This is the accepted manuscript made available via CHORUS. The article has been published as:

# Restless Tuneup of High-Fidelity Qubit Gates

M. A. Rol, C. C. Bultink, T. E. O'Brien, S. R. de Jong, L. S. Theis, X. Fu, F. Luthi, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, D. Deurloo, R. N. Schouten, F. K. Wilhelm, and L. DiCarlo

Phys. Rev. Applied **7**, 041001 — Published 24 April 2017

DOI: 10.1103/PhysRevApplied.7.041001

# Restless tuneup of high-fidelity qubit gates

M. A. Rol,[1,2] C. C. Bultink,[1,2] T. E. O'Brien,[3] S. R. de Jong,[1,2] L. S. Theis,[4]
X. Fu,[1] F. Luthi,[1,2] R. F. L. Vermeulen,[1,2] J. C. de Sterke,[5,1] A. Bruno,[1,2]
D. Deurloo,[6,1] R. N. Schouten,[1,2] F. K. Wilhelm,[4] and L. DiCarlo[1,2]

[1] *QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands*
[2] *Kavli Institute of Nanoscience, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands*
[3] *Instituut-Lorentz for Theoretical Physics, Leiden University, Leiden, The Netherlands*
[4] *Theoretical Physics, Saarland University, 66123 Saarbrucken, Germany*
[5] *Topic Embedded Systems B.V., P.O. Box 440, 5680 AK Best, The Netherlands*
[6] *Netherlands Organisation for Applied Scientific Research (TNO), P.O. Box 155, 2600 AD Delft, The Netherlands*

We present a tuneup protocol for qubit gates with tenfold speedup over traditional methods reliant on qubit initialization by energy relaxation. This speedup is achieved by constructing a cost function for Nelder-Mead optimization from real-time correlation of non-demolition measurements interleaving gate operations without pause. Applying the protocol on a transmon qubit achieves 0.999 average Clifford fidelity in one minute, as independently verified using randomized benchmarking and gate set tomography. The adjustable sensitivity of the cost function allows detecting fractional changes in gate error with nearly constant signal-to-noise ratio. The restless concept demonstrated can be readily extended to the tuneup of two-qubit gates and measurement operations.

Reliable quantum computing requires the building blocks of algorithms, quantum gates, to be executed with low error. Strategies aiming at quantum supremacy without error correction [1, 2] require $\sim 10^3$ gates, and thus gate errors $\sim 10^{-3}$. Concurrently, a convincing demonstration of quantum fault tolerance using the circuits Surface-17 and -49 [3, 4] under development by several groups worldwide requires gate errors one order of magnitude below the $\sim 10^{-2}$ threshold of surface code [5, 6].

The quality of qubit gates depends on qubit coherence times and the accuracy and precision of the pulses realizing them. With the exception of a few systems known with metrological precision [7], pulsing requires meticulous calibration by closed-loop tuning, i.e., pulse adjustment based on experimental observations. Numerical optimization algorithms have been implemented to solve a wide range of tuning problems with a cost-effective number of iterations [8–13]. However, relatively little attention has been given to quantitatively exploring the speed and robustness of the algorithms used. This becomes crucial with more complex and precise quantum operations, as the number of parameters and requisite precision of calibration grow.

Though many aspects of tuning qubit gates are implementation independent, some details are specific to physical realizations. Superconducting transmon qubits are a promising hardware for quantum computing, with gate times already exceeding coherence times by three orders of magnitude. Conventional gate tuneup relies on qubit initialization, performed passively by waiting several times the qubit energy-relaxation time $T_1$ or actively through feedback-based reset [14]. Passive initialization becomes increasingly inefficient as $T_1$ steadily increases [15, 16], while feedback-based reset is technically involved [17].

In this Letter, we present a gate tuneup method that dispenses with $T_1$ initialization and achieves tenfold speedup over the state of the art [9] without active reset. Restless tuneup exploits the real-time correlation of quantum-non-demolition (QND) measurements to interleave gate operations without pause, and the evaluation of a cost function for numerical optimization with adjustable sensitivity at all levels of gate fidelity. This cost function is obtained from a simple modification of the gate sequences of conventional randomized benchmarking (CRB) to penalize both gate errors within the qubit subspace and any leakage from it. We quantitatively match the signal-to-noise ratio of this cost function with a model that includes measured $T_1$ fluctuations. Restless tuneup robustly achieves $T_1$-dominated gate fidelity of 0.999, verified using both CRB with $T_1$ initialization and a first implementation of gate set tomography (GST) [18] in a superconducting qubit. While this performance matches that of conventional tuneup, restless is tenfold faster and converges in one minute.

