Sampling and Scrambling on a Chain of Superconducting Qubits

Michael R. Geller*
Center for Simulational Physics, University of Georgia, Athens, Georgia 30602, USA

(Received 9 December 2017; revised manuscript received 16 March 2018; published 31 August 2018)

We study a circuit, the Josephson sampler, that embeds a real vector into an entangled state of n qubits, and optionally samples from it. We estimate two fidelity measures and a variety of entropies and entanglement measures for this circuit, versus depth, on the 16-qubit ibmqx5 chip. This provides a quantitative assessment of the actual performance of the circuit on a first-generation quantum computer. To assess its expressiveness, we also measure its ability to generate quantum chaos as measured by out-of-time-order correlators (OTOCs). We conjecture that OTOCs can be used to validate a source of pseudorandom unitaries with only a single instance, which has application beyond the Josephson sampler. The circuit requires nearest-neighbor controlled-Z gates on a chain and is especially well suited for first-generation superconducting architectures.

DOI: 10.1103/PhysRevApplied.10.024052

I. INTRODUCTION

A common task in quantum computing is to embed large amounts of classical data into a quantum state of n qubits. Large means more than n bits, so the data must be input through a circuit with adjustable parameters. This has application to quantum machine learning [1], where the data of interest (e.g., images) are kilobytes in size or larger, and also to quantum simulation, where the data can be used as variational parameters [2–4]. Also, inputting pseudorandom classical data into such a circuit can be used to approximate Haar random unitaries [5–7], which have wide application in quantum computing and quantum information [5,8–11].

In this work, we will study the performance of a practical embedding circuit—*the Josephson sampler*—on the IBM Quantum Experience ibmqx5 device, which has 16 transmon qubits. We study samplers up to size $n=6$. The circuit acts on a 1D chain of qubits with nearest-neighbor controlled-NOT (CNOT) or controlled-Z (CZ) gates, and has a layered construction

$$U = U_L \cdots U_2 U_1, \tag{1}$$

with L layers, as shown in Fig. 1. Each gate u is a single-qubit rotation gate

$$u(\theta,\phi) \equiv e^{-i\phi Z/2} e^{-i\theta Y/2}, \tag{2}$$

where the y rotation is applied first and the z rotation is done in software and carries no depth. The vertical two-qubit gates are CZ gates. Important features of the design are that the gateset is universal and the columns of CZ gates rapidly generate entanglement. A sampler circuit with L layers maps a real vector $x \in \mathbb{R}^m$ of dimension

$$m = 2(2n-2)L, \tag{3}$$

to (ideally) a unitary $U \in U(2^n)$. Additional details about the circuit are provided in the Appendix.

Our work is inspired by Aaronson and Arkhipov [12], Boixo *et al.* [13], Kandala *et al.* [4], and Neill *et al.* [14]. Although our objective is not quantum supremacy, which is unlikely on a chain, we will apply many of the techniques introduced in Refs. [13] and [14]. A different but related problem of boson sampling with superconducting resonators has been discussed in Refs. [15] and [16]. The Josephson sampler is an alternative to the hardware-efficient circuit introduced by Kandala *et al.* [4], trading some performance for simpler portability. We will also extend previous work by measuring four-point out-of-time-order correlation functions [17–28]. These quantities probe the butterfly effect, a dynamical signature of quantum chaos, and information scrambling.

II. FIDELITY

The Josephson sampler is designed for use on first-generation gate-based quantum computers, which are not error corrected. It is therefore important to assess its performance on real devices. One way to measure the quality of a circuit implementation is to estimate the fidelity,

$$F \equiv \left(\mathrm{Tr}\sqrt{\sqrt{\rho_t}\rho\sqrt{\rho_t}}\right)^2 = \mathrm{Tr}\,\rho\rho_t, \tag{4}$$

*mgeller@uga.edu

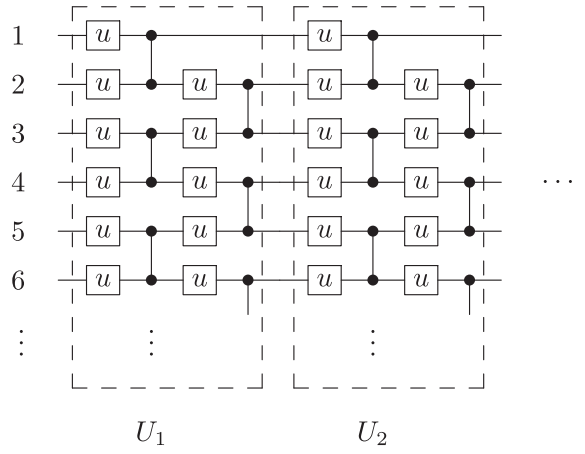

FIG. 1. Josephson sampler construction with L layers.

of the resulting final state ρ with its ideal pure target ρ_t. Given a classically precomputed ρ_t, there is a randomized protocol to estimate F, by Flammia and Liu [29], which we find to converge very quickly. In Fig. 2, we plot the state fidelity versus the number of sampler layers L. For each sampler size (n, L), the average state fidelity over s pseudorandom circuits is measured, with $s = 4$. The Flammia-Liu protocol [29] is based on an expansion of the density matrix in the Pauli basis, but requires expectation-value measurements for only a small number r of the more probable ones. In Fig. 2, we use $r = 8$ Pauli operator samples, which is sufficient to keep the sampling error to a few percent.

Boixo *et al.* [13] introduced alternative information-theoretic tools to quantify circuit fidelity. Their approach

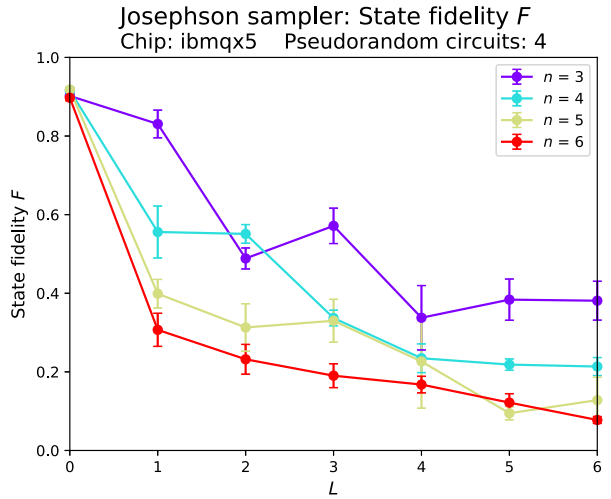

FIG. 2. The measured state fidelity given in Eq. (4) averaged over s pseudorandom circuits of size (n, L), after application to initial state $|0\rangle^{\otimes n}$. Each state-fidelity measurement consists of r Pauli expectation-value measurements. Here, $s = 4$ and $r = 8$. The error bars denote one standard error.

starts with a classically precomputed ideal probability distribution $p_{\text{ideal}}(x)$, with $x \in \{0, 1\}^n$ a classical state, or an algorithm for computing it. The goal is to experimentally distinguish between the actual distribution $p_{\text{meas}}(x)$ and the ideal one $p_{\text{ideal}}(x)$. However, we want to do this *without* an accurate estimate of $p_{\text{meas}}(x)$, which would require $O(2^n)$ measurement samples.

