



# CHORUS

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

## Robust and efficient in situ quantum control

Christopher Ferrie and Osama Moussa

Phys. Rev. A **91**, 052306 — Published 7 May 2015

DOI: [10.1103/PhysRevA.91.052306](https://doi.org/10.1103/PhysRevA.91.052306)

# Robust and efficient *in situ* quantum control

Christopher Ferrie<sup>1,2,\*</sup> and Osama Moussa<sup>3,†</sup>

<sup>1</sup>*Center for Quantum Information and Control, University of New Mexico, Albuquerque, New Mexico, 87131-0001*

<sup>2</sup>*Centre for Engineered Quantum Systems, School of Physics,  
The University of Sydney, Sydney, NSW, Australia*

<sup>3</sup>*Institute for Quantum Computing and Department of Physics and Astronomy,  
University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

Precision control of quantum systems is the driving force for both quantum technology and the probing of physics at the quantum and nano-scale. We propose an implementation independent method for *in situ* quantum control that leverages recent advances in the direct estimation of quantum gate fidelity. Our algorithm takes account of the stochasticity of the problem and is suitable for closed-loop control and requires only a constant number of fidelity estimating experiments per iteration independent of the dimension of the control space. It is efficient and robust to both statistical and technical noise.

## I. INTRODUCTION

Precision quantum control enables practical goals such as quantum computation [1] and quantum metrology [2] which in turn can provide probes of fundamental physics such as gravity wave detection [3]. To process the information in a quantum device requires enacting quantum logic gates with high fidelity. State of the art methods calculate the solution to Schrodinger’s equation in iterative gradient climbing algorithms [4] to arrive at a control pulse that has high fidelity to some target evolution with respect to some physical model. Contrary to this approach, which we call *ex situ* control, it would be convenient if the quantum device guided itself to a desired state—*in situ* quantum control.

Control of the *in situ* type can be divided into two categories: those that use a fresh copy of the system with each measurement and those that use a single copy in a continuous measurement scenario. The former was pioneered by Rabitz *et al* in the context of state transfer using genetic algorithms [5], and is the type of *in situ* control we consider. The latter is often called *quantum feedback control* and the interested reader is referred to [6]. Both are also referred to as either *feedback* or *closed-loop* control [7].

Until recently, however, it was not known whether the fidelity to some target gate could be estimated efficiently from experiment. Here we leverage recent advances in fidelity estimation [8–12] to design an *in situ* quantum control algorithm which can in principle efficiently and robustly find the optimal control sequence. In particular, we show via numerical experiments that our *in situ* algorithm converges in fidelity to the target unitary gate at rate given by  $O(1/N_{\text{tot}})$  where  $N_{\text{tot}}$  is the total number of experiments performed.

In very broad strokes, the current paradigm for quantum control proceeds in four steps: (i) modeling, (ii) estimation, (iii) optimization and (iv) implementation and

verification. At each of these steps, many things can go wrong. We briefly overview each step, recall what could go wrong and describe how *in situ* control remedies the problem.

The first task is to model the system; that is, to create a physical model of the dynamics experienced by the system. Much can be learned from our current understanding of the physics involved. However, as physics is the art of approximation, the model will never be precisely correct. It can be very good, but the precision in control will be limited by the precision in the model [13]. Alternatively, if the system itself is used, then the “model” used to design controls is perfect—it is an exact replica of the physics because it is the physics! Thus, there are no limitations due to modeling errors for *in situ* control. This is illustrated in Fig. 1.

Given a model, the next task is to estimate the parameters of the model. There are many ingenious ways to do this (from full process tomography [14] to Hamiltonian learning using quantum resources [15]). Although estimation of physical parameters is limited in precision due to statistical and technical errors, control algorithms can be made “robust” to these imperfections. But there is no free lunch here either: a control strategy that is robust to a range of parameters is typically more taxing to find, and will not be strictly optimal for the “true” parameters [16]. For *in situ* control, we have noted already that a model is not required. Therefore, the task of estimating the parameters of the model is not required and any problems incurred by this step are nonexistent.

Given a model and set of parameters, the task now is to find the optimal set of controls. This is usually done through some iterative optimization algorithm performed on a classical computer [4]. For each candidate set of controls, a classical simulation is required. It is thus inefficient in general to solve such an optimization problem. Moreover, it is doubly inefficient since a classical simulation is used to propagate the state vector only to compute a single number. While *in situ* control trivially sidestepped the previous problems, here is where it provides a technical advantage. Since the system itself will perform its own simulation, it may seem efficient. How-

\* csferrie@gmail.com

† omoussa@gmail.com

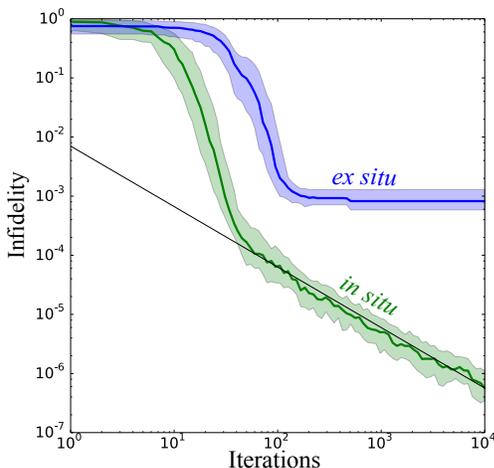


FIG. 1. (Color online) (Color online) The median *infidelity* (1-fidelity) of our *in situ* algorithm and an *ex situ* algorithm as a function of the number of iterations. The problem is that of designing a random single qubit gate and is described in more detail in Section III. The *in situ* algorithm was afforded only  $10^3$  experiments per fidelity estimate while the *ex situ* algorithm was given the ability to compute the fidelity to machine precision but saturates since the assumed Hamiltonian was incorrect by  $\|\Delta H\| = 10^{-2}\|H\|$ .

ever, the key difference between classical simulation (as used in control design) and quantum simulation is that a quantum simulator can only produce sample outcomes. Thus, it may be the case that an inefficient number of experiments need to be run at each step in the optimization protocol. The main result of this paper is to show this is not the case. That is, our *in situ* control strategy is efficient.