In many tuneup routines [Fig. 1(a)], the relevant information from the measurements can be expressed as the fraction $\varepsilon$ of non-ideal outcomes ($m_n$). In conventional gate tuneup, a qubit is repeatedly initialized in the ground state $|0\rangle$, driven by a set of gates ($\{G\}$) whose net operation is ideally identity, and measured [Fig. 1(b)]. The conventional cost function is the raw infidelity,

$$\varepsilon_{\mathrm{C}} = \sum_{n=1}^{N} (m_n \neq 0)/N.$$

The central idea of restless tuning [Fig. 1(c)] is to remove the time-costly initialization step, by measuring the correlation between subsequent QND measurements and interleaving gate operations without any rest [19]. For example, when the net ideal gate operation is a bit flip,

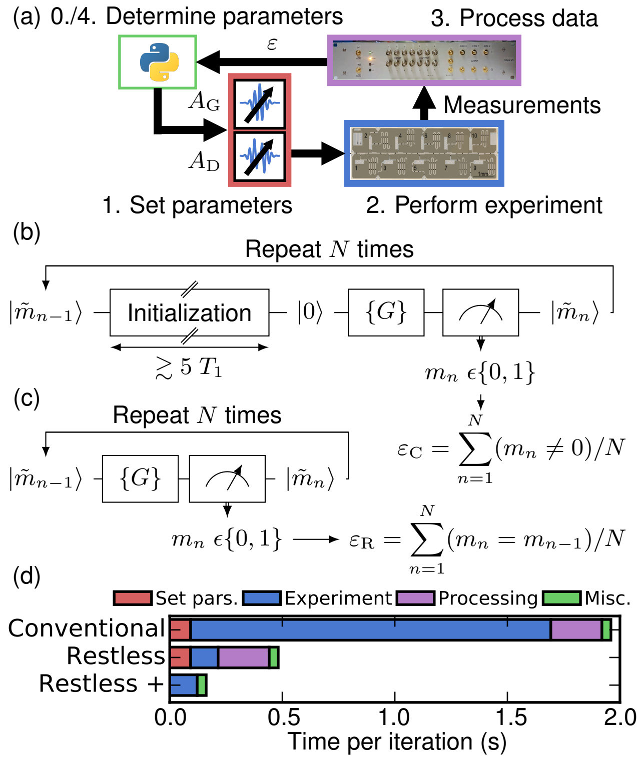


FIG. 1. (a) A general qubit gate tuneup loop. In conventional tuneup (b), the qubit is initialized before measuring the effect of $\{G\}$. In restless tuneup (c), the qubit is not initialized, and instead $m_{n-1}$ is used to estimate the initial state ($|\tilde{m}_{n-1}\rangle$). (d) Benchmark of various contributions to the time per iteration in conventional and restless tuneup, without and with technical improvements (see text for details).

we can define the error fraction

$$\varepsilon_{\mathrm{R}} = \sum_{n=2}^{N}(m_n = m_{n-1})/N. \quad (1)$$

We demonstrate restless tuneup of DRAG pulses [20] on the transmon qubit recently reported in [12] (summary of device parameters in [21]). We choose DRAG pulses (duration $\tau_{\mathrm{p}} = 20$ ns) for their proven ability to reduce gate error and leakage [22, 23] with few-parameter analytic pulse shapes. These pulses consist of Gaussian (G) and derivative of Gaussian (D) envelopes of the in- and quadrature-phase components of a microwave drive at the transition frequency $f$ between qubit levels $|0\rangle$ and $|1\rangle$. These components are generated using four channels of an arbitrary waveform generator (AWG), frequency upconversion by sideband modulation of one microwave source, and two I-Q mixers. The G and D components are combined inside a vector switch matrix (VSM) [24] (details in [21]). A key advantage of this scheme using four channels is the ability to independently set the G and D amplitudes ($A_{\mathrm{G}}$ and $A_{\mathrm{D}}$, respectively), without uploading new waveforms to the AWG.

