output; we do not add the additional layer of complexity to try to numerically understand which gate parameters could be dropped for each particular circuit.

The second reason is the fact that, as described in Sec. IV, the data carry evidence that coherent behavior persists unaccounted for by the more complex models we considered. Since these models have more degrees of freedom to optimize the fit and the data are not entirely compatible with the model, the parameters adapt in ways that might not reflect the physics from which they were derived. Nevertheless, it is clear that the best overall results in terms of goodness of fit as well as less problematic values for the optimal parameters are obtained with the models considering SQGEs and the BCNOT_5 gate.

Note that, in general, β_j and θ_k parameters shown in different rows of Table II need not match because these correspond

TABLE II. Optimal model parameters obtained with different single- and two-qubit gate models, evaluated by resorting to the ten different fits used in tenfold cross-validation. Underlined values are unstable cases in which some of the fits returned the value at one of the imposed bounds, occurring at only one of them in all ten different splits when precision errors indicate zero.

		No SQGEs		SQGEs					
		CNOT		CNOT		BCNOT$_2$		BCNOT$_5$	
$\mathcal{D}$	η	0.90206	±0.00011	0.91506	±0.00453	0.92386	±0.00352	0.92707	±0.00238
	ε	0.00261	±0.00009	0.00261	±0.00011	0.00240	±0.00013	0.00006	±0.00147
	θ_1			$\underline{0.01839}\,\pi$	$\pm0.06088\,\pi$	$0.00471\,\pi$	$\pm0.07067\,\pi$	$-0.01176\,\pi$	$\pm0.05021\,\pi$
	θ_2			$\underline{-0.06065}\,\pi$	$\pm0.07914\,\pi$	$\underline{-0.04070}\,\pi$	$\pm0.09111\,\pi$	$\underline{-0.06065}\,\pi$	$\pm0.07914\,\pi$
	θ_3			$0.02689\,\pi$	$\pm0.00462\,\pi$	$0.02131\,\pi$	$\pm0.00151\,\pi$	$0.01876\,\pi$	$\pm0.00183\,\pi$
	θ_4			$0.02604\,\pi$	$\pm0.00430\,\pi$	$0.02161\,\pi$	$\pm0.00174\,\pi$	$0.01876\,\pi$	$\pm0.00183\,\pi$
	θ_5			$0.00037\,\pi$	$\pm0.00002\,\pi$	$-0.02377\,\pi$	$\pm0.00405\,\pi$	$-0.02842\,\pi$	$\pm0.00646\,\pi$
	β_1					0.03751	±0.00629	0.01340	±0.00759
	β_2					0.03800	±0.00628	0.02078	±0.00539
	β_3							0.06064	±0.00880
	β_4							0.01920	±0.00587
	β_5							0.03955	±0.01229
$\langle X \rangle$	η	0.92975	±0.00020	0.93039	±0.00009	0.94163	±0.00077	0.94054	±0.00170
	ε	0.03707	±0.00013	0.03700	±0.00005	0.03706	±0.00005	0.03706	±0.00005
	θ_1			$-0.00525\,\pi$	$\pm0.00330\,\pi$	$-0.01759\,\pi$	$\pm0.00056\,\pi$	$-0.00769\,\pi$	$\pm0.00142\,\pi$
