Multiplexed readout of transmon qubits with Josephson bifurcation amplifiers

V. Schmitt,[1] X. Zhou,[1] K. Juliusson,[1] B. Royer,[2] A. Blais,[2,3] P. Bertet,[1] D. Vion,[1] and D. Esteve[1]

[1]*Quantronics group, Service de Physique de l'Etat Condensé, CNRS UMR 3680, CEA Saclay, 91191 Gif-sur-Yvette, France*
[2]*Départment de Physique, Université de Sherbrooke, Sherbrooke, Québec, Canada*
[3]*Canadian Institute for Advanced Research, Toronto, Ontario, Canada*

Achieving individual qubit readout is a major challenge in the development of scalable superconducting quantum processors. We have implemented the multiplexed readout of a four transmon qubit circuit using nonlinear resonators operated as Josephson bifurcation amplifiers. We demonstrate the simultaneous measurement of Rabi oscillations of the four transmons. We find that multiplexed Josephson bifurcation is a high-fidelity readout method, the scalability of which is not limited by the need of a large-bandwidth, nearly quantum-limited amplifier as is the case with linear readout resonators.

　　　　　　　　　　PACS number(s): 03.67.Lx, 74.78.Na, 85.25.Cp

I. INTRODUCTION

Since the demonstration of quantum coherence in single Cooper pair boxes [1,2], the coherence time of superconducting quantum bits (qubits) has increased by orders of magnitude [3–5], and high-fidelity operation has been achieved [6,7]. Quantum speedup of the Deutsch-Josza [8], Grover search [9], and Shor's factorization [10] algorithms, as well as deterministic teleportation [11] and measurement-based entanglement [12,13] protocols, were recently demonstrated in circuits with a few (2–5) qubits. Nevertheless, no superconducting quantum processor able to run algorithms demonstrating the power of quantum computation [14] has been operated yet. Making operational processors with a large number of qubits faces the challenge of maintaining quantum coherence in complex circuits, of implementing multiple individual qubit readout, and of performing high-fidelity gates in parallel with quantum error correction. Much effort is presently devoted to solving these different scalability issues [15].

We address here the problem of simultaneous readout of transmon qubits [3] in a single shot. Readout of Josephson qubits is commonly performed by coupling each of them to a linear microwave resonator whose resonance frequency is shifted by a qubit-state-dependent value $\pm\chi$ [16]. Measuring the reflection or the transmission of a microwave pulse by the resonator then reveals the qubit state [17]. High-fidelity readout has been reached in several experiments [13,18] by using quantum-limited Josephson parametric amplifiers [19]. Besides, simultaneous readout of several qubits was achieved by using resonators with staggered frequencies, all coupled to a single line on which microwave readout pulses were frequency multiplexed [20]. However, reaching single-shot fidelity in this case requires parametric amplifiers with both large bandwidth to accommodate all of these frequencies and large saturation power to linearly amplify all simultaneous pulses. The recent implementation of this method in a four-transmon circuit [21] achieved fast readout with a fidelity compatible with surface-code error correction.

An alternative method for transmon readout that does not require a Josephson parametric amplifier consists in turning each readout resonator into a nonlinear one, operated as a Josephson bifurcation amplifier (JBA) [22–24]. Indeed, driving a JBA with a suitable microwave pulse yields a fast

and hysteretic transition between dynamical states with widely different field amplitude and phase, which can discriminate with high fidelity the transmon ground state $|0\rangle$ from its excited states $|i\rangle = |1\rangle$, $|2\rangle$. The determination of the dynamical state then requires a subsequent longer measuring time at a lower power level [25]. In this work, we demonstrate multiplexed high-fidelity single-shot readout of four transmons using the circuit described in Fig. 1.