To measure the speedup obtained from the restless method, we must take the complete iteration into account. The traditional iteration of a tuneup routine involves: (1) setting parameters (4 channel amplitudes on a Tektronix 5014 AWG); (2) acquiring $N = 8000$ measurement outcomes; (3) sending the measurement outcomes to the computer and processing them; and (4) miscellaneous overhead that includes determining the parameters for the next iteration, as well as saving and plotting data. In Fig. 1(d), we visualize these costs for an example optimization experiment. We intentionally penalize the restless method by choosing a large number of gates ($\sim 550$). Even in these conditions, restless sequences reduce the acquisition time from 1.60 to 0.12 s. However, the improvement in total time per iteration (from 1.98 to 0.50 s) is modest due to 0.38 s of overhead.

We take two steps to reduce overhead. The 0.23 s required to send all measurement outcomes to the computer and then calculate the error fraction is reduced to $< 1$ ms by calculating the fraction in real time, using the same FPGA system that digitizes and processes the raw measurement signals into bit outcomes. The 0.09 s required to set the four channel amplitudes in the AWG is reduced to 1 ms by setting $A_{\mathrm{G}}$ and $A_{\mathrm{D}}$ in the VSM. With these two technical improvements, the remaining overhead is dominated by the miscellaneous contributions (40 ms). This reduces the total time per restless (conventional) iteration to 0.16 s (1.64 s).

A quantity of common interest in gate tuneup is the average Clifford fidelity $F_{\mathrm{Cl}}$, which is typically measured using CRB. In CRB, $\{G\}$ consists of sequences of $N_{\mathrm{Cl}}$ random Clifford gates, including a final recovery Clifford gate that makes the ideal net operation identity. Following [25], we compose the 24 single-qubit Clifford gates from the set of $\pi$ and $\pm\pi/2$ rotations around the $x$ and $y$ axes, which requires an average of 1.875 gates per Clifford. Gate errors make $\varepsilon_{\mathrm{C}}$ increase with $N_{\mathrm{Cl}}$ as [26, 27]

$$1 - \varepsilon_{\mathrm{C}} = A \cdot (p_{\mathrm{Cl}})^{N_{\mathrm{Cl}}} + B. \quad (2)$$

Here, $A$ and $B$ are constants determined by state preparation and measurement error (SPAM), and $1 - p_{\mathrm{Cl}}$ is the average depolarizing probability per gate, making $F_{\mathrm{Cl}} = \frac{1}{2} + \frac{1}{2}p_{\mathrm{Cl}}$. Extracting $F_{\mathrm{Cl}}$ from a CRB experiment involves measuring $\varepsilon_{\mathrm{C}}$ for different $N_{\mathrm{Cl}}$ and fitting Eq. (2). However, for tuning it is sufficient to optimize $\varepsilon_{\mathrm{C}}$ at one choice of $N_{\mathrm{Cl}}$, because $\varepsilon_{\mathrm{C}}(N_{\mathrm{Cl}})$ decreases monotonically with $F_{\mathrm{Cl}}$ [9].

In the presence of leakage, CRB sequences and $\varepsilon_{\mathrm{C}}$ are not ideally suited for restless tuneup. Typically, there is significant overlap in the readout signals from the first- ($|1\rangle$) and second- ($|2\rangle$) excited state of a transmon. A transmon in $|2\rangle$ can produce a string of identical measurement outcomes until it relaxes back to the qubit subspace. If the ideal net operation of $\{G\}$ is identity, the measurement outcomes can be indistinguishable from ideal behavior. Although the leakage on single-qubit gates is typically small ($10^{-5} - 10^{-3}$ per Clifford for the range of $A_{\mathrm{D}}$ considered [23, 24]), a simple modification to the sequence allows penalizing leakage. By choosing the recovery Clifford for restless randomized benchmark-