Finally, given the optimized set of controls, the final task is to implement those controls in an experiment and verify that the intended gate was produced. High precision experiments are an impressive feat of physics and engineering—there are many imperfections in machined or fabricated components, stray fields and vibrations, noisy electronics and so on that need to be overcome. Thus, it is not often the case that the optimal controls will reach the same value of the objective function in reality that they did in simulation. For *in situ* control, once the optimal control has been found, we are done. That is, the last experiment performed in our iterative optimization algorithm is the one sought after and verifies its own performance.

The imperfections due to the classical control devices are of systematic or stochastic origins. The systematic deviations (resulting from e.g. nonlinear or non-uniform response of amplifiers or resonators) can in principle be reversed/accounted for with pre-distortions of the optimized sequences, or using classical feedback, or even by inclusion in the modeling or optimization steps. The random distortions, or the *technical noise*, plagues the implementation and verification step for *ex situ* schemes, and countering this requires more data and averaging, which

scales quadratically with the desired precision. The technical noise challenges *in situ* control in a different way—if the decision making of the optimization method is rigid and based solely on a deterministic view of the fidelity landscape, it is not hard to imagine how it would be limited by randomness in the fidelity evaluations. Thus, it is important for our scheme to allow for the stochasticity of the fidelity estimation.

Although in the standard paradigm the challenges in modeling, estimation and implementation are in principle avoidable in a perfect world, those in actual control finding are not. That is, in general, classical optimization of the controls is inefficient. For our scheme to provide a solution to this problem, we need to show that it is efficient. It is illustrative to compare to two recent similar protocols, which use both classical and *in situ* control: Ad-HOC [17] and ORBIT [18]. Both protocols rely on the Nelder-Mead (NM) optimization algorithm. Since these have such sexy names, we need to give our algorithm a name—we call it Adaptive Control via Randomized Optimization Nearly Yielding Maximization (ACRONYM).

ACRONYM can be divided into two distinct pieces: fidelity estimation and stochastic optimization. Fidelity estimation can be achieved by various means. For certain classes of gates, efficient algorithms exist [9–12] and have already been experimentally implemented [11, 19–22]. However, any fidelity estimation scheme requires many repeated experiments to reduce the noise due to statistical fluctuations.

In Ad-HOC, the noise is modeled by a depolarizing channel, and in ORBIT, the noise is some combination of technical experimental noise and statistical noise due to finite sampling. The latter provides an ultimate lower bound on the achievable accuracy in Ad-HOC and ORBIT, whereas ACRONYM is not limited by statistical noise. The reason, as observed in [17], is due to the sensitivity of the NM algorithm to fluctuations in fidelity. After describing ACRONYM, we demonstrate that it is efficient in this sense of requiring only a fixed number of experiments per iteration of the algorithm. Broadly speaking, ACRONYM is implementation independent, requiring only that changes in the control space produce changes in the fidelity which can be estimated via experiment.

The outline of the paper is as follows. In Section II we introduce the problem and our control algorithm ACRONYM. Section III reports on numerical experiments benchmarking the performance of the algorithm. Section IV contains some further discussion and concludes the paper.

## II. ACRONYM: ADAPTIVE CONTROL VIA RANDOMIZED OPTIMIZATION NEARLY YIELDING MAXIMIZATION

### A. Problem statement

Generally, the task is to select a set of controls  $\mathbf{c} \in \mathbb{R}^p$ , where  $p$  is the dimension of the control space, such that the implemented channel  $\Lambda_{\mathbf{c}}$  is as close as possible to some target  $U_{\text{T}}$ , assumed to be a unitary. The “closeness” is measured by the channel fidelity  $f(\mathbf{c}) := F(\Lambda_{\mathbf{c}}, U_{\text{T}})$ , where we have defined the object function  $f$  to maximize. When the implemented gate is also unitary—the case we consider—we have

$$F(U_{\mathbf{c}}, U_{\text{T}}) = \frac{1}{d^2} |\text{Tr}(U_{\text{T}}^\dagger U_{\mathbf{c}})|^2, \quad (1)$$

where  $d$  is the dimension of the quantum system.

A classical simulation can provide an exact numerical calculation of  $f$ , whereas a quantum simulation only provides a single datum—many from which we can estimate  $f$ . Happily, there exist efficient protocols to estimate  $f$ —for example, randomized benchmarking [12], direct fidelity estimation via Monte Carlo [9, 10], or certification via twirling [11]. Let us suppose that we are in some regime where we can consider a good estimate of the fidelity to be

$$\hat{f}(\mathbf{c}) = \mathbb{E}[f(\mathbf{c})], \quad (2)$$

where the expectation value is taken with respect to the distribution of data. Naturally, due to finite sample statistics, this function is *stochastic* and thus the optimization term “function call” takes on new meaning—the same set of controls  $\mathbf{c}$  might result in different function evaluations  $\hat{f}$ . Thus, the problem of maximizing the fidelity becomes one of *stochastic optimization* [23].