	θ_2			$-0.00573\,\pi$	$\pm0.00330\,\pi$	$-0.01707\,\pi$	$\pm0.00099\,\pi$	$-0.00767\,\pi$	$\pm0.00142\,\pi$
	θ_3			$\underline{0.05396}\,\pi$	$\pm0.07693\,\pi$	$\underline{0.10000}\,\pi$	$\pm0.00000\,\pi$	$0.00961\,\pi$	$\pm0.01214\,\pi$
	θ_4			$\underline{0.02049}\,\pi$	$\pm0.06694\,\pi$	$\underline{-0.01281}\,\pi$	$\pm0.07203\,\pi$	$\underline{-0.00520}\,\pi$	$\pm0.08215\,\pi$
	θ_5			$-0.00082\,\pi$	$\pm0.00004\,\pi$	$-0.05063\,\pi$	$\pm0.00196\,\pi$	$-0.00676\,\pi$	$\pm0.00396\,\pi$
	β_1					0.17064	±0.00658	0.10534	±0.02135
	β_2					-0.01194	±0.00029	-0.03753	±0.01307
	β_3							0.07033	±0.07038
	β_4							-0.01431	±0.01384
	β_5							-0.03618	±0.01304
$\langle X_0 \rangle$	η	0.91745	±0.00015	0.91740	±0.00018	0.91900	±0.00011	0.94262	±0.00014
	ε	0.03394	±0.00008	0.03360	±0.00007	0.03356	±0.00007	0.03237	±0.00005
	θ_1			$-0.00331\,\pi$	$\pm0.00002\,\pi$	$0.05805\,\pi$	$\pm0.00009\,\pi$	$0.00326\,\pi$	$\pm0.00171\,\pi$
	θ_2			$-0.00331\,\pi$	$\pm0.00002\,\pi$	$0.05805\,\pi$	$\pm0.00010\,\pi$	$0.00327\,\pi$	$\pm0.00172\,\pi$
	θ_3			$\underline{0.00645}\,\pi$	$\pm0.09506\,\pi$	$\underline{0.10000}\,\pi$	$\pm0.00000\,\pi$	$-0.01545\,\pi$	$\pm0.02125\,\pi$
	θ_4			$\underline{0.01769}\,\pi$	$\pm0.07071\,\pi$	$\underline{-0.01414}\,\pi$	$\pm0.08456\,\pi$	$\underline{0.00289}\,\pi$	$\pm0.09722\,\pi$
	θ_5			$-0.00331\,\pi$	$\pm0.00002\,\pi$	$\underline{-0.10000}\,\pi$	$\pm0.00000\,\pi$	$0.00512\,\pi$	$\pm0.00443\,\pi$
	β_1					-0.02477	±0.00063	-0.00394	±0.02631
	β_2					-0.05963	±0.00029	-0.02698	±0.02227
	β_3							0.23815	±0.00155
	β_4							-0.06506	±0.03454
	β_5							0.02333	±0.03759
$\langle X_1 \rangle$	η	0.93767	±0.00006	0.93765	±0.00005	0.93826	±0.00009	0.94136	±0.00007
	ε	0.03391	±0.00004	0.03379	±0.00006	0.03382	±0.00005	0.03379	±0.00006
	θ_1			$-0.00137\,\pi$	$\pm0.00047\,\pi$	$0.05483\,\pi$	$\pm0.00050\,\pi$	$0.00477\,\pi$	$\pm0.00164\,\pi$
	θ_2			$-0.00116\,\pi$	$\pm0.00030\,\pi$	$0.05485\,\pi$	$\pm0.00047\,\pi$	$0.00477\,\pi$	$\pm0.00164\,\pi$
	θ_3			$-0.02542\,\pi$	$\pm0.06461\,\pi$	$-0.03627\,\pi$	$\pm0.00287\,\pi$	$0.01919\,\pi$	$\pm0.00582\,\pi$
	θ_4			$\underline{0.00000}\,\pi$	$\pm0.09982\,\pi$	$\underline{-0.03084}\,\pi$	$\pm0.07377\,\pi$	$0.04070\,\pi$	$\pm0.09111\,\pi$
	θ_5			$0.00106\,\pi$	$\pm0.00030\,\pi$	$\underline{0.10000}\,\pi$	$\pm0.00000\,\pi$	$\underline{-0.00225}\,\pi$	$\pm0.00134\,\pi$
	β_1					-0.00272	±0.00027	-0.01068	±0.00611
	β_2					0.04713	±0.00304	0.00853	±0.00825
	β_3							0.06933	±0.00161
	β_4							0.07775	±0.01049
	β_5							-0.02379	±0.00541