II. EXPERIMENTAL SETUP

The chip consists of four cells, labeled $i = 1$–4, coupled to a single transmission line that carries the multiple qubit control and JBA signals. The sample is fabricated on a sapphire substrate in a two-step lithography process. The transmission line and the readout resonators are first patterned in a niobium film using optical lithography and reactive ion etching. The transmons and JBA junctions are then fabricated by electron lithography and double-angle evaporation of aluminum through a suspended shadow mask, with intermediate oxidation. It is measured in a dilution refrigerator with base temperature 30 mK. Each transmon B_i includes a superconducting quantum interference device (SQUID) [see Fig. 1(d)] that makes its $|0\rangle \leftrightarrow |1\rangle$ transition frequency f_{Bi}^{01} tunable with magnetic field [26]. In this experiment dedicated to readout, only a global magnetic field produced by a single coil can be applied to all transmons simultaneously. Each qubit is coupled to its JBA with a coupling constant $g_i/2\pi \simeq 85$ MHz. The JBAs have staggered frequencies f_{Ri} around 7.75 GHz separated by 61, 69, and 96 MHz, and quality factors of 2500, 2550, 2650, and 2200. All have the same Kerr nonlinearity [27] $K/2\pi \simeq -225$ kHz (that corresponds to a frequency reduction of 112.5 kHz per photon).

The qubits are controlled resonantly and we note θ_i^{kl} a rotation of qubit i by an angle θ between its states $|k\rangle$ and $|l\rangle$. The microwave control pulses at frequencies f_{Bi}^{kl} are obtained by translating the frequency f_{B0} of a single carrier using the technique of single-sideband mixing: using an IQ mixer, the carrier is multiplied by two signals I and Q delivered by an arbitrary waveform generator (AWG). I and Q are a sum of signals at frequencies $\delta_{Bi}^{kl} = f_{Bi}^{kl} - f_{B0}$, with suitable envelopes and phases. Another mixer is used to produce the JBA drive signals in the same way, that is by
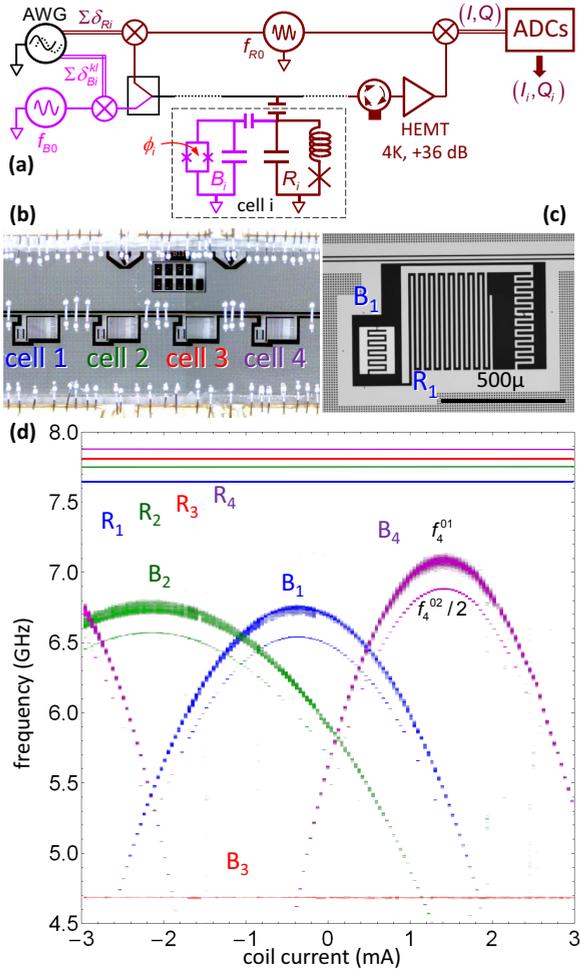
FIG. 1. (Color online) Readout of transmon qubits based on multiplexed JBAs. (a) Schematic electrical circuit. Four qubit-readout cells i (only one shown) are capacitively coupled to a microwave transmission line (black central line). Each cell is made of a tunable transmon qubit B_i [26] of transition frequencies f_{Bi}^{kl}, capacitively coupled to a JBA resonator R_i of frequency f_{Ri}. Control and readout pulses are produced and analyzed as described in the text. (b), (c) Optical micrographs showing (b) the measured chip with four cells, and (c) cell 1 with transmon B_1 and lumped element JBA R_1. (d) Spectroscopy of the four qubits B_i and readouts R_i as a function of the coil current inducing a global magnetic field. Frequencies f_{Ri} are indicated by lines, whereas qubit spectra are obtained by exciting the qubits with a 4 μs long single-frequency control pulse, reading out simultaneously the four JBAs, and color-plotting their switching probabilities.