### B. *in situ* control algorithm

A particularly useful set of techniques for multidimensional analysis goes by the name *simultaneous perturbation stochastic approximation* (SPSA) [24]. In short, SPSA is an iterative optimization technique which uses only two (noisy) function calls per iteration to estimate the gradient. Our algorithm, ACRONYM, is a direct quantum analog of SPSA. The steps of each iteration are outlined as follows. First fix a tolerance  $\epsilon > 0$ . For each  $k$ , so long as  $\|\mathbf{c}_{k+1} - \mathbf{c}_k\| > \epsilon$ , repeat:

1. Generate a random direction to search in defined by the vector  $\Delta_k$ . A recommended vector which we choose is where each element is selected according to a fair coin toss:  $\Delta_{kj} = \pm 1$ .

2. Calculate the estimated gradient

$$\mathbf{g}_k = \frac{\hat{f}(\mathbf{c}_k + \beta_k \Delta_k) - \hat{f}(\mathbf{c}_k - \beta_k \Delta_k)}{2\beta_k} \Delta_k. \quad (3)$$

3. Calculate the next iterate via

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \alpha_k \mathbf{g}_k. \quad (4)$$

The functions  $\alpha_k$  and  $\beta_k$  control the convergence and are user defined, although they are usually specified in the forms

$$\alpha_k = \frac{a}{(k+1)^s}, \quad \beta_k = \frac{b}{(k+1)^t}, \quad (5)$$

where  $a, b, s$  and  $t$  are chosen first roughly based on extensive numerical studies for many problems then tweaked based on numerical simulations for the problem at hand. For reference, generally good choices are [25]  $a = b = 1$ ,  $s = 0.602$  and  $t = 0.101$ , but we have found that the asymptotically optimal values [24]  $s = 1$  and  $t = 1/6$  give good results across all parameter regimes considered here.

### C. Convergence discussion

There are a number of convergence results on stochastic optimization and the variant we use [24, 25]. All conclude that the error in the design space decreases at rate  $O(k^{\beta_c})$  where typically  $\beta_c \in [-1/2, 0]$ . Note that the lower bound comes from standard statistical arguments. The actual performance achieved depends on a number of often competing factors, thus we will allow the data to decide and compare to what might be expected from asymptotic arguments.

Here, we will be interested in the convergence in the objective function—the fidelity—rather than the controls. First we argue that  $f$  will decrease as  $O(k^{\beta_f})$  for  $\beta_f = 2\beta_c$ . Assuming  $f$  is differentiable and obtains its minimum at  $\mathbf{c}_{\text{opt}}$ , the gradient  $\nabla f(\mathbf{c}_{\text{opt}}) = 0$  and we obtain the bound

$$|f(\mathbf{c}_k) - f(\mathbf{c}_{\text{opt}})| \leq \frac{K}{2} \|H[f](\mathbf{c}_{\text{opt}})\| \|\mathbf{c}_k - \mathbf{c}_{\text{opt}}\|_2^2, \quad (6)$$

$$\frac{K}{2} \|H[f](\mathbf{c}_{\text{opt}})\| k^{2\beta_c}, \quad (7)$$

for some constant  $K$  and where  $H[f](\mathbf{c}_{\text{opt}})$  is the Hessian of second derivatives, evaluated at  $\mathbf{c}_{\text{opt}}$  and  $\|H[f](\mathbf{c}_{\text{opt}})\|$  is the spectral norm, the largest eigenvalue, of  $H[f](\mathbf{c}_{\text{opt}})$ . Thus, we should still expect  $O(k^{\beta_f})$  convergence in fidelity, with optimality given by  $\beta_f \approx -1$ . In the examples we consider, we will extract the exponents from fits to the simulation data. Since we will not generally consider  $\beta_c$ , we drop the subscript  $f$  on  $\beta_f$  from now on.

In discussing the performance of the algorithms considered here, we will refer to three numbers:  $N$ , the number of experiments used for each fidelity estimate;  $M$ , the number of fidelity estimates required for each iteration of the algorithm; and  $k$ , the number of iterations. Thus, the total number of physical experiments required after  $k$  iterations is  $N_{\text{tot}} = N \cdot M \cdot k$ . For ACRONYM,  $M = 2$  regardless of the dimension of the control space or quantum

system. Moreover, we will show that the performance is roughly independent of  $N$ . Thus we can limit our attention to the performance as a function of  $k$ . For NM, the situation is more complicated since the performance depends crucially on  $N$ , and  $M$  randomly fluctuates. In all cases considered below, however, ACRONYM outperforms NM by any metric.

#### D. A tale of two fidelities

We acknowledge the apparent contradiction in claiming that the fidelity converges when it is only known up to statistical fluctuations of order  $1/\sqrt{N}$ . However, our claim is that the *underlying* fidelity (1) achieved by ACRONYM converges and not the *estimated* fidelity (2)—the estimated fidelity from noisy observation does not converge. The latter situation is simply a consequence of finite statistics *per iteration*.

Though both fidelities are important in their own right, the problem of control *is* to maximize the underlying fidelity, rather than to verify it. As a consequence, the fidelity of the currently selected control may be much closer to the optimal fidelity than can be verified by experiments with the same noise sources. In an experiment, the experimenter must trust that the algorithm converges or verify the controls found at the final iteration with more resources than used in their finding. In settings where quantum control is expected to be an automated subroutine, verifying the solution would not be expected in production.