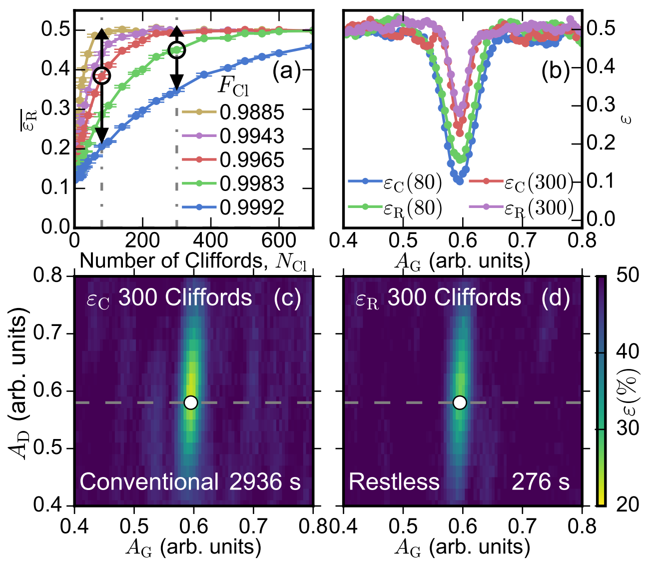


FIG. 2. (a) Average error fraction of RRB for different $F_{\rm Cl}$ vs $N_{\rm Cl}$. (b) $\varepsilon_{\rm C}$ and $\varepsilon_{\rm R}$ as a function of $A_{\rm G}$ for $N_{\rm Cl} = 80$ and $N_{\rm Cl} = 300$. The curves are denoted by a dashed line in (c-d). (c-d) $\varepsilon$ for $N_{\rm Cl} = 300$ as a function of $A_{\rm G}$ and $A_{\rm D}$. White circles indicate minimal $\varepsilon$. Total acquisition time is shown at the bottom right.

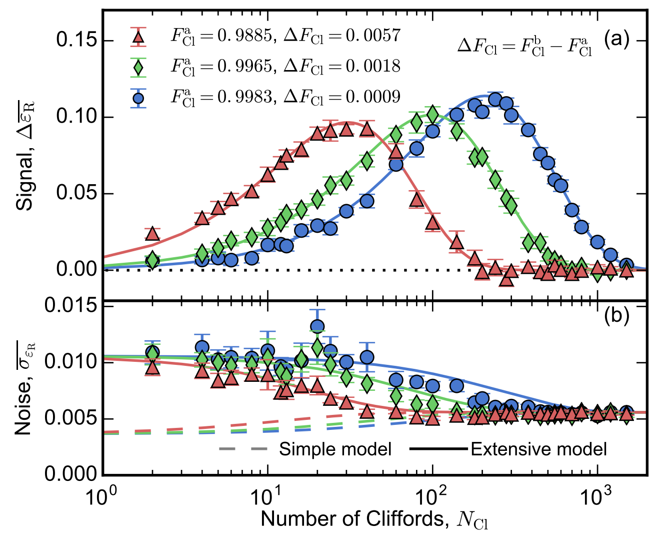


FIG. 3. (a) Signal $\Delta\overline{\varepsilon_{\rm R}}$ for a halving of the gate infidelity, plotted as a function $N_{\rm Cl}$ at $F^{\rm a}_{\rm Cl} \sim 0.989$ (red), 0.996 (green) and 0.998 (blue). (b) Noise dependence on $N_{\rm Cl}$ at the same fidelity levels. Added curves are obtained from the two models described in the main text.

ing (RRB) sequences so that the ideal net operation of $\{G\}$ is a bit flip, leakage produces an error. This simple modification makes $\varepsilon_{\rm R}$ a better cost function.

We now examine the suitability of the restless scheme for optimization (Fig. 2). Plots of the average $\varepsilon_{\rm R}(N_{\rm Cl})$ $[\overline{\varepsilon_{\rm R}}(N_{\rm Cl})]$ at various $F_{\rm Cl}$ (controlled via $A_{\rm G}$) behave similarly to $\varepsilon_{\rm C}$ in CRB. Furthermore, $\varepsilon_{\rm R}$ is minimized at the same $A_{\rm G}$ as $\varepsilon_{\rm C}$, with only a shallower dip because of SPAM. The ($A_{\rm G}$, $A_{\rm D}$) landscapes for both cost functions [Fig. 2(c-d)] are smooth around the optimum, making them suitable for numerical optimization. The fringes far from the optimum arise from the limited number of seeds (always 200) used to generate the RB sequences. Note that while the landscapes are visually similar, the difference in time required to map them is striking: $\sim 50$ min for $\varepsilon_{\rm C}$ versus $< 5$ min for $\varepsilon_{\rm R}$ at $N_{\rm Cl} = 300$.