IQ-mixing a carrier f_{R0} with a sum of signals at frequencies $\delta_{Ri} = f_{Ri} - f_{R0}$. Figure 1(a) illustrates the setup used. Driven at frequencies f_{Ri} chosen 9 MHz below their bare frequencies (with the transmon in its ground state), the JBAs switch at bifurcation from a state with average photon number 11 ± 1 to a state between 40 and 100 photons [27,28].

The transmon-JBA detuning $\Delta_i/2\pi = f_{Bi}^{01} - f_{Ri}$ determines both the readout sensitivity (through χ_i) and the Purcell energy relaxation rate $T_{P,i}^{-1} \simeq 2\pi f_{Ri}/Q_i(g_i/\Delta_i)^2$ of the qubit through the resonator input line [3]. The readout pulses have a first short 25 ns long step [see Fig. 2(b)] and a longer 2 μs
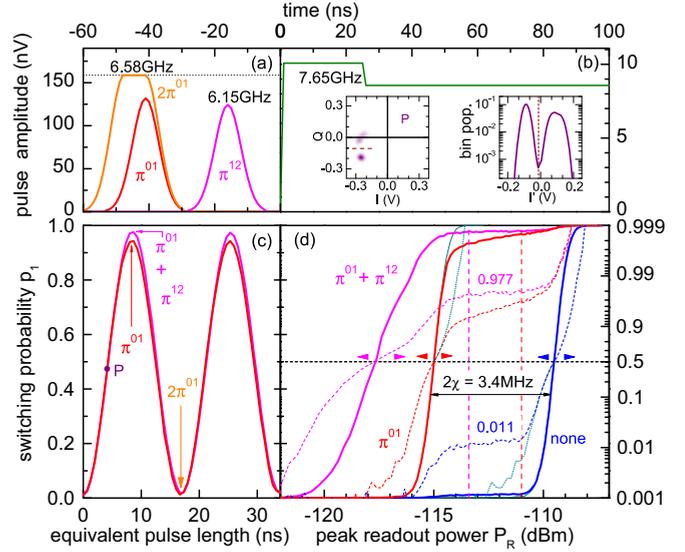
FIG. 2. (Color online) Characterization of cell 1 at detuning $\Delta_1/2\pi = 1.08$ GHz. (a) Microwave control pulse envelopes for π_1^{01}, π_1^{12}, and $(2\pi)_1^{01}$ rotations (see text). The dotted line shows the maximum amplitude used. (b) Beginning of the microwave readout pulse envelope (solid green line). Left inset: Density plot of (I_1,Q_1) obtained from 10^5 repetitions of a $(\pi/2)_1^{01}$ pulse [purple dot P in (c)] followed by a readout pulse. Right inset: Corresponding histogram (population in 10 mV wide bins) along the direction I' joining the two cloud centers. (c) Rabi oscillation of p_1 as a function of the equivalent control pulse length (duration of a rectangular pulse with maximum amplitude), without (red) and with (magenta) shelving (see text). (d) Probability p_1 with no qubit control pulse (blue) and after a π_1^{01} pulse alone (red) or with shelving (magenta). Solid lines represent p_1 on a linear scale (left axis) whereas dashed and dotted lines show it using a double-logarithmic scale below and above 0.5 (right axis). Thin solid and dotted lines represent "ideal" S curves (see text). The vertical dashed lines indicate the pulse power yielding the highest readout contrasts with (left) and without (right) shelving.

latching step at 85% of the peak power. In practice, bifurcation develops (or not) between 50 ns and 500 ns, whereas I and Q are averaged between 325 ns and 1325 ns. Readout pulses can overlap in time [see for instance Fig. 4(a)] so that the output signal contains contributions of different JBAs. In order to extract these contributions, the output signal is demodulated in two steps: An analog demodulation at the readout carrier frequency f_{R0} is first performed; the resulting signal is then digitized at 2 Gsample/s with a 1 GHz analog bandwidth that widely covers the 250 MHz frequency range spanned by the four JBAs; it is then demodulated numerically by a dedicated PC, which directly multiplies it with cosine functions at frequencies δ_{Ri} and averages the result. The outcome of a readout sequence is four points (I_i,Q_i) in the in-phase and quadrature plane (one for each JBA frequency f_{Ri}), as shown in Fig. 3(b). The outcome of a readout sequence is thus four points (I_i,Q_i) in the in-phase and quadrature plane.