The method of Boixo *et al.* [13] focuses on the rare, high-information events and their statistics. Given $p_{\text{ideal}}(x)$ define the *information* $s(x) \equiv \log_2[1/p_{\text{ideal}}(x)]$ contained in each classical state x, measured in bits. First, we ask some purely theoretical questions about the ideal distribution $p_{\text{ideal}}(x)$. For example, we can calculate the average information

$$S_{\text{ideal}} \equiv \langle s(x)\rangle_{\text{ideal}} = \sum_x p_{\text{ideal}}(x)s(x), \tag{5}$$

when sampling from $p_{\text{ideal}}(x)$, which is the classical Shannon entropy. We can also calculate the uniform average

$$S_{\text{unif}} \equiv \langle s(x)\rangle_{\text{unif}} = \frac{1}{N}\sum_x s(x), \quad N = 2^n, \tag{6}$$

which is larger than the entropy because it overrepresents the rare prob $< 1/N$ high-information events. Therefore, the difference

$$S_{\text{unif}} - S_{\text{ideal}}, \tag{7}$$

is sensitive to the statistics of rare events.

In the information-theoretic approach, the quantity directly estimated is the cross entropy

$$H_c \equiv \langle s(x)\rangle_{\text{meas}} = \sum_x p_{\text{meas}}(x)s(x). \tag{8}$$

Note that it is not necessary to explicitly reconstruct $p_{\text{meas}}(x)$, only to sample from it. So a possible fidelity measure is

$$F_{\text{in}} \equiv \frac{S_{\text{unif}} - H_c}{S_{\text{unif}} - S_{\text{ideal}}}, \tag{9}$$

which was recently used by Neill *et al.* [14]. Our measured values of F_{in} are shown in Fig. 3. An advantage of information-fidelity measurement is that it only requires a single H_c estimate per sampler circuit, an r-fold reduction relative to state-fidelity estimation. (Going beyond the small circuits studied here, cross-entropy estimation should scale better than fidelity estimation [13].) The information-fidelity data are also less noisy than the state data. However, a possible weakness of the definition given in Eq. (9) is that there is not much n dependence in Fig. 3, which does not seem physical.

The remarkable similarity between Figs. 2 and 3 was predicted in [13]. Here, we explain the connection in a different but equivalent way. Let us assume a depolarizing

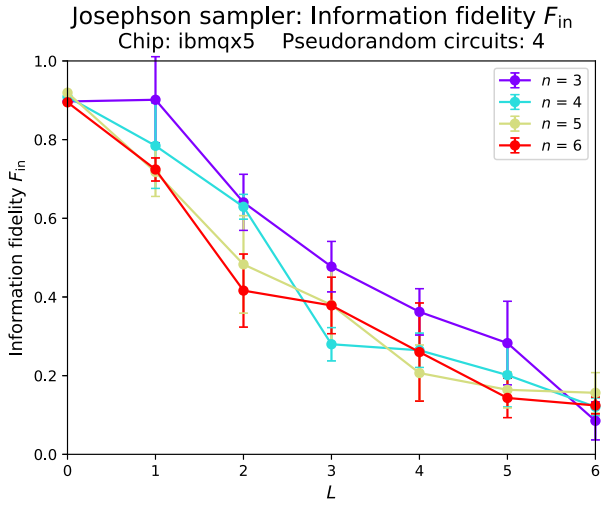

FIG. 3. The measured information fidelity given in Eq. (9) averaged over s pseudorandom circuits (the same circuits as used in Fig. 2).

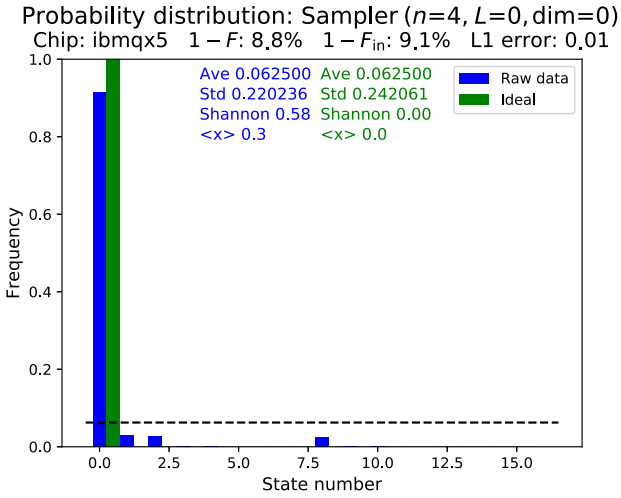

FIG. 4. The measured (blue) and simulated ideal (green) probability distributions resulting from state preparation immediately followed by read-out in the computational basis.

error model for the physical density matrix,

$$\rho = (1-\epsilon)\rho_{\text{ideal}} + \epsilon\rho_{\text{unif}}, \quad \rho_{\text{unif}} = I/N, \quad 0 \le \epsilon \le 1. \tag{10}$$

In this model, $p_{\text{meas}}(x) = (1-\epsilon)p_{\text{ideal}}(x) + \epsilon p_{\text{unif}}$ and, by a direct calculation, $F_{\text{in}} = 1-\epsilon$. We can say that the information fidelity is a direct measurement of the depolarizing error ϵ. Evaluation of Eq. (4) for the same error model gives

$$1 - F = (1 - F_{\text{in}})\left(1 - \frac{1}{N}\right). \tag{11}$$

Apart from a $1/N$ correction, where $N = 2^n$, the state and information fidelities are identical for a depolarizing channel. The relation shown in Eq. (11) is not expected to hold for other error models, but in the chaotic regime the physical errors become symmetrized, due to the action of a Haar random (or even two-design) circuit, to a depolarized form [13,30,31]. This is how we interpret the overall agreement between Figs. 2 and 3.

III. SAMPLING

Figure 4 shows the probability distribution $p(x)$ after preparing the state $|0000\rangle$ and then immediately measuring. This serves to measure the read-out errors and explain the subsequent figures. Here, $x \in \{0, 1, \ldots, N-1\}$ labels the N classical states, with $N = 2^n$. The dashed line is $1/N$. Each probability distribution $p(x)$ is separately characterized in four ways: the average probability or event frequency (Ave), which is always $1/N$, the width of the frequency distribution as measured by the standard deviation (Std), the classical entropy (Shannon), and the mean index ($\langle x \rangle$). We also quantify the difference between p_{meas} and p_{ideal} in three different ways: the state-fidelity loss $1-F$, the cross-entropy error $1-F_{\text{in}}$, and the L1 error,

$$\frac{1}{N}\sum_x |p_{\text{meas}} - p_{\text{ideal}}|, \tag{12}$$

which is the L1 distance divided by N. After one layer (Fig. 5), the probability distributions begin to spread out, and after four layers (Fig. 6), they appear to be highly scrambled but poorly correlated with each other. A probability distribution from the six-qubit sampler is shown in Fig. 7.

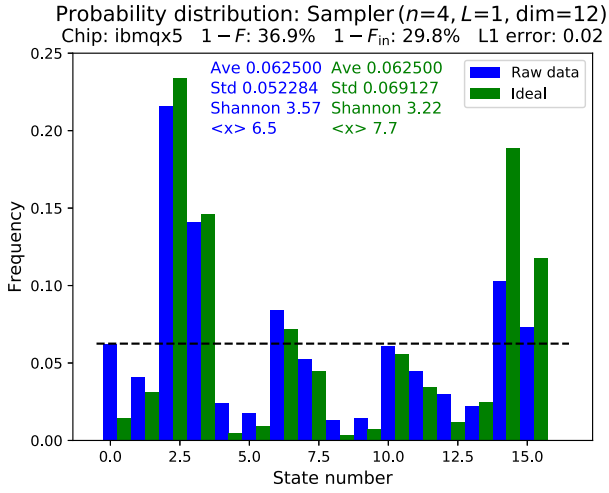

FIG. 5. The measured (blue) and ideal (green) probability distributions for a pseudorandom sampler of size $(n = 4, L = 1)$, which embeds a vector of dimension $m = 12$. The dashed line is $1/N$.