### III. NUMERICAL EXPERIMENTS

#### A. Random single-qubit gates

The simplest demonstration is that of a qubit subjected to a Hamiltonian, in some frame,

$$H(t) = \omega_0 Z + \omega_c(t) X, \quad (8)$$

where  $Z$  and  $X$  are the single qubit Pauli operators,  $\omega_0 Z$  is the fixed drift Hamiltonian, and  $\omega_c(t) X$  is the control. The task is to choose  $\omega_c$  such that

$$U_c = \exp(-iH(t)) = U_T. \quad (9)$$

Even for this simple problem, there is no closed-form solution, which is why we consider problem of fidelity optimization. For this example, we will average over targets randomly generated from the Haar measure. We work in dimensionless units where  $\omega_0 = 1$  and the control is piecewise constant [26] on the time intervals between  $\{0, 1, 2, \dots, p\}$  so that  $\mathbf{c} = (\omega_c(1), \omega_c(2), \dots, \omega_c(p)) \in \mathbb{R}^p$ . We assume the fidelity is estimated via finite sample statistics such that the objective function in Eq. (2) is distributed according to  $N\hat{f}(\mathbf{c}) \sim \text{Bin}(N, f(\mathbf{c}))$ , a Binomial distribution, where  $N$  is the number of experiments used

to estimate the fidelity and the Binomial parameter is the true fidelity of the controls  $\mathbf{c}$ . This model for fidelity estimation was chosen as it conveniently and intuitively demonstrates the key important limitation of using non-stochastic optimization techniques: that fidelity must be estimated from experimental samples. The actual model will vary depending on the details of the experimental implementation, but what remains is that fidelity estimation is limited by statistical fluctuations—and here we will explore the ultimate shot-noise limits.

First, we demonstrate the utility of *in situ* control in general before comparing different *in situ* strategies. Suppose in an *ex situ* algorithm (such as [4]), the model above is not quite correct. For example, suppose that the drift Hamiltonian is incorrectly modeled such that  $H_{\text{drift}} = Z + \Delta H$ , for some random perturbation  $\Delta H$ . In particular, suppose  $\|\Delta H\| = 0.01$ . Then the offline optimization will rapidly design a control scheme for this model to hit the target  $U_T$ . However, these controls will implement a different unitary under the dynamics of the true model. On the other hand, ACRONYM uses the true model by fiat, but, as any *in situ* protocol, suffers a penalty for not being able to exactly compute the objective function—it must do so through experimental trials. In Fig. 1, we show the performance of an offline algorithm with a bad model and ACRONYM. As expected, the offline algorithm does quite well until the errors in the model dominate. On the other hand, ACRONYM continues to learn the optimal controls. Next, we will determine the asymptotic rate of learning and compare ACRONYM to the NM algorithm.

In the following, then, we report on numerical experiments which demonstrate that ACRONYM converges in fidelity to the target at rate  $O(k^\beta)$  given by  $\beta \approx -1$ . Since we show the asymptotic performance is independent of the the number of experiments per iteration, the overall performance appears to scale as  $O(1/N_{\text{tot}})$ . The simulations and our optimization algorithm, as well as comparisons [27] to the NM algorithm, were implemented in Python using SciPy [28]. The results for the single qubit problem are summarized in Fig. 2. As expected, the more experimental samples taken per iteration, the less the statistical noise, and the higher the fidelity given a fixed number of iterations for NM. We also see that NM optimization ceases to improve when the statistical noise dominates its ability to guess the correct search direction.

For all qubit simulations the control space dimension is  $p = 10$  which would naively suggest that  $M = 20$  fidelity evaluations are required to estimate the gradient. We will discuss the NM algorithm further in the discussion of Section IV, but here we note that for the simulations presented in Fig. 2 the average number of function calls per iteration used by NM was  $M = 3.13$ , 50% more than ACRONYM which, recall, requires exactly two fidelity estimates,  $M = 2$ . Moreover, ACRONYM's performance is roughly independent of the number of experiments per iteration, as shown in Fig. 3. We see that after  $N \approx 10^4$  experiments per fidelity estimate, there is no additional