The sensitivity of $\varepsilon_{\rm R}$ to the tuning parameters depends on both the gate fidelity and $N_{\rm Cl}$. This can be seen in the variations between curves in Fig. 2(a). In order to quantify this sensitivity, we define a signal-to-noise ratio (SNR). For signal we take the average change in the error fraction, $\Delta\overline{\varepsilon_{\rm R}} = \overline{\varepsilon_{\rm R}}(F^{\rm b}_{\rm Cl}) - \overline{\varepsilon_{\rm R}}(F^{\rm a}_{\rm Cl})$, from $F^{\rm a}_{\rm Cl}$ to $F^{\rm b}_{\rm Cl} \approx \frac{1}{2} + \frac{1}{2}F^{\rm a}_{\rm Cl}$ (halving the infidelity). For noise we take $\overline{\sigma_{\varepsilon_{\rm R}}}$, the average standard deviation of $\varepsilon_{\rm R}$ between $F^{\rm a}_{\rm Cl}$ and $F^{\rm b}_{\rm Cl}$. We find that the maximal SNR remains $\sim 15$ for an optimal choice of $N_{\rm Cl}$ that increases with $F^{\rm a}_{\rm Cl}$ (Fig. 3 and details in [21]). This allows tuning in logarithmic time, since reducing error rates $p \rightarrow p/2^M$ requires only $M$ optimization steps.

A simple model describes the measurement outcomes as independent and binomially distributed with error probability $\varepsilon_{\rm R}$, as per Eq. (2) with $\varepsilon_{\rm C} \rightarrow \varepsilon_{\rm R}$. This model captures all the essential features of the signal. However, it only quantitatively matches the noise at high $N_{\rm Cl}$. Experiment shows an increase in noise at low $N_{\rm Cl}$. In this range, $\varepsilon_{\rm R}$ is dominated by SPAM, which is primarily due to $T_1$. We surmise that the increase stems from $T_1$ fluctuations [28] during the acquisition of statistics in these RRB experiments. To test this hypothesis, we develop an extensive model incorporating $T_1$ fluctuations into the calculation of both signal and noise [21]. We find good agreement with experimental results using independently measured values of $\overline{T_1}$ and $\sigma_{T_1}$. The good agreement confirms the non-demolition character of the measurement previously reported in [12].

Following its validation, we now employ $\varepsilon_{\rm R}$ in a two-step numerical optimization protocol (Fig. 4). We choose the Nelder-Mead algorithm [29] as it is derivative-free and easy to use, requiring only the specification of a starting point and initial step sizes. The first step using $\varepsilon_{\rm R}(N_{\rm Cl} = 80)$ ensures convergence even when starting relatively far from the optimum, while the second step using $\varepsilon_{\rm R}(N_{\rm Cl} = 300)$ fine tunes the result. We test the optimization for four realistic starting deviations from the optimal parameters $(A_{\rm D}^{\rm opt}, A_{\rm G}^{\rm opt})$. $A_{\rm G}$ is chosen at both approximately 6% above and below $A_{\rm G}^{\rm opt}$, selected as a worst-case estimate from a Rabi oscillation experiment. $A_{\rm D}$ is chosen at both approximately half and double $A_{\rm D}^{\rm opt}$. The initial step sizes are $\Delta A_{\rm G} \approx -0.03 A_{\rm G}^{\rm opt}$, $\Delta A_{\rm D} \approx -0.25 A_{\rm D}^{\rm opt}$ for the first step, and $\Delta A_{\rm G} \approx -0.01 A_{\rm G}^{\rm opt}$, $\Delta A_{\rm D} \approx -0.08 A_{\rm D}^{\rm opt}$ for the second step.

We assess the accuracy of the above optimization and compare to traditional methods. A CRB experiment

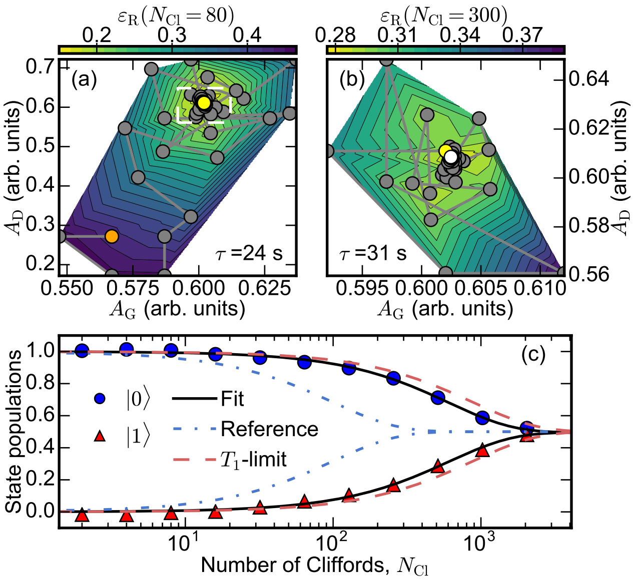


FIG. 4. Two-parameter restless tuneup using a two-step optimization, first at $N_{Cl} = 80$ (a) and then at $N_{Cl} = 300$ (b). Contour plots show a linear interpolation of $\varepsilon_R$. The starting point, intermediate result and final result are marked by orange, yellow, and white dots respectively. (c) CRB of tuned pulses ($F_{Cl} = 0.9991$), compared to $F_{Cl}^{(T_1)} = 0.9994$ and $F_{Cl} = 0.995$ for reference.