TABLE II. *(Countinued.)*

		No SQGEs		SQGEs					
		CNOT		CNOT		BCNOT$_2$		BCNOT$_5$	
$\mathcal{D}_m$	η			1.00000	± 0.00000	0.99999	± 0.00002	0.98819	± 0.03324
	ε	-0.00370	± 0.00004	-0.00370	± 0.00007	0.01448	± 0.00224	-0.32232	± 0.01093
	θ_1			$0.00425\,\pi$	$\pm 0.07391\,\pi$	$-0.01039\,\pi$	$\pm 0.07781\,\pi$	$-0.00380\,\pi$	$\pm 0.06965\,\pi$
	θ_2			$-0.02309\,\pi$	$\pm 0.06652\,\pi$	$-0.00139\,\pi$	$\pm 0.08569\,\pi$	$0.00209\,\pi$	$\pm 0.07158\,\pi$
	θ_3			$0.03880\,\pi$	$\pm 0.07545\,\pi$	$0.09986\,\pi$	$\pm 0.00032\,\pi$	$0.06215\,\pi$	$\pm 0.04653\,\pi$
	θ_4			$-0.00000\,\pi$	$\pm 0.00000\,\pi$	$-0.00591\,\pi$	$\pm 0.06771\,\pi$	$-0.01688\,\pi$	$\pm 0.06911\,\pi$
	θ_5			$-0.03308\,\pi$	$\pm 0.07253\,\pi$	$-0.03750\,\pi$	$\pm 0.08945\,\pi$	$-0.09324\,\pi$	$\pm 0.01471\,\pi$
	β_1					-0.01573	± 0.00234	0.12234	± 0.04425
	β_2					-0.00364	± 0.00140	-0.37500	± 0.04456
	β_3							0.05512	± 0.01691
	β_4							-0.10399	± 0.04620
	β_5							-0.39596	± 0.05204

to different experiments, performed with different calibrations of the device as well as different qubits. However, for the models of $\langle X \rangle$ and $\mathcal{D}$, as well as $\langle X_0 \rangle$ and $\langle X_1 \rangle$, for which data were obtained at the same execution time in the device, it seems some of the β_j and θ_k values in the last column of Table II are comparable while others are not. This inconsistency is in agreement with what we argued before: there is still some physical effect unaccounted for, which gives some freedom to the different models to find different optimal values for the parameters included in the model. Furthermore,

if overparametrization is present, the final models are also optimizable with different values of the parameters, which could also explain some differences.

To summarize, we have shown that our models capture most of the physics of the operations, doing a very good job at modeling the data in an interpretable manner. However, we are also sure there is still some small degree of incompleteness in them, not only because of the remnant lack of fit and nonuniform distribution of residuals, but also due to some unstable fit parameters.

[1] N. Bohr, The quantum postulate and the recent development of atomic theory, Nature (London) **121**, 580 (1928).

[2] L. De Broglie, Recherches sur la théorie des quanta, Ph.D. thesis, Migration-Université en Cours d'Affectation, 1924.

[3] W. Heisenberg, Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik, Z. Phys. **43**, 172 (1927).

[4] N. Bohr, *Discussion with Einstein on Epistemological Problems in Atomic Physics, in Albert Einstein: Philosopher-Scientist* (Library of Living Philosophers, Evanston, IL, 1949), Chap. 7, pp. 201–241.

[5] W. Heisenberg, *The Physical Principles of the Quantum Theory* (University of Chicago, Chicago, 1930).

[6] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1965), Vol. 3.

[7] J. A. Wheeler, The "past" and the "delayed-choice" double-slit experiment, in *Mathematical Foundations of Quantum Theory*, edited by A. Marlow (Academic, New York, 1978), pp. 9–48.

[8] W. K. Wootters and W. H. Zurek, Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle, Phys. Rev. D **19**, 473 (1979).