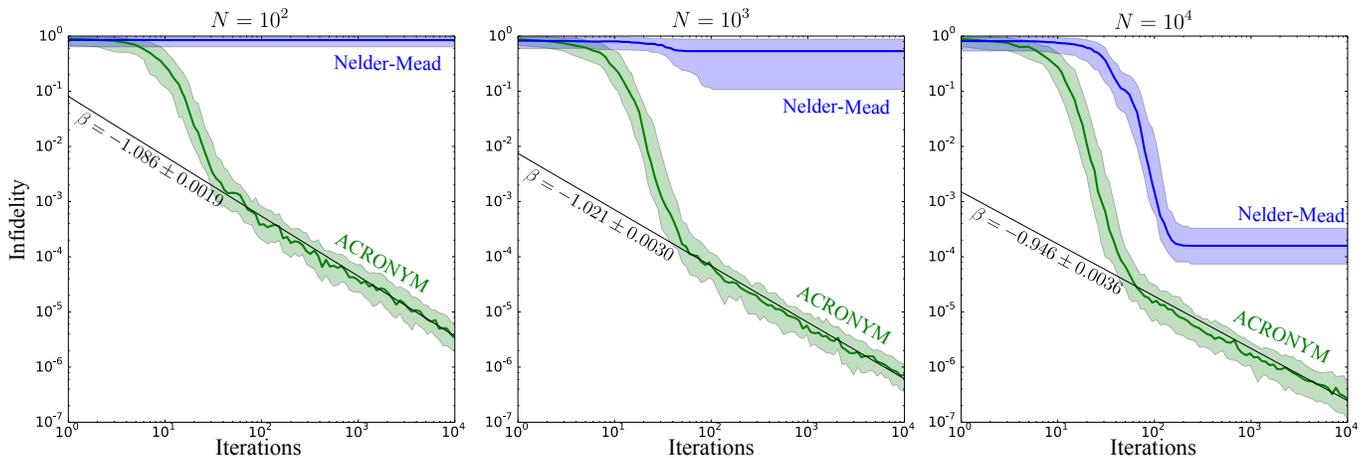


FIG. 2. (Color online) For the single qubit random gate design problem, plotted is the *infidelity* (1-fidelity) of the ACRONYM and NM algorithms as a function of the number of iterations for three different numbers of experiments per iteration: (Left to right)  $10^2$ ,  $10^3$  and  $10^4$ . The thick lines are the median of the data and the shaded region around each is the interquartile range (the middle 50% of the data). The thin black lines are fits to the theoretical  $O(k^\beta)$  scaling. We see that NM is limited by the noise in the function call while ACRONYM is not.

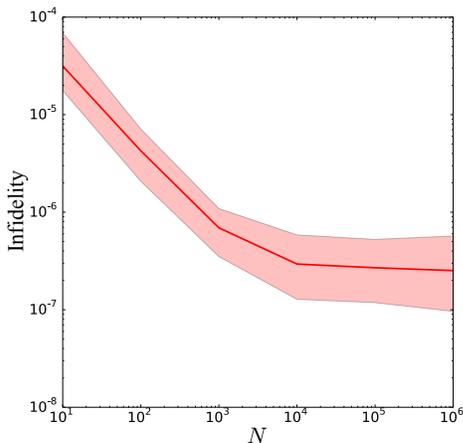


FIG. 3. (Color online) For the single qubit random gate design problem, plotted is the *infidelity* (1-fidelity) of the ACRONYM as a function of  $N$ , the number of experiments per fidelity estimate at  $k = 10^4$  iterations. The dimension of the control space is  $p = 10$ . The thick line is the median of the data and the shaded region is the interquartile range (the middle 50% of the data). This demonstrates, with the results of Fig. 2, that more experiments per iteration of the algorithm are not necessary to converge. At some point (here  $\approx 10^4$ ), a constant number of experiments suffices.

gain in performance. This is important as it implies that very few overall experiments indeed are needed to converge to the target unitary. This is another way of seeing that additional experiments are not needed at each iteration in order to converge—a constant number of experiments suffices to converge.

## B. C-NOT gate

In the first example, we considered single qubit gates, which are not sufficient for quantum computation. To enable universal quantum computing we require a two-qubit entangling gate, the most commonly considered of which is the C-NOT [1]. Let the target  $U_T$  then be the C-NOT gate and the Hamiltonian be

$$H(t) = H_{\text{drift}} + H_c(t), \quad (10)$$

where the drift Hamiltonian consists of local Zeeman terms and a Heisenberg exchange interaction:

$$H_{\text{drift}} = \delta_1 Z_1 + \delta_2 Z_2 + \frac{J}{2}(X_1 X_2 + Y_1 Y_2 + Z_2 Z_2), \quad (11)$$

and the control contains the transverse terms

$$H_c = c_{x,1}(t)X_1 + c_{x,2}(t)X_2 + c_{y,1}(t)Y_1 + c_{y,2}(t)Y_2, \quad (12)$$

where the subscripts denote which qubit the operator acts on. For the simulation, we take  $\delta_1 = -1\delta_2 = 10J = 1$  and have the controls be piecewise constant on the intervals  $\{0, 1, 2, \dots, q\}$  so that

$$\mathbf{c} = (c_{x,1}(1), c_{x,1}(2), \dots, c_{x,1}(q), \\ c_{x,2}(1), \dots, c_{y,1}(1), \dots, c_{y,2}(1), \dots, c_{y,2}(q)), \quad (13)$$

which is now a vector in  $\mathbb{R}^p$  with  $p = 4q$ .

Again, we consider the case of estimating the fidelity from finitely many experiments. In Fig. 4, the infidelity is plotted versus the number of iterations for both  $N = 10^3$  and  $N = 10^4$  number of experiments per fidelity for both ACRONYM and NM. In this case we used  $q = 10$ , which means the dimension of the search space was  $p = 40$ . A standard finite difference gradient approximation would

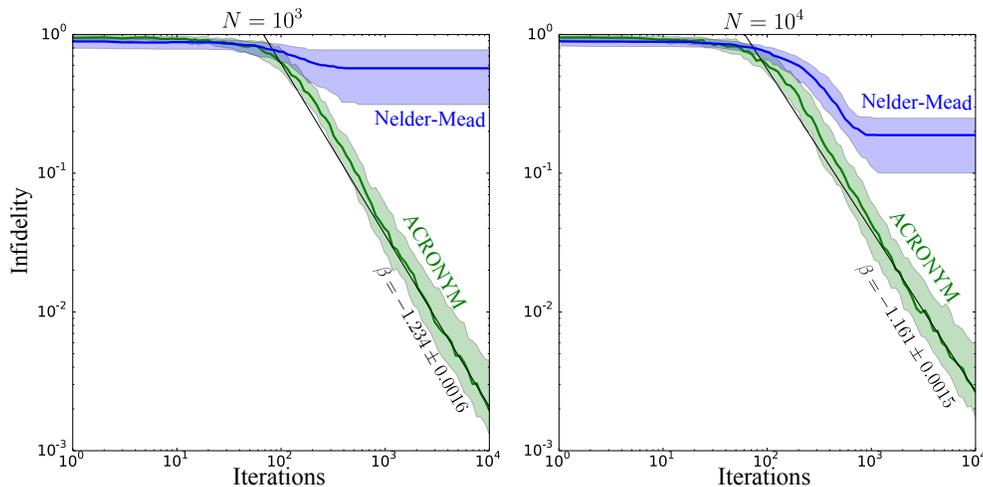


FIG. 4. (Color online) For the control problem of designing a C-NOT gate, plotted is the *infidelity* (1-fidelity) of ACRONYM and NM as a function of the number of iterations for  $N = 10^3$  and  $N = 10^4$  experiments per iteration. The thick lines are the median of the data and the shaded region is the interquartile range (the middle 50% of the data). As shown for the single qubit problem in the figures above, ACRONYM converges while NM does not.