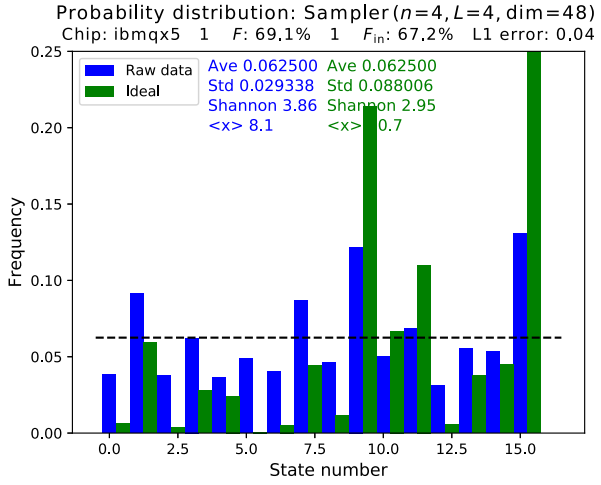

FIG. 6. The measured (blue) and simulated (green) probability distributions after four layers.

IV. ENTROPY AND ENTANGLEMENT

It is also interesting to measure the classical entropy generated by the sampler as a function of L; this is shown in Fig. 8. Note that measured entropies almost reach their maximum of n bits, but that the ideal entropies are about one bit short (we will discuss this in Sec. V). The data suggest that $n-1$ of the n bits of generated entropy come from the unitary dynamics of the circuit, and that decoherence does not have a strong effect on the entropy production (relative to its effect on fidelity). Perhaps this is because the entropy is already increasing very rapidly due to the unitary evolution.

There are several entanglement measures that we will study, most of which are forms of subsystem quantum entropies, where the subsystem is one of the qubits on the chain. It is simple to measure these quantities here because

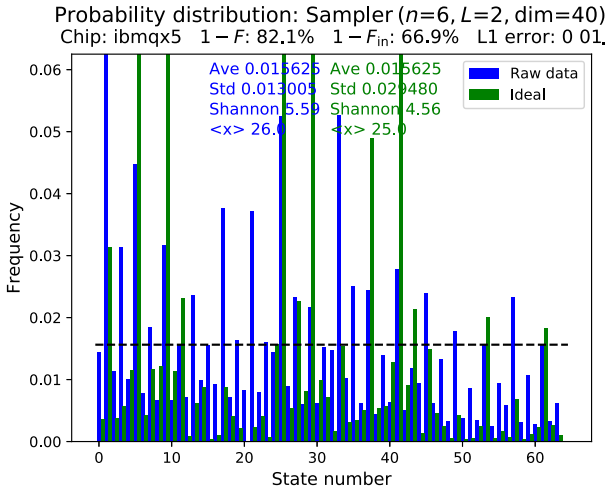

FIG. 7. The measured (blue) and ideal (green) probability distributions for a six-qubit sampler after two layers.

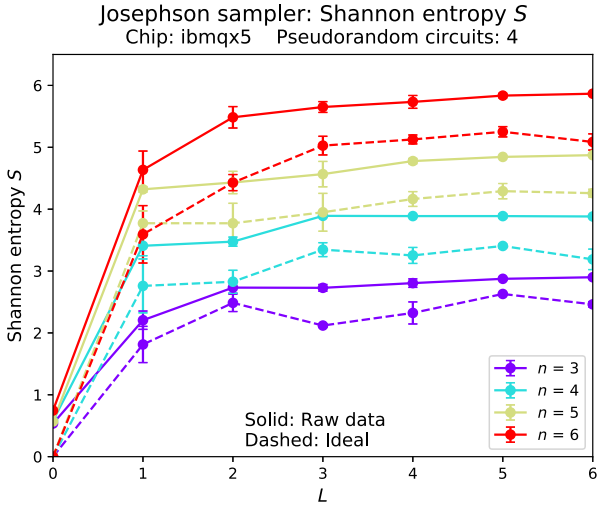

FIG. 8. The measured (solid curves) and simulated (dashed curves) entropy versus the number of sampler layers. The classical simulations includes the ideal circuit unitaries but no gate errors or decoherence.

we can directly reconstruct the single-qubit reduced density matrix

$$\rho_i = \mathrm{Tr}_{j\neq i}(\rho), \tag{13}$$

by tomography, which we carry out, one qubit at a time, tracing over (or not reading out) the other qubits. The average bipartite entanglement [5]

$$Q = 2(1-\overline{\gamma}), \quad \overline{\gamma} = \frac{1}{n}\sum_{i=1}^{n}\gamma_i, \quad \gamma_i = \mathrm{Tr}(\rho_i^2), \tag{14}$$

is plotted in Fig. 9. Here, γ_i is the purity of the reduced density matrix given in Eq. (13). For each sampler size (n, L), Q is measured for a set of s pseudorandom circuits, each requiring $3n$ measurements. The factor of 3 comes from the tomography operations required to measure each γ_i, and there are n of them to measure. (The data in Fig. 9 require the implementation and measurement of 1512 distinct circuits.) The results are encouraging, given that a high degree of entanglement is obtained after two layers. We can also look for an n dependence with regard to how quickly entanglement is achieved: we would normally expect that longer chains would entangle more slowly (after more circuit depth) than smaller ones, but the data show strikingly little n dependence.

A weakness of the entanglement measure given in Eq. (14) is that it is based on purity loss, which is also caused by decoherence. Thus, some of the entanglement that we are detecting is really decoherence. One indication of this is that in the $n=3$ and $n=4$ cases, the measured values exceed their Haar averages [5]

$$\langle Q\rangle_{\mathrm{Haar}} = \frac{N-2}{N+1}, \quad N = 2^n, \tag{15}$$

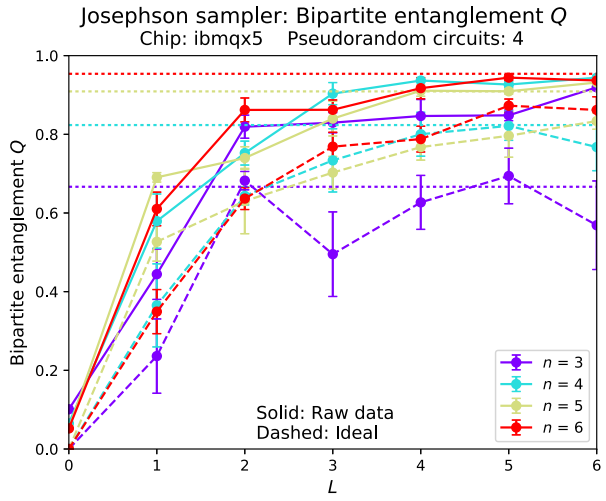

FIG. 9. The measured entanglement given in Eq. (14) averaged over s pseudorandom circuits (solid curves), along with their ideal classically simulated values (dashed). The horizontal dotted lines show the Haar averages given in Eq. (15).