[9] M. O. Scully, B.-G. Englert, and J. Schwinger, Spin coherence and Humpty-Dumpty. III. The effects of observation, Phys. Rev. A **40**, 1775 (1989).

[10] M. O. Scully and H. Walther, Quantum optical test of observation and complementarity in quantum mechanics, Phys. Rev. A **39**, 5229 (1989).

[11] D. Deutsch, Uncertainty in Quantum Measurements, Phys. Rev. Lett. **50**, 631 (1983).

[12] P. Mittelstaedt, A. Prieur, and R. Schieder, Unsharp particle-wave duality in a photon split-beam experiment, Found. Phys. **17**, 891 (1987).

[13] D. M. Greenberger and A. Yasin, Simultaneous wave and particle knowledge in a neutron interferometer, Phys. Lett. A **128**, 391 (1988).

[14] G. Jaeger, A. Shimony, and L. Vaidman, Two interferometric complementarities, Phys. Rev. A **51**, 54 (1995).

[15] B.-G. Englert, Fringe Visibility and Which-Way Information: An Inequality, Phys. Rev. Lett. **77**, 2154 (1996).

[16] G. Björk and A. Karlsson, Complementarity and quantum erasure in Welcher WEG experiments, Phys. Rev. A **58**, 3477 (1998).

[17] S. Dürr, Quantitative wave-particle duality in multibeam interferometers, Phys. Rev. A **64**, 042113 (2001).

[18] M. Jakob and J. A. Bergou, Complementarity and entanglement in bipartite qudit systems, Phys. Rev. A **76**, 052107 (2007).

[19] M. Jakob and J. A. Bergou, Quantitative complementarity relations in bipartite systems: Entanglement as a physical reality, Opt. Commun. **283**, 827 (2010).

[20] P. J. Coles, J. Kaniewski, and S. Wehner, Equivalence of wave-particle duality to entropic uncertainty, Nat. Commun. **5**, 5814 (2014).

[21] P. J. Coles, Entropic framework for wave-particle duality in multipath interferometers, Phys. Rev. A **93**, 062111 (2016).

[22] R. M. Angelo and A. D. Ribeiro, Wave-particle duality: An information-based approach, Found. Phys. **45**, 1407 (2015).

[23] E. Bagan, J. A. Bergou, S. S. Cottrell, and M. Hillery, Relations Between Coherence and Path Information, Phys. Rev. Lett. **116**, 160406 (2016).

[24] S. Dürr, T. Nonn, and G. Rempe, Origin of quantum-mechanical complementarity probed by a "which-way" experiment in an atom interferometer, Nature (London) **395**, 33 (1998).

[25] P. Bertet, S. Osnaghi, A. Rauschenbeutel, G. Nogues, A. Auffeves, M. Brune, J. Raimond, and S. Haroche, A complementarity experiment with an interferometer at the quantum-classical boundary, Nature (London) **411**, 166 (2001).

[26] T. J. Herzog, P. G. Kwiat, H. Weinfurter, and A. Zeilinger, Complementarity and the Quantum Eraser, Phys. Rev. Lett. **75**, 3034 (1995).

[27] Y.-H. Kim, R. Yu, S. P. Kulik, Y. Shih, and M. O. Scully, Delayed "Choice" Quantum Eraser, Phys. Rev. Lett. **84**, 1 (2000).

[28] S. P. Walborn, M. O. Terra Cunha, S. Pádua, and C. H. Monken, Double-slit quantum eraser, Phys. Rev. A **65**, 033818 (2002).

[29] L. Neves, G. Lima, J. Aguirre, F. A. Torres-Ruiz, C. Saavedra, and A. Delgado, Control of quantum interference in the quantum eraser, New J. Phys. **11**, 073035 (2009).

[30] A. Peruzzo, P. Shadbolt, N. Brunner, S. Popescu, and J. L. O'Brien, A quantum delayed-choice experiment, Science **338**, 634 (2012).