require  $M = 80$  fidelity estimations per iteration while ACRONYM requires exactly  $M = 2$  (for this problem NM used an average of  $M = 3.23$  function calls per iteration—62% more than ACRONYM). From Fig. 4, we see that ACRONYM continues to convergence after NM has saturated which occurs quite rapidly for even  $N = 10^4$  experiments per fidelity estimate. The convergence of ACRONYM is again  $O(k^\beta)$  with fits giving  $\beta \approx -1$ . These numerical examples suggest ACRONYM has superior performance for *in situ* control.

### C. Robustness to control noise

Finally, we show that ACRONYM is robust by considering an additional source of noise, namely imperfect implementation of the controls. To model this, we add to every experimental iteration independent zero-mean Gaussian noise on each component  $\mathbf{c}_j$  of the control vector. To be precise, each fidelity estimate is expected to be performed using the control parameters  $\mathbf{c}$ , but what actually happens is  $\mathbf{c} + \boldsymbol{\epsilon}$ , where  $\boldsymbol{\epsilon}$  is drawn according to a zero-mean multivariate normal distribution with diagonal covariances of  $10^{-2}$ . The results are shown in Fig. 5. As in Fig. 2 (right), ACRONYM used  $N = 10^4$  experiments for each fidelity evaluation (2 per iteration, recall). However, we gave NM exact fidelity values—essentially infinitely many experiments for free. Both methods were then subjected to control noise of strength (standard deviation)  $10^{-2}$ . We see that even when NM has perfect fidelity evaluation, the added noise renders the method useless (additional noise on fidelity estimation leave the NM approach learning nothing at all). Whereas, ACRONYM is robust to the added noise maintaining converge at a rate given by  $\beta \approx -1$ . This shows that ACRONYM is robust to many simultaneous sources

of noise, both statistical and technical.

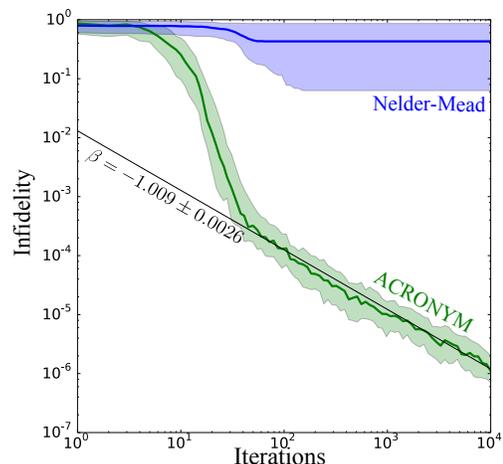


FIG. 5. (Color online) Back to the single qubit problem, now with added with control noise. Plotted is the *infidelity* (1-fidelity) of the ACRONYM and NM algorithms as a function of the number of iterations. ACRONYM used  $N = 10^3$  experiments per function evaluation, but NM was given exact fidelity values. Both methods were subjected to independent zero-mean Gaussian noise on each control component with standard deviation  $10^{-2}$ . Even with exactly fidelity values, NM is limited by the control noise. ACRONYM is robust to both statistical and control noise.

## IV. CONCLUSION AND DISCUSSION

Before we conclude, we distinguish some further discussion points, speculation and directions for future research.

### A. More on noise

So far, we have assumed that the noise on the fidelity estimation instances is completely independent and identically distributed, and we have shown that ACRONYM, by virtue of accounting for the stochasticity of the fidelity evaluations, is not limited by this type of noise. However, in real experiments, noise comes in more flavors. So, what other kinds of noise is ACRONYM robust to? The original convergence proof of Spall [24] only required that the expected difference of the noise at each function evaluation be zero. In other words, the noise need not be assumed independent. Thus, for example, low-frequency drifts should not be a problem, but higher-frequency drifts may result in a bias in the gradient estimate at each evaluation. Conveniently, in most scenarios, it is the low frequency noise that is more difficult to detect and correct while the problem high frequency noise is more routinely dealt with.

Another common example of noise in the quantum device characterization literature is state preparation and measurement (SPAM) error. As demonstrated in randomized benchmarking and *self-consistent* approaches [31, 32], however, SPAM errors can be cleverly mitigated. In any case, the story remains the same as above: if the noise biases the gradient estimates, the algorithm will not necessarily converge.

We note that while ACRONYM solves the optimization problem even in the presence of additional control noise, the final implementation itself will not be robust to the same noise. That is, ACRONYM finds the controls which would have the highest fidelity without noise, but does not find a set of controls have the same robustness properties as typically sought after, such as optimal average or worst-case fidelity when varying over the distribution of noise (see, for example, [29, 30]). In a sense, ACRONYM solves the problem—in the presence of noise—as if the noise were not present. One can imagine, for example, using such a technique to “tune-up” gates first in the presence of noise before spending resources to fine-tune the control mechanism. On the other hand, it would also be interesting to combine this approach to the robustness techniques in, for example, [29, 30].