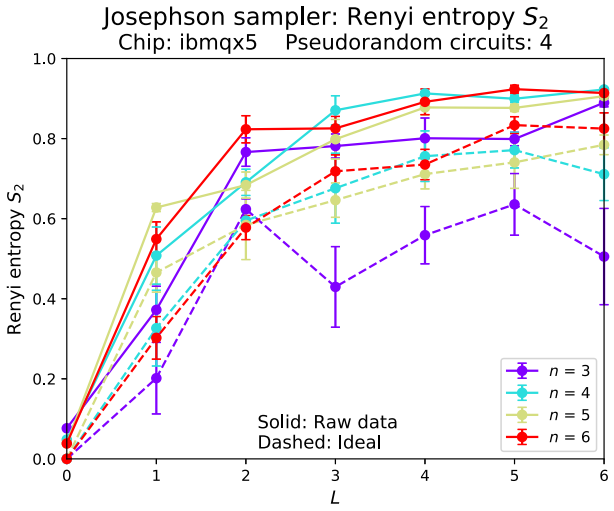

FIG. 10. The measured (solid curves) and simulated (dashed curves) Rényi entropy S_2, using the same data as presented in Fig. 9.

which are listed in Table I and plotted as horizontal dotted lines in Fig. 9. While exceeding the Haar average is theoretically possible, it is also exponentially unlikely due to measure concentration [32]. A second indication comes from comparing the data (solid curves) to a simulation (dashed curves) of the idealized problem with no gate errors and no decoherence. The measured data are typically 5–10% higher in Q. So we might conclude that as much as 10% of the Q we are observing is not genuine entanglement. This would still be quite impressive given how much entanglement is generated. However, it is easy to get misled by this entanglement measure because it can remain finite after the state has decohered (although it should eventually vanish when qubits relax to their nonentangled ground state $|0\rangle^{\otimes n}$).

Another entanglement measure derived from the single-qubit purity γ_i is the average second Rényi entropy

$$S_2 = -\frac{1}{n}\sum_{i=1}^{n}\log_2\gamma_i, \tag{16}$$

which is plotted in Fig. 10. The second Rényi entropy for a single qubit satisfies $0 \le S_2 \le 1$, the same bounds as for Q. Apart from the low-entanglement (small L) regime, we find that S_2 is almost identical to Q. The reason for the similarity is that the fluctuations in the γ_i are small, at least for $n > 3$ (note the small error bars in Fig. 9), so we can approximate $\gamma_i \approx \overline{\gamma}$. Then, linearizing S_2 about a reference purity γ_0, we have

$$S_2 \approx -\log_2(\gamma_0) + \frac{\gamma_0 - \overline{\gamma}}{\gamma_0 \ln 2} = a - b\overline{\gamma}, \tag{17}$$

where $a = \log_2(1/\gamma_0) + 1/\ln(2)$ and $b = 1/(\gamma_0 \ln 2)$. The linearized Rényi entropy given in Eq. (17) would be identical to Q if $a = b = 2$. Although this condition is not met for any choice of γ_0, the choice $\gamma_0 = 0.7$ leads to $a = 1.96$ and $b = 2.06$, which is quite close. Therefore, we can view the average bipartite entanglement given in Eq. (14) as a linearized second Rényi entropy, corrected to assure the desired behavior in the limits $\overline{\gamma} \to 1/2$ and $\overline{\gamma} \to 1$.

Finally, we determine an average entanglement entropy

$$S_e = -\frac{1}{n}\sum_{i=1}^{n}\mathrm{Tr}(\rho_i \log_2 \rho_i), \tag{18}$$

measured in bits. This is shown in Fig. 11. As a point of reference, we also plot the Haar-averaged entanglement entropy [33]

$$\langle S_e\rangle_{\mathrm{Haar}} = \frac{1}{\ln 2}\left[\sum_{k=d_B+1}^{d_A d_B}\frac{1}{k} - \frac{d_A - 1}{2d_B}\right] \tag{19}$$

for a bipartition of the chain into subsystems A and B with Hilbert-space dimensions $d_A = 2$ and $d_B = 2^{n-1}$. We note that the ideal entropies do not quite reach their Haar average values, while the measured ones exceed them. The entanglement entropy (in different contexts) was also measured in superconducting qubits in Refs. [34] and [35].

TABLE I. Haar average values of Q from Eq. (15).

n	3	4	5	6
$\langle Q\rangle_{\mathrm{Haar}}$	0.667	0.824	0.909	0.954

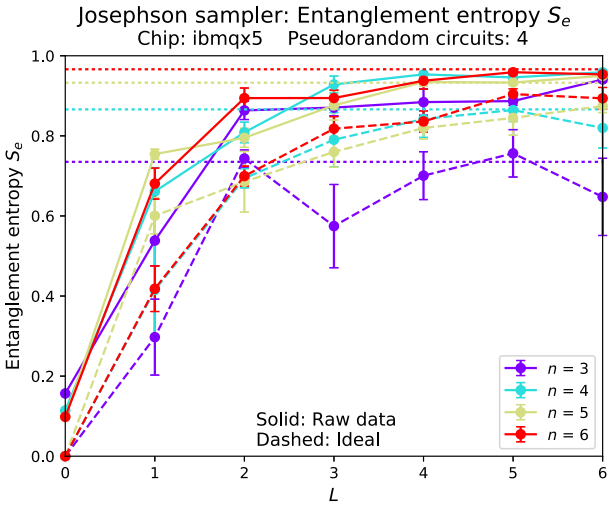

FIG. 11. The measured (solid curves) and simulated (dashed curves) entanglement entropy, using the same data as in Fig. 9.

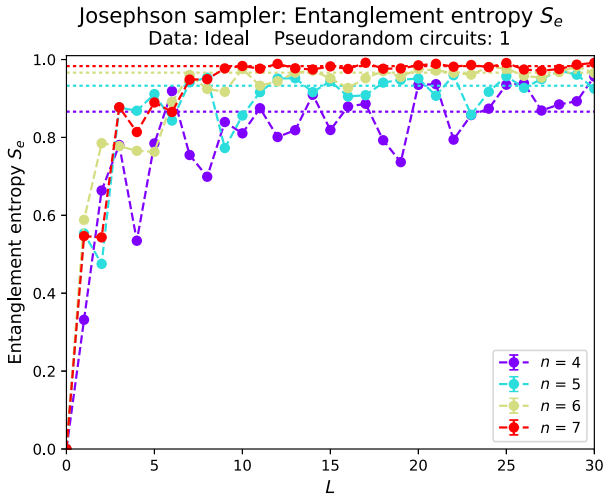

FIG. 12. The ideal entanglement entropy given in Eq. (18) for a single pseudorandom circuit (dashed curves). The horizontal dotted lines show the Haar averages given in Eq. (19). This is the same as Fig. 11, except that here we use only one circuit, explore larger n and L, and only include classical simulation results.

V. HAAR TYPICALITY

Next, we use the Josephson sampler to approximate Haar random unitaries. We will do this in a standard way, by inputting pseudorandom vectors $x \in \mathbb{R}^m$ of rotation angles into the circuit, thereby making the circuit itself pseudorandom. Our goal is to experimentally measure the quality of the resulting random unitaries and the scrambling of quantum information that they produce. There are different aspects of the random unitaries that we might want to assess. A measure of quality for an ensemble of random unitaries can be defined by regarding them as ϵ-approximate k-designs [7] and determining ϵ as a function of k. However, this might be more useful as a theoretical tool, where one can try to calculate ϵ versus k. Instead, we will focus on a different aspect of the random unitaries: whether or not they are Haar *typical*.

In this section, we will investigate an aspect of random unitaries inspired by the following proposition:

Conjecture 1 (Haar typicality). Let C be a quantum circuit on a register of n qubits for generating approximate Haar random unitaries $U \in U(2^n)$. Then, it is possible to experimentally validate C with only a single random instance.