[31] F. Kaiser, T. Coudreau, P. Milman, D. B. Ostrowsky, and S. Tanzilli, Entanglement-enabled delayed-choice experiment, Science **338**, 637 (2012).

[32] E. Buks, R. Schuster, M. Heiblum, D. Mahalu, and V. Umansky, Dephasing in electron interference by a "which-path" detector, Nature (London) **391**, 871 (1998).

[33] D. Sprinzak, E. Buks, M. Heiblum, and H. Shtrikman, Controlled Dephasing of Electrons via a Phase Sensitive Detector, Phys. Rev. Lett. **84**, 5820 (2000).

[34] I. Neder, M. Heiblum, D. Mahalu, and V. Umansky, Entanglement, Dephasing, and Phase Recovery via Cross-Correlation Measurements of Electrons, Phys. Rev. Lett. **98**, 036803 (2007).

[35] G. Teklemariam, E. M. Fortunato, M. A. Pravia, T. F. Havel, and D. G. Cory, NMR Analog of the Quantum Disentanglement Eraser, Phys. Rev. Lett. **86**, 5845 (2001).

[36] G. Teklemariam, E. M. Fortunato, M. A. Pravia, Y. Sharf, T. F. Havel, D. G. Cory, A. Bhattaharyya, and J. Hou, Quantum erasers and probing classifications of entanglement via nuclear magnetic resonance, Phys. Rev. A **66**, 012309 (2002).

[37] X. Peng, X. Zhu, X. Fang, M. Feng, M. Liu, and K. Gao, An interferometric complementarity experiment in a bulk nuclear magnetic resonance ensemble, J. Phys. A: Math. Gen. **36**, 2555 (2003).

[38] S.-B. Zheng, Y.-P. Zhong, K. Xu, Q.-J. Wang, H. Wang, L.-T. Shen, C.-P. Yang, J. M. Martinis, A. N. Cleland, and S.-Y. Han, Quantum Delayed-Choice Experiment with a Beam Splitter in a Quantum Superposition, Phys. Rev. Lett. **115**, 260403 (2015).

[39] K. Liu, Y. Xu, W. Wang, S.-B. Zheng, T. Roy, S. Kundu, M. Chand, A. Ranadive, R. Vijay, Y. Song, L. Duan, and L. Sun, A twofold quantum delayed-choice experiment in a superconducting circuit, Sci. Adv. **3**, e1603159 (2017).

[40] A. Bienfait, Y. P. Zhong, H.-S. Chang, M.-H. Chou, C. R. Conner, E. Dumur, J. Grebel, G. A. Peairs, R. G. Povey, K. J. Satzinger, and A. N. Cleland, Quantum Erasure Using En-

tangled Surface Acoustic Phonons, Phys. Rev. X **10**, 021055 (2020).

[41] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Observation of a "quantum eraser": A revival of coherence in a two-photon interference experiment, Phys. Rev. A **45**, 7729 (1992).

[42] P. D. D. Schwindt, P. G. Kwiat, and B.-G. Englert, Quantitative wave-particle duality and nonerasing quantum erasure, Phys. Rev. A **60**, 4285 (1999).

[43] IBM Quantum, https://quantum-computing.ibm.com (accessed 2021-04-30).

[44] M. Amico and C. Dittel, Simulation of wave-particle duality in multipath interferometers on a quantum computer, Phys. Rev. A **102**, 032605 (2020).

[45] N. Schwaller, M.-A. Dupertuis, and C. Javerzac-Galy, Evidence of the entanglement constraint on wave-particle duality using the IBM Q quantum computer, Phys. Rev. A **103**, 022409 (2021).

[46] M. B. Pozzobom, M. L. W. Basso, and J. Maziero, Experimental tests of the density matrix's property-based complementarity relations, Phys. Rev. A **103**, 022212 (2021).

[47] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, Proc. R. Soc. A **454**, 339 (1998).