Spectroscopic data of the qubits and readout resonators as a function of the coil current are displayed in Fig. 1(d). These data were recorded at high excitation power to show spectroscopic lines at both f_{Bi}^{01} and $f_{Bi}^{02}/2$. Frequencies f_{Bi}^{01} of tunable qubits $B_{1,2,4}$ peak at about 0.7–1 GHz below the
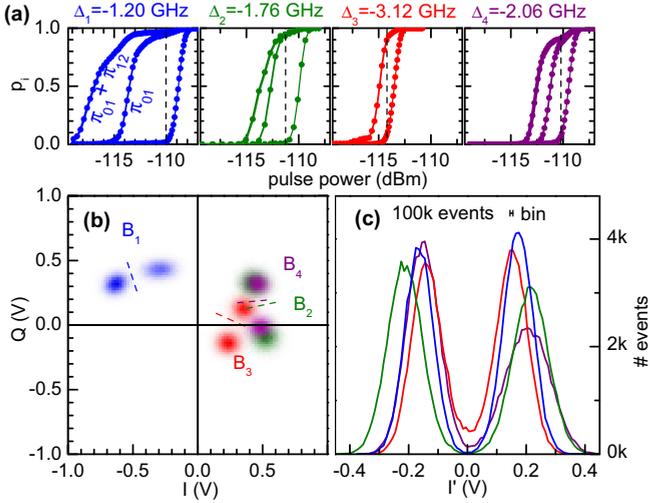
FIG. 3. (Color online) Simultaneous readout of the four qubits at a magnetic field such that $\Delta_{1-4}/2\pi = (-1.2, -1.76, -3.12, -2.06)$ GHz. (a) Switching probabilities p_i of the four readouts as a function of readout power P_R, after no control pulse (right curves in each panel), and after a π_i^{01} pulse without (middle) and with shelving (left curves for $B_{1,2,3}$). Dashed vertical lines indicate the optimal readout powers used in (b) and (c) and in Fig. 4 (shelving used only for B_2 and B_4). (b) Density plots of the four (I_i, Q_i) obtained from 10^5 measurements. Segments indicate the separatrices between switching and nonswitching events. (c) Corresponding histograms along the lines perpendicular to separatrices.

frequency of their respective JBA, and the anharmonicity $\alpha = f_{Bi}^{12} - f_{Bi}^{01} \simeq -434 \pm 2$ MHz. The measured relaxation times of all transmons are found to be in the range $T_1 = 1.7$–3.2 μs for $|\Delta_i/2\pi| \gtrsim 1$ GHz. This is significantly below the Purcell limit $T_P > 8$ μs and shorter than in comparable 2D transmon circuits [21], probably due to dielectric losses [29].

III. READOUT PERFORMANCE

All qubit-readout cells yielded similar performances at equal detuning Δ_i. Performance of cell 1, operated at a qubit-JBA detuning $\Delta_1/2\pi = -1.08$ GHz, is summarized in Fig. 2. All qubit control pulses have 3σ long Gaussian rises and falls with $\sigma = 4$ ns, as shown in Fig. 2(a). Numerical simulations of the transmon dynamics including its three lowest levels show that such control pulses do not introduce preparation errors larger than 0.1% [30]. Readout is performed either immediately after applying a θ_1^{01} Rabi pulse, or after a subsequent π_1^{12} pulse that shelves the excited qubit in state $|2\rangle$, as in [23]. This shelving decreases the error made in measuring the excited qubit by blocking its relaxation down to state $|0\rangle$ before the measurement is completed [31].