In practical terms, we are saying that we can experimentally test whether or not a given quantum circuit is random. The complexity of C does not matter here, although validation will require a constant number of measurements of C, so in practice C will be efficient. The intuition for a single-instance diagnosis comes from the exponentially sharp distributions of Lipschitz-continuous functions of $U \in U(2^n)$ about their Haar averages [32]. This means that, with overwhelming probability, every U, when drawn uniformly from the unitary group, will be chaotic and will possess (some) universal attributes equal to their Haar-averaged values. The attributes are universal in the sense that they only depend on the register size n. In the upside-down world of random unitaries, all U's look alike, so in the error-free limit the Haar-typicality conjecture would follow from the results of Hayden and coworkers [32] on the concentration of entanglement entropy. The single-instance test would be to measure entanglement entropies on various bipartitions of the register and compare them with the Haar-averaged values. Our main assertion, then, is that a practical typicality test is possible in the presence of small but finite gate errors and decoherence.

Suppose that we were to use the entanglement entropy as shown in Eq. (18) to diagnose Haar typicality. In Fig. 12, we plot the ideal S_e for a single pseudorandom sampler. The entanglement entropy rapidly reaches the Haar-typical values, plotted as horizontal dotted lines. But, as we observe from Fig. 11, this measure is not sufficiently robust against gate errors and decoherence to diagnose Haar typicality. We can say that a necessary condition for a circuit to be Haar typical is a measured S_e at least as large as $\langle S_e \rangle_{\text{Haar}}$, but this condition is not sufficient. The bipartite entanglement given in Eq. (14) suffers from the same sensitivity to errors.

A better approach is to use the fact that Haar-typical unitaries are quantum chaotic [13,32], and to probe that chaos. First, we use the technique of Boixo *et al.* [13] and Neill *et al.* [14], and measure quantum chaos by its effect on the statistics of the sampled probability amplitudes. Probability density functions are shown in Figs. 13–15. For $n = 4$ and smaller, the data are very noisy, but show the main features that we find for larger n: There is good agreement for

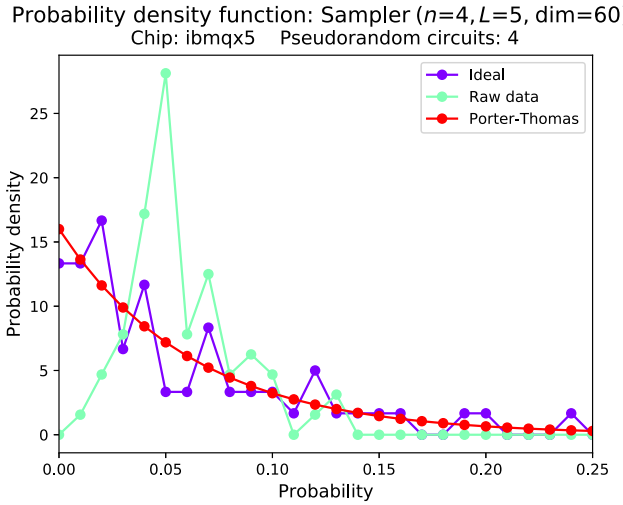

FIG. 13. The probability density function averaged over four pseudorandom samplers. The ideal data results from classical simulation of the circuits with no decoherence or gate errors. Here, $1/N = 0.0625$.

probabilities larger than $1/N$, but the frequencies of small probabilities are suppressed relative to the Porter-Thomas distribution. Furthermore, we do not find a clear convergence to the Porter-Thomas distribution as a function of L. The agreement is already quite good after two layers, as shown in Fig. 14 for six qubits, but it then declines as low-probability amplitudes become less frequent, as in Fig. 15.

Porter-Thomas (PT) statistics is also reflected in the classical entropy taking the universal value [13]

$$S_{\mathrm{PT}} = n - 1 + \gamma, \tag{20}$$

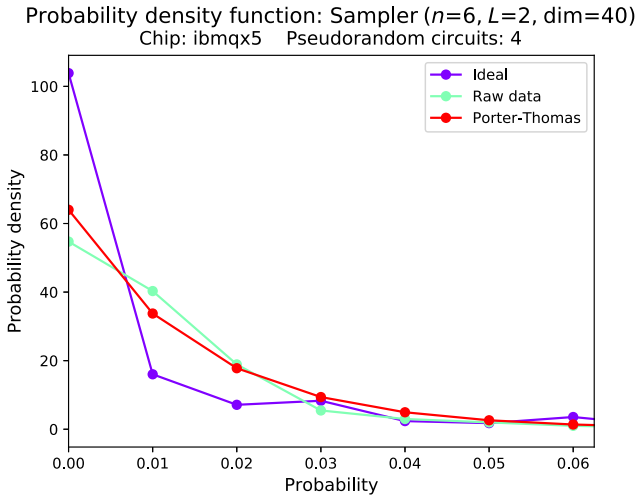

FIG. 14. The probability density function averaged over four pseudorandom samplers. In this case, $1/N = 0.0156$.

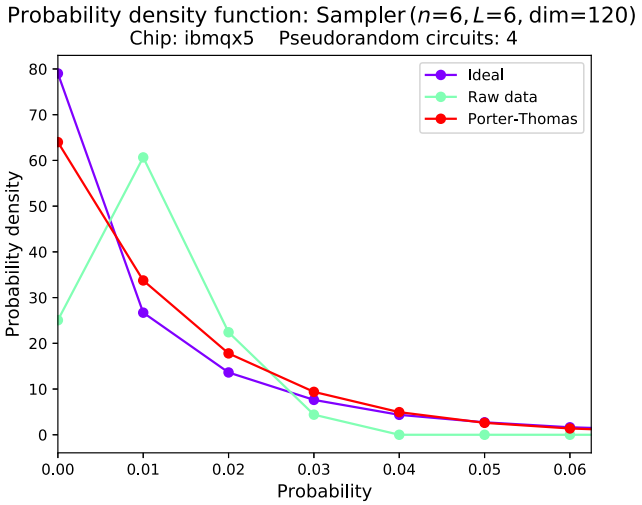

FIG. 15. The probability density function averaged over four pseudorandom samplers. $1/N = 0.0156$.

where $\gamma \approx 0.577$ is the Euler constant. In Fig. 16, we show the ideal sampler entropy converging to (almost) the universal values. Note that PT-chaotic circuits are not maximally randomizing, as their classical entropy generation is short by 0.423 bits. However, classical entropy measurement suffers from the same problem as do S_e and Q.

The probing of quantum chaos via Porter-Thomas statistics is probably best done with many circuit samples, and is not ideally suited for single-instance typicality testing. An alternative approach is to probe the butterfly effect, a dynamical manifestation of chaos. In the original setting

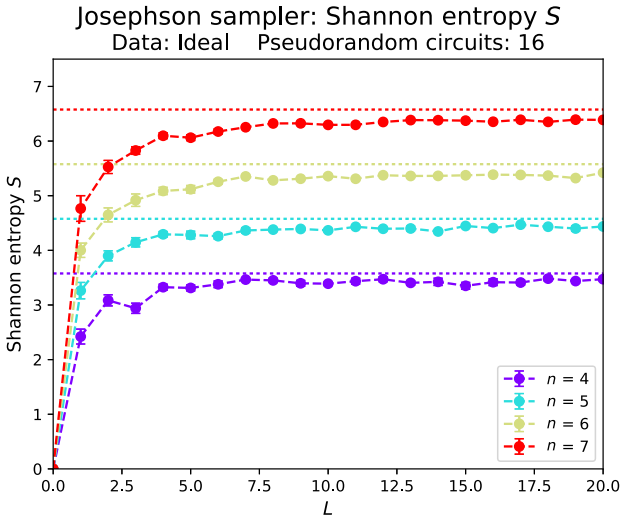

FIG. 16. The simulated (dashed curves) entropy versus the number of sampler layers. The dotted horizontal lines are the Porter-Thomas values from Eq. (20). In this plot, we average over 16 pseudorandom samples to allow a better comparison with Eq. (20).