| | 2-par. ($A_G$, $A_D$) | | 3-par. ($A_G$, $A_D$, $f$) | |
|---|---|---|---|---|
| | conv. | restl. | conv. | restl. |
| $\overline{F_{Cl}}$ | 0.9991 | 0.9991 | 0.9990 | 0.9990 |
| $\sigma_{F_{Cl}}$ | $3 \cdot 10^{-5}$ | $3 \cdot 10^{-5}$ | 0.0001 | 0.0001 |
| $\overline{\tau}$ | 660 s | 59 s | 610 s | 66 s |
| $\sigma_\tau$ | 110 s | 11 s | 110 s | 13 s |
| $\overline{N_{it}}$ | 400 | 370 | 370 | 420 |
| $\sigma_{N_{it}}$ | 70 | 70 | 70 | 80 |
| $\overline{F_{Cl}^{(T_1)}}$ | 0.9994 | | 0.9993 | |
| $\overline{T_1}$ | 21.4 $\mu$s | | 19.3 $\mu$s | |

TABLE I. Tuning protocol performance. Mean (overlined) and standard deviations (denoted by $\sigma$) of $F_{Cl}$, time to convergence $\tau$, and number of iterations $N_{it}$ for restless and conventional tuneups with 2 and 3 parameters. Average $T_1$ measured throughout these runs and corresponding average $F_{Cl}^{(T_1)}$ are also listed.

[Fig. 4(c)] following two-parameter restless optimization indicates $F_{Cl} = 0.9991$. This value matches the average achieved by both restless and conventional tuneups for the different starting conditions. We also implement GST to independently verify results obtained using CRB. From the process matrices we extract the average GST Clifford fidelity, $F_{Cl}^{GST} = 0.99907 \pm 0.00003$ ($0.99909 \pm 0.00003$) for restless (conventional) tuneup [21], consistent with the value obtained from CRB.

The robustness of the optimization protocol is tested by interleaving tuneups with CRB and $T_1$ measurements over 11 hours (summarized in Table I, and detailed in [21]). Both tuneups reliably converge to $F_{Cl} = 0.9991$, close to the $T_1$ limit [30]:

$$F_{Cl}^{(T_1)} \approx \frac{1}{6}\left(3 + 2e^{-\tau_c/2T_1} + e^{-\tau_c/T_1}\right) = 0.9994, \qquad (3)$$

with $\tau_c = 1.875\ \tau_p$. However, restless tuneup converges in one minute, while conventional tuneup requires eleven.

It remains to test how restless tuneup behaves as additional parameters are introduced. Many realistic scenarios also require tuning the drive frequency $f$. As a worst case, we take an initial detuning of $\pm 250$ kHz. The initial step size in the first (second) step is 100 kHz (50 kHz). The 3-parameter optimization converges to $F_{Cl} = 0.9990 \pm 0.0001$ for both restless and conventional tuneups. We attribute the slight decrease in $F_{Cl}$ achieved by 3-parameter optimization to the observed reduction in average $T_1$.

In summary, we have developed an accurate and robust tuneup method achieving a tenfold speedup over the state of the art [9]. This speedup is achieved by avoiding qubit initialization by relaxation, and by using real-time correlation of measurement outcomes to build the cost function for numerical optimization. We have applied the restless concept to the tuneup of Clifford gates on a transmon qubit, reaching a $T_1$-dominated fidelity of 0.999 in one minute, verified by conventional randomized benchmarking and gate set tomography. We have shown experimentally that the method can detect fractional reductions in gate error with nearly constant signal-to-noise ratio. An interesting next direction is to develop an algorithm that makes optimal use of this tunable sensitivity while maintaining the demonstrated robustness. The enhanced speed combined with the generic nature of the optimizer would also allow exploring other, more generic non-adiabatic gates without analytic pulse shapes, in a fashion analogous to optimal control theory [31, 32]. Immediate next experiments will extend the restless concept to the tuneup of two-qubit controlled-phase gates [33, 34] exploiting interactions with non-computational states [35], in which leakage errors often dominate ($\sim 10^{-2}$). In this context, we anticipate that the RRB modification and the $\varepsilon_R$ cost function will prove essential to reach 0.999 fidelity. Finally, we also envision applying the restless concept to the simultaneous tuneup of single-qubit gates in the many-qubit setting (e.g, a logical qubit).