The density plot of (I_1, Q_1) obtained from 10^5 repetitions of the readout after a $(\pi/2)_1^{01}$ pulse is shown in the left inset of Fig. 2(b). The two clouds with a small relative overlap of order 10^{-5} (estimated from the corresponding histogram in the right inset) reveal an excellent discrimination of the JBA states. The fidelity of the qubit to JBA mapping is investigated by measuring the variations of the switching probability p_1 as

a function of the peak readout power P_R. These so-called S curves are shown in Fig. 2(d) in three different cases: when the qubit is left in its ground state $|0\rangle$ with no applied control pulse (blue), after a π_1^{01} pulse aiming at preparing state $|1\rangle$ (red), and after a π_1^{01} pulse followed by a π_1^{12} shelving pulse (magenta). One observes that the S curves for the two states $|0\rangle$ and $|1\rangle$ are separated in P_R by about 5.5 dB (or equivalently by $2\chi = 3.4$ MHz in resonator or drive frequency), which is much larger than the 2.4 dB (1.5 MHz) width of the ground-state S curve, defined here by $1\% < p < 99\%$. This result implies that, in the absence of preparation errors and relaxation before and during measurement, readout errors would be negligible.

In practice, at the optimal powers P_R [see Fig. 2(d)], the measured total errors are 1.1% for $|0\rangle$, and 3.1% and 2.2% for $|1\rangle$ without and with shelving, respectively. These errors result from two effects. First, the 1.1% error in the ground state is due to a residual thermal excitation of the qubit (corresponding to a qubit temperature of 70 mK), as evidenced by the flat shoulder on the ground-state S curve at low power. This spurious excitation is also responsible for the same absolute 1.1% error in preparing state $|1\rangle$. The remaining errors in $|1\rangle$ are thus 2.0% and 1.1% without and with shelving. Second, numerical simulations including relaxation during the control pulses, using the independently measured relaxation time $(\Gamma_1^{10})^{-1} = 2.0$ μs, account for absolute errors of 0.6% and 1.1% without and with shelving. The shelving case is thus fully understood: errors in $|1\rangle$ are only due to thermal population and relaxation at preparation, and relaxation to $|0\rangle$ during readout is efficiently blocked as proven by the horizontal plateau at $p_1(P_R) \simeq 0.98$. The intrinsic readout fidelity with shelving is thus excellent.

Without shelving, the remaining readout error is 1.4% at the optimal P_R, but with a slow increase of $p_1(P_R)$ as it approaches 1. This behavior is not understood and quantum simulation of the JBA + excited qubit dynamics is needed to address this question. Nevertheless, to infer what would be the intrinsic readout fidelity in the absence of preparation errors and extra relaxation at readout, we reconstruct ideal S curves: for the ground state, the lower part $p_1 \leqslant 0.5$ is replaced by the S curve measured for the qubit excited state and shifted in power to remove the effect of residual thermal excitation; for the excited state, its upper part $p_1 > 0.5$ is replaced by the one measured in the ground state and shifted in power to remove the effect of relaxation at readout. These ideal S curves, illustrated by thin solid and dotted lines in Fig. 2(d), give intrinsic readout errors lower than 2×10^{-3} both for the ground and excited states.

We now discuss the simultaneous readout of the four qubits. Given the lack of individual transmon tunability, a magnetic field leading to not too large detunings $\Delta_i/2\pi = (-1.2, -1.76, -3.12, -2.06)$ GHz was applied. In addition, longer readout pulses with 50 ns measurement step and 2 μs latching step are used, as shown in Fig. 4(a). The measurement outcomes for the four qubits prepared with control pulses close to $(\pi/2)_{01}$ and with π_{12} shelving only for B_2 and B_4 [32] are shown in Fig. 3. The density plots in the (I_i, Q_i) planes are shown with their best separatrix between switching and nonswitching events. As illustrated in Fig. 3(c), the switching histograms measured along an axis perpendicular to the separatrix show a good separation, albeit smaller than obtained at the optimal working point of each
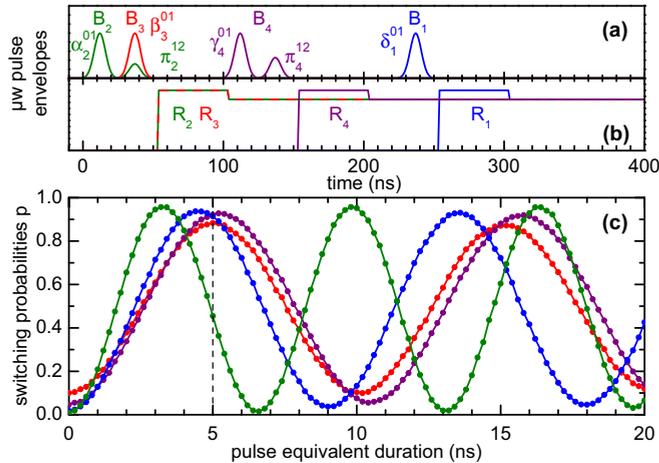
FIG. 4. (Color online) Simultaneous measurement of Rabi oscillations of the four qubits at readout powers indicated in Fig. 3. Microwave control (a) and readout (b) pulse envelopes used at 5 ns equivalent Rabi pulse duration. Only B_2 and B_4 are shelved on their second excited levels before readout. (c) Simultaneous Rabi oscillations of p_{1-4} as a function of the equivalent control pulse duration.