### B. More on Nelder-Mead

While ACRONYM requires only  $M = 2$  function calls per iteration, NM is still doing impressively well at only  $M \approx 3.5$ , on average. Unfortunately, there is little more we can say since, although NM is widely popular and successful, it still lacks (after 50 years!) a satisfactory convergence proof [33]. Moreover, in many applications, it is found to underperform in high dimensions [33].

As countless others have empirically observed, NM does work very well in some cases. In fact, we have seen that there are some regimes where NM outperforms ACRONYM. These are typically earlier in the search

where the noise in the function call is dominated by the distance in the objective function to the optimal point. Perhaps, then, a more efficient adaptive protocol exists which begins with NM and switches to ACRONYM?

Although we have harped on NM for not performing well in the presence of noise, we do note that some variants of NM have been proposed to deal with stochasticity in the objective function [34]. It is unclear how well such methods would perform on quantum control problems.

### C. Learning-type control

Recall the update rule for the controls:

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \alpha_k \mathbf{g}_k. \quad (14)$$

Generically, one could have a law of the form:

$$\mathbf{c}_{k+1} = \mathcal{A}(\mathbf{c}_k), \quad (15)$$

with the obvious demand that the optimal solution be the unique fixed point of  $\mathcal{A}$ . Specific algorithms in classical control theory of this type, dubbed *learning-type control*, typically use linear models for the output (see, for example, [35]) and are hence not immediately applicable to the quantum control problem. It would be interesting, however, to look to this more general formalism for potential improvements over ACRONYM. Ideally, we would like to replace the two function calls demanded by SPSA by one—perhaps allowing that the updated controls are a specification of experiment to be performed. This would further minimize the role of additional classical computation and may give new insights into quantum automation and learning problems.

### D. Parametrization freedom

The *ex situ* paradigm of pulse design, due to the enduring inefficiency of simulating arbitrary time-dependent Hamiltonians on a classical computer, imposes a preference on the parametrization of the controls. In most cases, the preferred parametrization is one that leads to the Hamiltonian being piecewise time-independent in some frame, simplifying the computation of the overall quantum propagator on a classical computer (as a time-ordered product of exponentials instead of the more costly time-ordered exponentiation of an integral.) This parametrization is further encouraged by the ability to approximate the gradient of the fidelity given the already computed step-wise propagators [4].

Switching to an *in situ* setting allows for more freedom in the choice of parametrization, for one is no longer required to simulate the quantum dynamics on a classical computer. That is to say, the choice of the parametrization can be driven by other requirements or preferences.

For example, one could potentially choose a parametrization with few parameters; greatly reducing the parameter space over which the optimization is performed. Another could look for a parametrization that has a better fidelity landscape; making the identification or optimization problem easier. Yet another parametrization could more readily and naturally specify other constraints on the pulse form like frequency bandwidth or frequency selectivity. In general, we posit that any parametrization that deterministically maps the parameter space onto realizable wave forms can be used with *in situ* algorithms.

It is understood of course that time-ordering is important in quantum mechanics. Yet, one wonders if thinking about global approaches as opposed to temporally-local parametrizations would be beneficial, since, after all, the goal is to design an overall propagator as opposed to an instantaneous Hamiltonian. That is to say, what matters in quantum gate design is the destination not the route.

At any rate, we expect that this freedom of parametrization will allow researchers to choose parametrizations that are more natural or better suited to their specific problems, and will encourage the diversification of approaches, from which an evolved approach better suited to the general problem of pulse design will arise.

## E. Conclusion

In this work, we have introduced ACRONYM, a stochastic optimization algorithm to design *in situ* control sequences for quantum information processing tasks. The fact that fidelity estimation can be done efficiently (via randomized benchmarking, for example) and our algorithm requires a constant number of experiments per iteration—regardless of the dimension of the control space—implies that ACRONYM is efficient. We have also demonstrated that it is robust not only to the statistical noise inherent in *in situ* fidelity estimation but also to noise on the control fields. Moreover, ACRONYM is implementation independent—it requires only that the controls produce changes in the fidelity which can be estimated via experiment.

## ACKNOWLEDGMENTS

O.M. acknowledges the difficulty inherent in pinpointing the time of birth of ideas, let alone the moment of conception of insight, and in this light, thanks everyone he has talked with over the years about *in situ* control, in particular M. P. da Silva, C. A. Ryan, J. Emerson, R. Laflamme, and D. G. Cory. CF was supported by National Science Foundation grant number PHY-1212445, the Canadian Government through the NSERC PDF program, the IARPA MQCO program, the ARC via EQuS project number CE11001013, and by the US Army Research Office grant numbers W911NF-14-1-0098 and W911NF-14-1-0103.