[48] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[49] M. O. Scully, B.-G. Englert, and H. Walther, Quantum optical tests of complementarity, Nature (London) **351**, 111 (1991).

[50] R. Ionicioiu and D. R. Terno, Proposal for a Quantum Delayed-Choice Experiment, Phys. Rev. Lett. **107**, 230406 (2011).

[51] Qiskit: An open-source SDK for working with quantum computers at the level of pulses, circuits, and algorithms, https://qiskit.org (accessed 2021-04-30).

[52] C. Chamberland, G. Zhu, T. J. Yoder, J. B. Hertzberg, and A. W. Cross, Topological and Subsystem Codes on Low-Degree Graphs with Flag Qubits, Phys. Rev. X **10**, 011022 (2020).

[53] A. Blais, J. Gambetta, A. Wallraff, D. I. Schuster, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Quantum-information processing with circuit quantum electrodynamics, Phys. Rev. A **75**, 032329 (2007).

[54] 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).

[55] J. Majer, J. Chow, J. Gambetta, J. Koch, B. Johnson, J. Schreier, L. Frunzio, D. Schuster, A. A. Houck, A. Wallraff *et al.*, Coupling superconducting qubits via a cavity bus, Nature (London) **449**, 443 (2007).

[56] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer's guide to superconducting qubits, Appl. Phys. Rev. **6**, 021318 (2019).

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

[58] D. Alsina and J. I. Latorre, Experimental test of Mermin inequalities on a five-qubit quantum computer, Phys. Rev. A **94**, 012314 (2016).

[59] D. García-Martín and G. Sierra, Five experimental tests on the 5-qubit IBM quantum computer, J. Appl. Math. Phys. **6**, 1460 (2018).

[60] https://www.rigetti.com (accessed 2021-04-30).

[61] https://quantumai.google (accessed 2021-04-30).

[62] https://www.quantum-inspire.com (accessed 2021-04-30).

[63] Y. Chen, M. Farahzad, S. Yoo, and T.-C. Wei, Detector tomography on IBM quantum computers and mitigation of an imperfect measurement, Phys. Rev. A **100**, 052315 (2019).

[64] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, Quantum **4**, 257 (2020).

[65] A. Asfaw, L. Bello, Y. Ben-Haim, S. Bravyi, N. Bronn, L. Capelluto, A. C. Vazquez, J. Ceroni, R. Chen, A. Frisch, J. Gambetta, S. Garion, L. Gil, S. D. L. P. Gonzalez, F. Harkins, T. Imamichi, D. McKay, A. Mezzacapo, Z. Minev, R. Movassagh *et al.*, Learn quantum computation using qiskit, http://community.qiskit.org/textbook (2020).

[66] A. Wald and J. Wolfowitz, On a test whether two samples are from the same population, Ann. Math. Stat. **11**, 147 (1940).

[67] N. Sundaresan, I. Lauer, E. Pritchett, E. Magesan, P. Jurcevic, and J. M. Gambetta, Reducing unitary and spectator errors in cross resonance with optimized rotary echoes, PRX Quantum **1**, 020318 (2020).

[68] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Geometric theory of nonlocal two-qubit operations, Phys. Rev. A **67**, 042313 (2003).

[69] M. J. Bremner, C. M. Dawson, J. L. Dodd, A. Gilchrist, A. W. Harrow, D. Mortimer, M. A. Nielsen, and T. J. Osborne, Practical Scheme for Quantum Computation with any Two-Qubit Entangling Gate, Phys. Rev. Lett. **89**, 247902 (2002).

[70] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Exact Two-Qubit Universal Quantum Circuit, Phys. Rev. Lett. **91**, 027903 (2003).

[71] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algorithm, Quantum Info. Comput. **6**, 81 (2006).

[72] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, Phys. Rev. A **94**, 052325 (2016).

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

[74] M. Malekakhlagh, E. Magesan, and D. C. McKay, First-principles analysis of cross-resonance gate operation, Phys. Rev. A **102**, 042605 (2020).