cell. The standard deviation of the switching probability for all cells was checked to decrease as expected for independent events.

Having characterized simultaneous readout of the four qubits, we now include qubit drive (see Fig. 4). For simplicity, the control pulses are not applied simultaneously in order to avoid having to take into account the qubit ac Stark shift resulting from other qubit drives. The control and readout pulses are shown in Fig. 4(a). The switching curves of the four JBAs are shown in Fig. 4(b) after no qubit pulse, and after a π_i^{01} pulse without or with shelving. Rabi oscillations of the four qubits, measured at the optimal powers indicated in Fig. 4(b),

are shown in Fig. 4(c). These data show that JBA readout is compatible with qubit driving and simultaneous multiplexed operation. The overall performance of our multiplexed JBA is thus comparable with that achieved using linear dispersive readout and parametric amplifiers [21], albeit with larger errors *not* due to the readout method itself.

IV. CONCLUSION

A natural question that arises is the maximum number of transmons that multiplexed JBA could handle. Indeed, due the nonlinear character of JBAs, bifurcation of a given JBA can be affected by the dynamics of other JBAs that are close in frequency. How close their frequencies can be without inducing readout crosstalk is not known. In the present setup, this phenomenon was quantified by preparing B_1 in $|0\rangle$ or $|1\rangle$ and B_2 in a superposition $(|0\rangle + |1\rangle)/\sqrt{2}$. The difference between the values of p_2 for the two B_1 states gives a crosstalk of only $0.2\% \pm 0.05\%$. This low value shows that a JBA frequency separation of 60 MHz is conservative, and therefore that more qubits could be read out in parallel.

In conclusion, multiplexed JBA readout of transmons has an excellent intrinsic readout fidelity when shelving is used, and is compatible with driving and reading transmons in a small qubit register. Its scalability, limited by the interactions between JBAs with close frequencies, is still under investigation, but the present results suggest that reading out a ten-qubit register is possible.

[1] Y. Nakamura, Yu. A. Pashkin, and J. S. Tsai, Nature (London) **398**, 786 (1999).

[2] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret, Science **296**, 886 (2002).

[3] J. Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **76**, 042319 (2007).

[4] J. A. Schreier, A. A. Houck, Jens Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. B **77**, 180502 (2008).

[5] Hanhee Paik, D. I. Schuster, Lev S. Bishop, G. Kirchmair, G. Catelani, A. P. Sears, B. R. Johnson, M. J. Reagor, L. Frunzio, L. I. Glazman, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Phys. Rev. Lett. **107**, 240501 (2011).

[6] A. Dewes, F. R. Ong, V. Schmitt, R. Lauro, N. Boulant, P. Bertet, D. Vion, and D. Esteve, Phys. Rev. Lett. **108**, 057002 (2012).

[7] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis, Nature (London) **508**, 500 (2014).

[8] T. Yamamoto, M. Neeley, E. Lucero, R. C. Bialczak, J. Kelly, M. Lenander, M. Mariantoni, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, Y. Yin, A. N. Cleland, and J. M. Martinis, Phys. Rev. B **82**, 184515 (2010).

[9] A. Dewes, R. Lauro, F. R. Ong, V. Schmitt, P. Milman, P. Bertet, D. Vion, and D. Esteve, Phys. Rev. B **85**, 140503 (2012).

[10] Erik Lucero, R. Barends, Y. Chen, J. Kelly, M. Mariantoni, A. Megrant, P. O'Malley, D. Sank, A. Vainsencher, J. Wenner, T. White, Y. Yin, A. N. Cleland, and John M. Martinis, Nat. Phys. **8**, 719 (2012).