[18,19], we consider a time-independent Hamiltonian H and a pair of local spatially nonoverlapping (hence commuting) Hermitian observables V and W. Consider the overlap between states

$$|a\rangle = W(t)V(0)|x\rangle \quad \text{and} \quad |b\rangle = V(0)W(t)|x\rangle, \tag{21}$$

that only differ in the order in which V and W are applied. Here, $W(t) = U^\dagger(t)WU(t)$, with $U(t) = e^{-iHt/\hbar}$ the time-evolution operator, and $|x\rangle$ is an initial classical state $x \in \{0,1\}^n$. The operation $W(t)$ propagates the state forward in time, applies W, and then propagates backwards in time. The states $|a\rangle$ and $|b\rangle$ differ by whether V is applied before this excursion or after it. If $t = 0$, $|a\rangle$ and $|b\rangle$ are identical, because V and W commute. The key idea—or definition—is that chaotic dynamics in $U(t)$ will make them fail to commute at later times, resulting in a decrease of the overlap $|\langle a|b\rangle|^2$, which we can measure through the four-point correlation function

$$|\langle x|W(t)V(0)W(t)V(0)|x\rangle|^2 = |\langle a|b\rangle|^2. \tag{22}$$

We will probe the chaos generated by a Josephson-sampler circuit by using it place of the time evolution, and using the inverse circuit for the reverse direction. For observables, we use single-qubit Pauli matrices Y_1 and X_n sitting at the ends of the chain, to mimic the classical limit $(V,W) \to (p,x)$. The butterfly effect, we consider is defined by a nonvanishing of

$$C = \mathrm{Tr}([W,V]^\dagger[W,V])/N, \tag{23}$$

where

$$V = Y_1, \quad W = U^\dagger X_n U, \tag{24}$$

with U and $U^\dagger$ provided by the sampler. Expanding the commutators leads to $C = 2(1-F)$, where

$$F = \frac{\mathrm{Tr}\, WVWV}{N} = \frac{1}{N}\sum_{x=0}^{N-1} G(x) \tag{25}$$

is real and

$$G(x) \equiv \langle x|WVWV|x\rangle \tag{26}$$

is complex. Here, $G(x)$ gives the contribution from computational basis state $x \in \{0,1\}^n$ to the trace F. Note that $|G(x)|^2$ is the four-point function and overlap discussed in Eq. (22). F has the standard form of an out-of-time-order correlator (OTOC) used to diagnose the butterfly effect, but with the time-evolution operator replaced by the Josephson sampler circuit. We will measure the absolute-value $|F|$, which the butterfly effect causes to vanish. An OTOC was also measured recently in a nuclear-magnetic-resonance quantum simulator [24].

The measurement of F would appear to require the measurement of $N = 2^n$ individual $G(x)$ correlators, one for each computational basis state x. But in the cases studied here, the real part of the complex quantity $G(x)$ is essentially independent of x, allowing us to estimate the trace F with a small set $\{x'\}$ of ν classical states chosen at random from $\{0,1\}^n$:

$$F \approx \frac{1}{\nu}\sum_{x\in\{x'\}} \mathrm{Re}\, G(x). \tag{27}$$

Because $\mathrm{Re}\, G(x)$ is approximately independent of x,

$$|F| \approx \frac{1}{\nu}\sum_{x\in\{x'\}} |\mathrm{Re}\, G(x)|. \tag{28}$$

In addition to the stochastic trace evaluation, we introduce a second approximation, by ignoring the imaginary part of $G(x)$. While the real part of $G(x)$ is essentially x-independent and changes in magnitude from 1 to 0 as scrambling develops, the imaginary part is small (typically <0.10 in the examples studied here) and random. Therefore, we use

$$|F| \approx \frac{1}{\nu}\sum_{x\in\{x'\}} |G(x)|. \tag{29}$$

Swingle *et al.* [21] discussed a similar approximation.

Before discussing the data, we will validate the approximation given in Eq. (29). In Fig. 17, we show classically simulated OTOCs for the pseudorandom Josephson sampler. The input data is a vector of pseudorandom real numbers in the range $[-2\pi, 2\pi]$. The dashed curves show exact values of F calculated from Eq. (25). The dotted curves

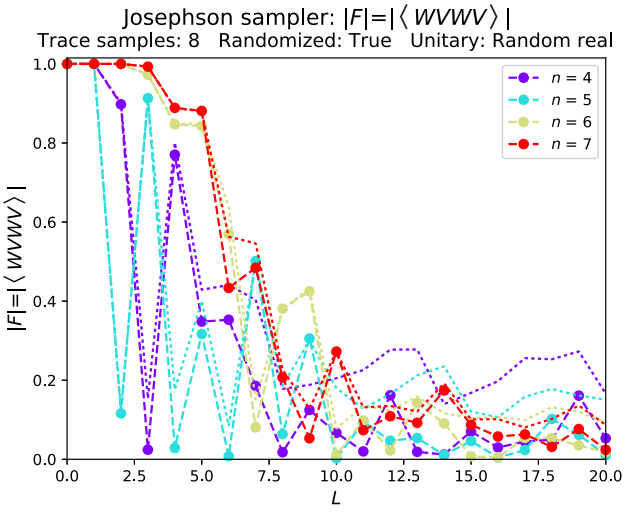

FIG. 17. Classically simulated OTOCs for the pseudorandom Josephson sampler. The dashed curves are exact, while the dotted curves use the approximation given in Eq. (29).

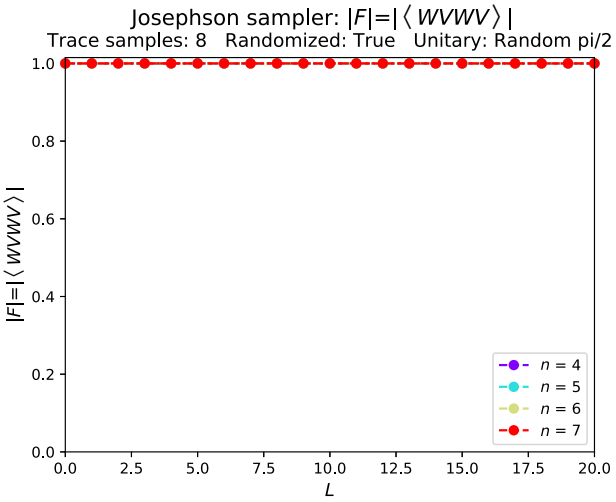

FIG. 18. Classically simulated OTOCs for the pseudorandom Josephson sampler with Clifford gates only. The dashed curves are exact, while the dotted curves use the approximation given in Eq. (29). Neither show scrambling (all curves are 1).

are the ideal results within the approximation given in Eq. (29), and we see that they easily diagnose the pronounced effect of scrambling in the chaotic regime. A second example, shown in Fig. 18, uses a *Clifford* sampler, i.e., a pseudorandom Josephson sampler with only Clifford gates. To realize this, we input a data vector consisting of random multiples of $\pi/2$ between -2π and 2π. In this example, the sampler does not exhibit OTOC decay. Although the scrambling and OTOC decay by unitary k-designs is a topic of current investigation [36,37], four-designs are expected to be sufficient for OTOC decay, whereas the ideal Clifford sampler should be a three-design [38], suggesting that $k \geq 4$ is actually necessary. The inclusion of random non-Clifford-gates, by instead using multiples of $\pi/4$, again leads to scrambling, as shown in Fig. 19.