## ACKNOWLEDGMENTS

We thank R. Sagastizabal for experimental assistance, C. Dickel, J. Helsen, and S. Poletto for discussions, A. Johnson for support with Microsoft QCoDeS and K. Ruddinger, E. Nielsen and R. Blume-Kohout for support with GST/pyGSTi. This research is supported by the Office of the Director of National Intelligence

(ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the U.S. Army Research Office grant W911NF-16-1-0071. Additional funding provided by the ERC Synergy Grant QC-lab, the China Scholarship Council (X.F.) and Microsoft Corporation Station Q. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

[1] 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," arXiv:1608.00263 (2016).

[2] Pierre-Luc Dallaire-Demers and Frank K. Wilhelm, "Quantum gates and architecture for the quantum simulation of the fermi-hubbard model," Phys. Rev. A **94**, 062304 (2016).

[3] Clare Horsman, Austin G Fowler, Simon Devitt, and Rodney Van Meter, "Surface code quantum computing by lattice surgery," New J. Phys. **14**, 123011 (2012).

[4] Yu Tomita and Krysta M. Svore, "Low-distance surface codes under realistic quantum noise," Phys. Rev. A **90**, 062320 (2014).

[5] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, "Surface codes: Towards practical large-scale quantum computation," Phys. Rev. A **86**, 032324 (2012).

[6] John M. Martinis, "Qubit metrology for building a fault-tolerant quantum computer," npj Quantum Inf. **1**, 15005 (2015).

[7] B. E. Anderson, H. Sosa-Martinez, C. A. Riofrío, Ivan H. Deutsch, and Poul S. Jessen, "Accurate and robust unitary transformations of a high-dimensional quantum system," Phys. Rev. Lett. **114**, 240401 (2015).

[8] D. J. Egger and F. K. Wilhelm, "Adaptive hybrid optimal quantum control for imprecisely characterized systems," Phys. Rev. Lett. **112**, 240503 (2014).

[9] J Kelly, R Barends, B Campbell, Y. Chen, Z Chen, B Chiaro, A Dunsworth, A. G. Fowler, I.-C. Hoi, E Jeffrey, A Megrant, J Mutus, C Neill, P. J. J. OMalley, C Quintana, P Roushan, D Sank, A Vainsencher, J Wenner, T. C. White, A. N. Cleland, and John M Martinis, "Optimal quantum control using randomized benchmarking," Phys. Rev. Lett. **112**, 240504 (2014).

[10] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Yu Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Lucero, M. Neeley, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, and John M. Martinis, "Scalable *in situ* qubit calibration during repetitive error detection," Phys. Rev. A **94**, 032321 (2016).

[11] D. T. McClure, Hanhee Paik, L. S. Bishop, M. Steffen, Jerry M. Chow, and Jay M. Gambetta, "Rapid driven reset of a qubit readout resonator," Phys. Rev. Appl. **5**, 011001 (2016).

[12] C. C. Bultink, M. A. Rol, T. E. O'Brien, X. Fu, B. C. S. Dikken, C. Dickel, R. F. L. Vermeulen, J. C. de Sterke, A. Bruno, R. N. Schouten, and L. DiCarlo, "Active resonator reset in the nonlinear dispersive regime of circuit QED," Phys. Rev. Appl. **6**, 034008 (2016).

[13] Pascal Cerfontaine, Tim Botzem, Simon Sebastian Humpohl, Dieter Schuh, Dominique Bougeard, and Hendrik Bluhm, "Feedback-tuned noise-resilient gates for encoded spin qubits," arXiv:1606.01897 (2016).

[14] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, "Feedback control of a solid-state qubit using high-fidelity projective measurement," Phys. Rev. Lett. **109**, 240502 (2012).

[15] M. H. Devoret and R. J. Schoelkopf, "Superconducting circuits for quantum information: An outlook," Science **339**, 1169–1174 (2013).

[16] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf, "Surface participation and dielectric loss in superconducting qubits," Appl. Phys. Lett. **107**, 162601 (2015).

[17] D. Ristè and L. DiCarlo, "Digital feedback in superconducting quantum circuits," ArXiv:1508.01385 (2015).