[11] L. Steffen, Y. Salathe, M. Oppliger, P. Kurpiers, M. Baur, C. Lang, C. Eichler, G. Puebla-Hellmann, A. Fedorov, and A. Wallraff, Nature (London) **500**, 319 (2013).

[12] D. Ristè, M. Dukalski, C. A. Watson, G. de Lange, M. J. Tiggelman, Ya. M. Blanter, K. W. Lehnert, R. N. Schouten, and L. DiCarlo, Nature (London) **502**, 350 (2013).

[13] N. Roch, M. E. Schwartz, F. Motzoi, C. Macklin, R. Vijay, A. W. Eddins, A. N. Korotkov, K. B. Whaley, M. Sarovar, and I. Siddiqi, Phys. Rev. Lett. **112**, 170501 (2014).

[14] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[15] M. H. Devoret and R. J. Schoelkopf, Science **339**, 1169 (2013).

[16] A. Blais, J. Gambetta, A. Wallraff, D. I. Schuster, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, Phys. Rev. A **75**, 032329 (2007).

[17] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. Lett. **95**, 060501 (2005).

[18] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. M. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff, L. Frunzio, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Nature (London) **511**, 444 (2014).

[19] M. A. Castellanos-Beltran, K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert, Nat. Phys. **4**, 929 (2008).

[20] M. Jerger, S. Poletto, P. Macha, U. Hübner, E. Il'ichev, and A. V. Ustinov, Appl. Phys. Lett. **101**, 042604 (2012).

[21] Evan Jeffrey, Daniel Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and John M. Martinis, Phys. Rev. Lett. **112**, 190504 (2014).

[22] V. E. Manucharyan, E. Boaknin, M. Metcalfe, R. Vijay, I. Siddiqi, and M. H. Devoret, Phys. Rev. B **76**, 014524 (2007).

[23] F. Mallet, Florian R. Ong, A. Palacios-Laloy, F. Nguyen, P. Bertet, D. Vion, and D. Esteve, Nat. Phys. **5**, 791 (2009).

[24] P. C. de Groot, A. F. van Loo, J. Lisenfeld, R. N. Schouten, A. Lupascu, C. J. P. M. Harmans, and J. E. Mooij, Appl. Phys. Lett. **96**, 123508 (2010).

[25] Because a new readout can only start after the resonator field has decayed, this method is better suited for single-shot readout than for continuous measurements of a qubit [20].

[26] B_3 has one of its two Josephson junctions open and is consequently not tunable.

[27] F. R. Ong, M. Boissonneault, F. Mallet, A. Palacios-Laloy, A. Dewes, A. C. Doherty, A. Blais, P. Bertet, D. Vion, and D. Esteve, Phys. Rev. Lett. **106**, 167002 (2011).

[28] M. Boissonneault, A. C. Doherty, F. R. Ong, P. Bertet, D. Vion, D. Esteve, and A. Blais, Phys. Rev. A **85**, 022305 (2012).

[29] C. M. Quintana, A. Megrant, Z. Chen, A. Dunsworth, B. Chiaro, R. Barends, B. Campbell, Yu Chen, I.-C. Hoi, E. Jeffrey, J. Kelly, J. Y. Mutus, P. J. J. O'Malley, C. Neill, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and John M. Martinis, Appl. Phys. Lett. **105**, 062601 (2014).

[30] Consequently, we do not use the DRAG correction technique used in J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **82**, 040305 (2010).

[31] This blockade occurs because the direct relaxation from state $|2\rangle$ to $|0\rangle$ involves a very small matrix element; consequently, the decay to $|0\rangle$ results from the two relaxation processes $|2\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow |0\rangle$ in series, and varies quadratically at short time, as $1 - (\Gamma^{\mathrm{eff}} t)^2$ with $\Gamma^{\mathrm{eff}} = \sqrt{\Gamma_1^{21}\Gamma_1^{10}} \sim \sqrt{2}\Gamma_1^{10}$. Note however that shelving takes the qubit out of the computational Hilbert space and thus makes the readout destructive.

[32] Shelving was neither used for qubit B_1 because f_1^{12} being too close from f_4^{01}, it would have driven qubit B_4, nor for qubit B_3 whose frequency f_3^{12} was outside the bandpass of the driving line.