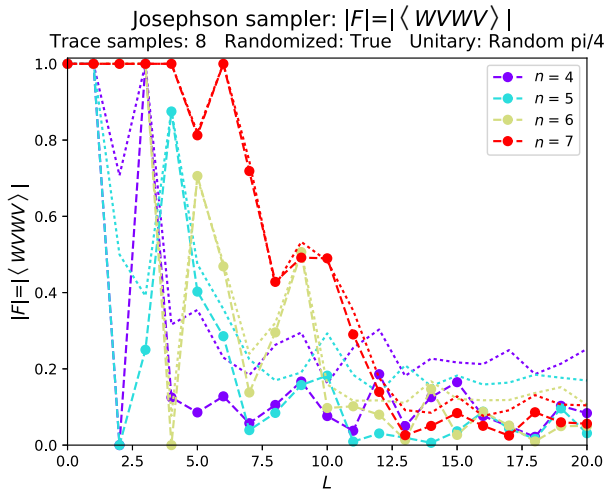

FIG. 19. The same as Fig. 18, but with single-qubit rotation angles given by random multiples of $\pi/4$.

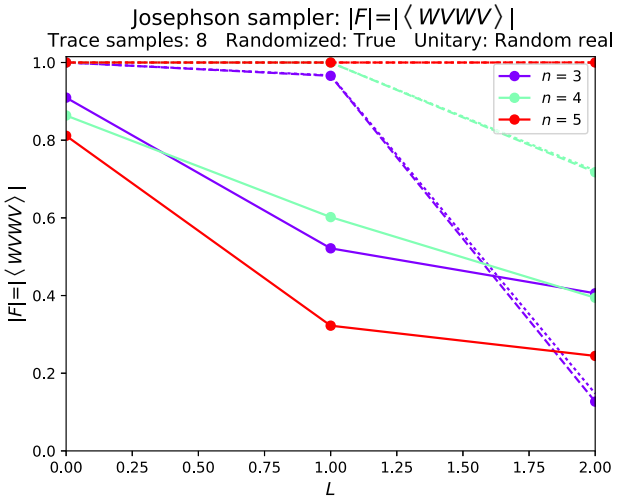

FIG. 20. The measured (solid curves) and classically simulated (dashed curves) OTOCs for the pseudorandom Josephson sampler. The dotted curves show the ideal result within the approximation given in Eq. (29).

These simulation results provide evidence that a rather small and shallow sampler circuit is capable of generating Haar-random unitaries of sufficient quality as to exhibit quantum chaos and scrambling. However, the measurement of each $|G(x)|$ requires a large complex circuit that, except for very small cases, exceeds the current size limit of the Quantum Experience API. We are able to measure OTOCs for samplers up to $n = 5$ and $L = 2$, as shown in Fig. 20. The data are consistent with scrambling, but how do we distinguish genuine unitary scrambling from simple gate errors and decoherence?

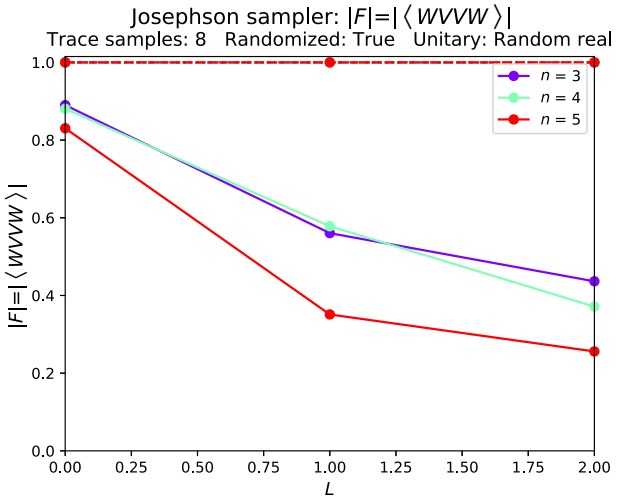

FIG. 21. The measured (solid curves) and classically simulated (dashed curves) OTOCs for the pseudorandom Josephson sampler. The dotted curves show the ideal result within the approximation given in Eq. (29).

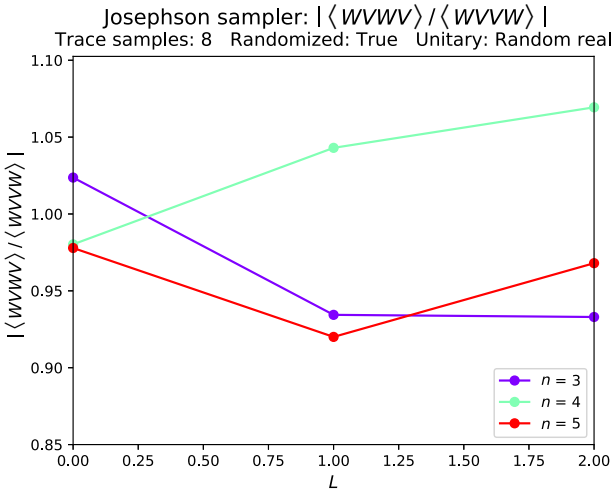

FIG. 22. Normalized (solid curves) OTOCs for the pseudorandom Josephson sampler.

In Fig. 21, we plot the measured four-point correlator $\langle WVVW\rangle$, which is ideally equal to one, and differs from Eq. (25) only in the order of the final two operators (VW versus WV). In the presence of genuine unitary scrambling, we would expect the decay in $\langle WVWV\rangle$ to be faster than in $\langle WVVW\rangle$, and in Fig. 22 we plot the ratio of these quantities, for a heuristic error-resistant metric. We conclude that the circuit fidelity is not yet sufficient to observe genuine unitary scrambling, but that it is not far off, giving support to the Haar-typicality conjecture.

VI. CONCLUSIONS

In this work, we study a general-purpose circuit for inputting classical information into a quantum computer, assuming a standard gateset and chain geometry. To assess the expressiveness of the circuit, we generate pseudorandom unitaries and experimentally estimate a variety of information and entanglement measures. This includes a study of the relationship between the state fidelity and the information-theoretic fidelity shown in Eq. (9), and introduced by Boixo *et al.* [13] and Neill *et al.* [14], which is easier to measure. We conjecture that it will be possible, with improvements in gate fidelity, to experimentally certify the presence of "high-quality" pseudorandom unitaries with only a single sample, by dynamic probes of chaos, such as the quantum butterfly effect. We also introduce and apply a practical method for measuring OTOCs in superconducting qubits. However, more work is needed to distinguish chaotic OTOC decay from noise [26–28].

ACKNOWLEDGMENTS

Data were taken on the IBM qx5 chip using the Quantum Experience API and the BQP software package developed by the author. The complete data set represented here consists of ∼3000 distinct circuits, each measured 8000 times. I am grateful to IBM Research and the IBM Quantum Experience team for making their devices available to the quantum-computing community. I also want to thank Sergio Boixo, Jerry Chow, Jordan Cotler, Andrew Cross, Charles Neill, Hanhee Paik, and Beni Yoshida for their private communications. I also thank Amara Katabarwa, Mingyu Sun, Jason Terry, and Phillip Stancil for their discussions and contributions to BQP. This work does not reflect the views or opinions of IBM or any of its employees.

Note added.—Recently, several papers have appeared on the question of distinguishing genuine unitary scrambling from simple gate errors and decoherence [26–28].