[18] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Kenneth Rudinger, Jonathan Mizrahi, Kevin Fortier, and Peter Maunz, "Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography," Nat. Commun. (2017).

[19] Except 3.25 $\mu$s needed for passive depletion of photons leftover from the 1 $\mu$s measurement [12].

[20] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, "Simple pulses for elimination of leakage in weakly nonlinear qubits," Phys. Rev. Lett. **103**, 110501 (2009).

[21] See supplemental material at [insert URL] for additional data.

[22] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, "Optimized driving of superconducting artificial atoms for improved single-qubit gates," Phys. Rev. A **82**, 040305 (2010).

[23] Zijun Chen, Julian Kelly, Chris Quintana, R. Barends, B. Campbell, Yu Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Lucero, E. Jeffrey, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Korotkov, and John M. Martinis, "Measuring and suppressing quantum state leakage in a superconducting qubit," Phys. Rev. Lett. **116**, 020501 (2016).

[24] S. Asaad, C. Dickel, S. Poletto, A. Bruno, N. K. Langford, M. A. Rol, D. Deurloo, and L. DiCarlo, "Independent, extensible control of same-frequency superconducting qubits by selective broadcasting," npj Quantum Inf. **2**, 16029 (2016).

[25] Jeffrey M. Epstein, Andrew W. Cross, Easwar Magesan, and Jay M. Gambetta, "Investigating the limits of randomized benchmarking protocols," Phys. Rev. A **89**, 062321 (2014).

[26] Easwar Magesan, J. M. Gambetta, and Joseph Emerson, "Scalable and robust randomized benchmarking of quantum processes," Phys. Rev. Lett. **106**, 180504 (2011).

[27] Easwar Magesan, Jay M. Gambetta, and Joseph Emerson, "Characterizing quantum gates via randomized benchmarking," Phys. Rev. A **85**, 042311 (2012).

[28] Clemens Müller, Jürgen Lisenfeld, Alexander Shnirman, and Stefano Poletto, "Interacting two-level defects as sources of fluctuating high-frequency noise in superconducting circuits," Phys. Rev. B **92**, 035442 (2015).

[29] J. A. Nelder and R. Mead, "A simplex method for function minimization," The Computer Journal **7**, 308–313 (1965).

[30] E. Magesan, private communication.

[31] Steffen J. Glaser, Ugo Boscain, Tommaso Calarco, Christiane P. Koch, Walter Köckenberger, Ronnie Kosloff, Ilya Kuprov, Burkhard Luy, Sophie Schirmer, Thomas Schulte-Herbrüggen, Dominique Sugny, and Frank K. Wilhelm, "Training schrödinger's cat: quantum optimal control," The European Physical Journal D **69**, 279 (2015).

[32] Shai Machnes, David J Tannor, Frank K Wilhelm, and Elie Assémat, "Gradient optimization of analytic controls: the route to high accuracy quantum optimal control," ArXiv:1507.04261 (2015).

[33] L. DiCarlo, M. D. Reed, L. Sun, B. R. Johnson, J. M. Chow, J. M. Gambetta, L. Frunzio, S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf, "Preparation and measurement of three-qubit entanglement in a superconducting circuit," Nature **467**, 574 (2010).

[34] 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, "Superconducting quantum circuits at the surface code threshold for fault tolerance." Nature **508**, 500 (2014).

[35] Frederick W. Strauch, Philip R. Johnson, Alex J. Dragt, C. J. Lobb, J. R. Anderson, and F. C. Wellstood, "Quantum logic gates for coupled superconducting phase qubits," Phys. Rev. Lett. **91**, 167005 (2003).

[36] M. A. Rol, C. Dickel, S. Asaad, C. C. Bultink, R. Sagastizabal, N. K. L. Langford, G. de Lange, B. C. S. Dikken, X. Fu, S. R. de Jong, and F. Luthi, "PycQED," (2016).

[37] A.C. Johnson, G. Ungaretti, *et al.*, "QCoDeS," (2016).

[38] T.E. O'Brien, B.Tarasinski, C.C.Bultink, M.A.Rol, R.Versluis, S.Poletto, and L. DiCarlo, "IARPA logiQ M6 report, QuSurf logical qubit and its simulated performance with current physical-qubit parameters, not publicly available," (2016).

[39] Erik Nielsen, Travis Scholten, Kenneth Rudinger, and Jonathan Gross, "pyGSTi: Version 0.9.1 beta," (2016) .