- 
- [1] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge University Press (2010).
  - [2] C. M. Caves, *Quantum-mechanical noise in an interferometer*, *Physical Review D* **23**, 1693 (1981).
  - [3] J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., *Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light*, *Nature Photonics* **7**, 613 (2013).
  - [4] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, *Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms*, *Journal of Magnetic Resonance* **172**, 296 (2005).
  - [5] R. S. Judson and H. Rabitz, *Teaching lasers to control molecules*, *Physical Review Letters* **68**, 1500 (1992).
  - [6] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control* (Cambridge University Press, 2010).
  - [7] C. Brif, R. Chakrabarti and H. Rabitz, *Control of quantum phenomena: past, present and future*, *New Journal of Physics* **12**, 075008 (2010).
  - [8] J. Emerson, M. P. da Silva, O. Moussa, C. A. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, *Symmetrized characterization of noisy quantum processes*, *Science* **317**, 1893 (2007).
  - [9] S. T. Flammia and Y.-K. Liu, *Direct fidelity estimation from few Pauli measurements*, *Physical Review Letters* **106**, 230501 (2011).
  - [10] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, *Practical characterization of quantum devices without tomography*, *Physical Review Letters* **107**, 210404 (2011).
  - [11] O. Moussa, M. P. da Silva, C. A. Ryan, and R. Laflamme, *Practical experimental certification of computational quantum gates using a twirling procedure*, *Physical Review Letters* **109**, 70504 (2012).
  - [12] E. Magesan, J. M. Gambetta, B. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, et al., *Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking*, *Physical Review Letters* **109**, 080505 (2012).
  - [13] A. Castro, J. Werschnik, and E. Gross, *Controlling the dynamics of many-electron systems from first principles: a combination of optimal control and time-dependent density-functional theory*, *Physical Review Letters* **109**, 153603 (2012).

- [14] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. G. Cory, *Quantum process tomography of the quantum Fourier transform*, *The Journal of Chemical Physics* **121**, 6117 (2004).
- [15] N. Wiebe, C. Granade, C. Ferrie, and D. Cory, *Hamiltonian Learning and Certification Using Quantum Resources*, *Physical Review Letters* **112**, 190501 (2014).
- [16] D. Egger and F. Wilhelm, *Superconductor Science and Technology* **27**, 014001 (2014).
- [17] D. Egger and F. Wilhelm, *Adaptive hybrid optimal quantum control for imprecisely characterized systems*, *Physical Review Letters* **112**, 240503 (2014).
- [18] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, I.-C. Hoi, E. Jeffrey, et al., *Optimal quantum control using randomized benchmarking*, *Physical Review Letters* **112**, 240504 (2014).
- [19] C. Ryan, M. Laforest, and R. Laflamme, *Randomized benchmarking of single- and multi-qubit control in liquid-state NMR quantum information processing*, *New Journal of Physics* **11**, 013034 (2009).
- [20] J. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, *Randomized benchmarking and process tomography for gate errors in a solid-state qubit*, *Physical Review Letters* **102**, 090502 (2009).
- [21] J. Gaebler, A. Meier, T. Tan, R. Bowler, Y. Lin, D. Hanneke, J. Jost, J. Home, E. Knill, D. Leibfried, et al., *Randomized Benchmarking of Multiqubit Gates*, *Physical Review Letters* **108**, 260503 (2012).
- [22] L. Steffen, M. P. da Silva, A. Fedorov, M. Baur, and A. Wallraff, *Experimental Monte Carlo Quantum Process Certification*, *Physical Review Letters* **108**, 260506 (2012).
- [23] J. C. Spall, *Introduction to stochastic search and optimization: estimation, simulation, and control*, John Wiley & Sons (2005).
- [24] J. C. Spall, *Multivariate stochastic approximation using a simultaneous perturbation gradient approximation*, *IEEE Transactions on Automatic Control* **37**, 332 (1992).
- [25] P. Sadegh and J. C. Spall, *Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation*, *IEEE Transactions on Automatic Control* **43**, 1480 (1996).
- [26] This parametrization of the controls as piecewise constant envelopes, viewed in the *in situ* paradigm, is largely a vestigial adaptation that arose due to the hardship associated with the classical simulation of quantum evolution, but is an unnecessary constraint. That being said, we do use it in the numerical experiments for the same reason it flourished in the first place – the relative ease of computation. Further discussion of this point is in Section IV D.
- [27] The Nelder-Mead method is the core of both Ad-HOC [17] and ORBIT [18]. Since we do not have access to the exact algorithms used in Ad-HOC and ORBIT, our comparisons to ACRONYM will be made to the Nelder-Mead algorithm as implemented by SciPy [28].
- [28] E. Jones, T. Oliphant, P. Peterson, et al., *SciPy: open source scientific tools for python* (2001), [scipy.org](http://scipy.org).
- [29] R. L. Kosut, M. D. Grace and C. Brif, *Robust control of quantum gates via sequential convex programming*, *Physical Review A* **88**, 052326 (2013).
- [30] D. Hocker, C. Brif, M. D. Grace, A. Donovan, T.-S. Ho, K. M. Tibbetts, R. Wu and H. Rabitz, *Characterization of control noise effects in optimal quantum unitary dynamics*, *Physical Review A* **90**, 062309 (2014).
- [31] Seth T. Merkel, Jay M. Gambetta, John A. Smolin, Stefano Poletto, Antonio D. Croles, Blake R. Johnson, Colm A. Ryan, and Matthias Steffen, *Self-consistent quantum process tomography*, *Physical Review A* **87**, 062119 (2013).
- [32] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk and P. Maunz, *Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit*, [arXiv:1310.4492](https://arxiv.org/abs/1310.4492) (2013).
- [33] F. Gao and L. Han, *Implementing the Nelder-Mead simplex algorithm with adaptive parameters*, *Computational Optimization and Applications* **51**, 259, (2012).
- [34] E. J. Anderson and M. C. Ferris, *A Direct Search Algorithm for Optimization with Noisy Function Evaluations*, *SIAM Journal on Optimization* **11**, 837 (2001).
- [35] Y. Wang, F. Gao and F. J. Doyle III, *Survey on iterative learning control, repetitive control, and run-to-run control*, *Journal of Process Control* **19**, 1589 (2009).