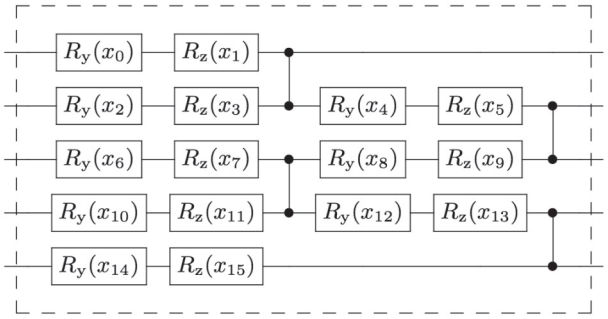

FIG. 23. The $n = 5$ Josephson sampler layer.

APPENDIX: THE JOSEPHSON SAMPLER CIRCUIT

Here, we provide additional details about the Josephson sampler. In Eq. (2), Y and Z are Pauli matrices. The circuit is a function of $x \in \mathbb{R}^m$, with each component being 4π-periodic. The design attempts to embed as many rotation parameters as possible, while making sure that there is no redundancy when applied to classical inputs or when

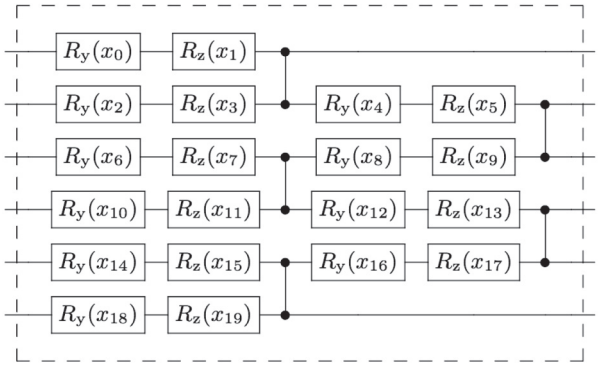

FIG. 24. The $n = 6$ Josephson sampler layer.

layered. On the imbqx5 chip, the CZ gate diag$(1, 1, 1, -1)$ is made from a CNOT gate and Hadamard gates. Explicit $L = 1$ circuits for $n = 5$ and 6 are shown in Figs. 23 and 24, which also show the particular mapping between the vector components and the gate angles used.

[1] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature **549**, 195 (2017).

[2] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, Nat. Commun. **5**, 4213 (2014).

[3] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Scalable Quantum Simulation of Molecular Energies, Phys. Rev. X **6**, 031007 (2016).

[4] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, 242 (2017).

[5] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, Science **302**, 2098 (2003).

[6] F. G. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, Commun. Math. Phys. **346**, 397 (2016).

[7] F. G. Brand ao, A. W. Harrow, and M. Horodecki, Efficient Quantum Pseudorandomness, Phys. Rev. Lett. **116**, 170502 (2016).

[8] D. P. DiVincenzo, M. Horodecki, D. W. Leung, J. A. Smolin, and B. M. Terhal, Locking Classical Correlations in Quantum States, Phys. Rev. Lett. **92**, 067902 (2004).

[9] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Randomizing quantum states: Constructions and applications, Comm. Math. Phys. **250**, 371 (2004).

[10] D. L. Lum, J. C. Howell, M. S. Allman, T. Gerrits, V. B. Verma, S. W. Nam, C. Lupo, and S. Lloyd, Quantum enigma machine: Experimentally demonstrating quantum data locking, Phys. Rev. A **94**, 022315 (2016).

[11] F. Dupuis, J. Florjanczyk, P. Hayden, and D. Leung, The locking-decoding frontier for generic dynamics, Proc. R. Soc. A **469**, 20130289 (2013).

[12] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theory Comput. **9**, 143 (2013).

[13] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[14] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, R. Barends, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, B. Foxen, R. Graff, E. Jeffrey, J. Kelly, E. Lucero, A. Megrant, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis, A blueprint for demonstrating quantum supremacy with superconducting qubits, Science **360**, 195 (2018).

[15] B. Peropadre, G. G. Guerreschi, J. Huh, and A. Aspuru-Guzik, Proposal for Microwave Boson Sampling, Phys. Rev. Lett. **117**, 140505 (2016).

[16] S. Goldstein, S. Korenblit, Y. Bendor, H. You, M. R. Geller, and N. Katz, Decoherence and interferometric sensitivity of boson sampling in superconducting resonator networks, Phys. Rev. B **95**, 020502 (2017).

[17] A. I. Larkin and Y. N. Ovchinnikov, Quasiclassical method in the theory of superconductivity, Sov. Phys. JETP **28**, 1200 (1969).

[18] S. H. Shenker and D. Stanford, Black holes and the butterfly effect, JHEP **3**, 67 (2014).

[19] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, JHEP **8**, 106 (2016).

[20] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Chaos in quantum channels, JHEP **2**, 4 (2016).

[21] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, Measuring the scrambling of quantum information, Phys. Rev. A **94**, 040302 (2016).

[22] G. Zhu, M. Hafezi, and T. Grover, Measurement of many-body chaos using a quantum clock, Phys. Rev. A **94**, 062329 (2016).

[23] N. Y. Yao, F. Grusdt, B. Swingle, M. D. Lukin, D. M. Stamper-Kurn, J. E. Moore, and E. Demler, Interferometric approach to probing fast scrambling, arXiv:1607.01801.

[24] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, Measuring Out-of-Time-Order Correlators on a Nuclear Magnetic Resonance Quantum Simulator, Phys. Rev. X **7**, 031011 (2017).

[25] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet, Nat. Phys. **13**, 781 (2017).

[26] B. Swingle and N. Y. Halpern, Resilience of scrambling measurements, arXiv:1802.01587.

[27] Y.-L. Zhang, Y. Huang, and X. Chen, Information scrambling in chaotic systems with dissipation, arXiv:1802.04492.

[28] K. A. Landsman, C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, and C. Monroe, Verified quantum information scrambling, arXiv:1806.02807.

[29] S. T. Flammia and Y.-K. Liu, Direct Fidelity Estimation from Few Pauli Measurements, Phys. Rev. Lett. **106**, 230501 (2011).

[30] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Symmetrized characterization of noisy quantum processes, Science **317**, 1893 (2007).

[31] E. Magesan, J. M. Gambetta, and J. Emerson, Characterizing quantum gates via randomized benchmarking, Phys. Rev. A **85**, 042311 (2012).

[32] P. Hayden, D. Leung, and A. Winter, Aspects of generic entanglement, arXiv:0407049.

[33] D. N. Page, Average Entropy of a Subsystem, Phys. Rev. Lett. **71**, 1291 (1993).

[34] C. Neill, P. Roushan, M. Fang, Y. Chen, M. Kolodrubetz, Z. Chen, A. Megrant, R. Barends, B. Campbell, B. Chiaro, A. Dunsworth, E. Jeffrey, J. Kelly, J. Mutus, P. J. J. O'Malley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. Polkovnikov, and J. M. Martinis, Ergodic dynamics and thermalization in an isolated quantum system, Nat. Phys. **12**, 1037 (2016).

[35] K. Xu, J.-J. Chen, Y. Zeng, Y. Zhan, C. Song, W. Liu, Q. Guo, P. Zhang, D. Xu, H. Deng, K. Huang, H. Wang, X. Zhu, D. Zheng, and H. Fan, Emulating Many-Body Localization with a Superconducting Quantum Processor, Phys. Rev. Lett. **120**, 050507 (2018).

[36] D. A. Roberts and B. Yoshida, Chaos and complexity by design, arXiv:1610.04903.

[37] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, Chaos, complexity, and random matrices, JHEP **11**, 48 (2017).

[38] Z. Webb, The Clifford group forms a unitary 3-design, Quantum Inf. Comput. **16**, 1379 (2016).