# Advantages of mixed unitary operators for quantum information processing

Anthony Polloreno*
*Rigetti Computing, Berkeley, CA*

Kevin C. Young
*Sandia National Laboratories, Livermore, CA*
(Dated: December 15, 2018)

...

## I. INTRODUCTION

The past decade has seen a dramatic increase in the performance and scale of quantum information processors (QIPs). Gate fidelities are now routinely in the 99% to 99.99% range, and dozens of individually-addressable qubits are becoming available on integrated devices. While these advances represent important steps forward on the path towards a computationally useful QIP, the quantum advantage milestone has yet to be definitively reached. The limiting factor, of course, is errors in the quantum gate operations.

The impact of an error in a quantum gate depends strongly on both the magnitude and the nature of the errors. Systematic, or *coherent*, errors can arise from poorly calibrated controls or imperfect gate compilations that induce repeatable, undesired unitary errors on the state of a QIP. Errors of this type are correlated in time and add up coherently. They are are computationally expensive to model and it is difficult to place analytic bounds on circuit performance. Contrast this against random, or *stochastic*, errors, which result from high-frequency noise in the controls or the environment. Systems with stochastic errors can be modeled by defining a rate of various discrete errors in the system, such as a bit flip or phase flip. These errors are significantly easier to simulate on a classical computer, and their impact on quantum circuits is much easier to estimate.

Despite the relative ease of modeling stochastic errors, coherent errors are often much more likely to appear in QIPs. Drifting control parameters or environmental variables can easily have long correlation times that result in errors which is strongly coherent over the length of a quantum circuit. While these errors can often be reconstructed using various tomographic techniques, their impact is difficult to predict. The diamond distance can be used to estimate the failure probability of a quantum circuit, but the diamond distance can grow quadratically with repeated application of gate with coherent errors. For long circuits, this can add up extremely quickly. Recent work by Hastings and Campbell, however, has shown that coherent noise can be strongly suppressed by probabilistic sampling over various implementations of the target quantum gates. This averaging results in quantum processes with diamond distances that grow only linearly quadratic in the over/under rotation angle of the component gates. [Clear this up.]

In this article we discuss various applications of mixed unitary controls, and show that the advantages of this approach can be made robust to drift in the target gates. In this article, we discuss an optimal control approach to the mixed unitary control design problem. We apply our methods in simulation where we construct single- and two-qubit mixed unitary controls which are robust to drift and uncertainty in the control parameters. We further present an experimental implementation of single-qubit mixed unitary controls on a superconducting qubit testbed at Rigetti Quantum Computing. Using randomized benchmarking, we are able to show a marked improvement in error rates, as well as a reduced variance in circuit outcome probabilities, indicating a reduction in the coherence of the error.

## II. MATHEMATICAL PRELIMINARIES

Quantum gate operations are implemented by applying a sequence of classical control fields to some set of qubits. Fluctuations in the environment or imperfections in the controls can cause the state of the qubits to change in a way that is different from what was intended. But if the gates are fairly stable with time and context, then we can describe their action on the qubit state using *process matrices*, linear, Markovian maps on the state of some qubits. When working with process matrices, it is convenient to write the system density operator using a vectorized representation, and in this article, we'll make use of the generalized Bloch vector,

$$\vec{\rho} = \mathrm{Tr}\,(\rho\vec{\Sigma}), \tag{1}$$

where $\vec{\Sigma}$ is a vector of all $4^n$ $n$-qubit Pauli operators. The action of a gate is then given by the usual matrix multiplication:

$$\rho \to \mathcal{G}\rho = \mathcal{E}\tilde{\mathcal{G}}\rho. \tag{2}$$

Here $\mathcal{G}_i$ is the target operation, $\tilde{\mathcal{G}}$ is the actual gate as implemented, and $\mathcal{E}$ is the effective error channel.

* Email: anthony@rigetti.com

$$\left(\begin{array}{c|c} 1 & \vec{0} \\ \hline \vec{m} & R \end{array}\right) \tag{3}$$

The top row of all trace-preserving (TP) maps is fixed to $\{1, 0, 0, 0, \cdots\}$. The first column, $\vec{m}$, describes any deviations from unitality, such as could arise from amplitude damping. If the error channel is unitary, then the error is coherent, and the submatrix $R$ is perfectly antisymmetric, corresponding to a rotation of the generalized Bloch vector. If $R$ is diagonal, then the error channel is Pauli stochastic, with each entry corresponding to the probability that the associated Pauli error occurs in each application of the gate. [Be more clear about what "associated" means here]. If R is symmetric but not diagonal, then the channel is still stochastic, but the random errors consist of correlated Pauli operators (such as $X+Y$). For a single qubit, this describes everything, but the situation can be slightly more complicated for more qubits.

** The vec representation may be a little more convenient here? The channel that results is:

$$\mathcal{M} = \sum_i p_i U_i^* \otimes U_i \tag{4}$$

$$= \sum_i p_i \exp(i\theta\hat{n}\cdot\vec{\sigma}^*) \otimes \exp(-i\theta\hat{n}\cdot\vec{\sigma}) \tag{5}$$

$$\simeq \sum_i p_i (\cos\theta + i\sin\theta H^*) \otimes (\cos\theta - i\sin\theta H) \tag{6}$$

$$\simeq \sum_i p_i (\cos\theta^2 I\otimes I + \sin\theta^2 H^*\otimes H \tag{7}$$

$$+ i\sin\theta\cos\theta H^*\otimes I - i\sin\theta\cos\theta I\otimes H) \tag{8}$$

In this case, $\sin\theta^2$ is always positive, so we cannot hope to eliminate the second term, but $\sin\theta\cos\theta$ may be positive or negative, and so there is hope that we could possibly combine various implementations to eliminate this term. The result would be a purely stochastic channel.

The *size* of an error in quantum gates may be quantified in a number of ways. Two of the most metrics are the average gate fidelity, $\mathcal{F}$, and the diamond norm, $||\cdot||_\diamond$. These may be represented in terms of the error maps as

$$||\mathcal{E}||_\diamond = \sup_\rho ||(I\otimes I)(\rho) - (\mathcal{E}\otimes I)(\rho)||_1 \tag{9}$$

$$\mathcal{F}(\mathcal{E}) = ... \tag{10}$$

The diamond norm is generally linear in the over-rotation angle of a quantum operation. Also definition of average gate infidelity. The AGI is generally quadratic in the over-rotation angle of a quantum operation.

A mixed unitary channel consists of a set of unitary channels, $\tilde{\mathcal{G}}_i$, and associated weights, $\sum_i p_i = 1$. The process matrices for a mixed unitary channel is then the weighted sum of the component channels, $\tilde{\mathcal{G}}_{\rm M} = \sum_i p_i \tilde{\mathcal{G}}_i$, and the associated error channel is simply the weighted sum of the associated error channels, $\mathcal{E}_{\rm M} = \sum_i p_i \mathcal{E}_i$.

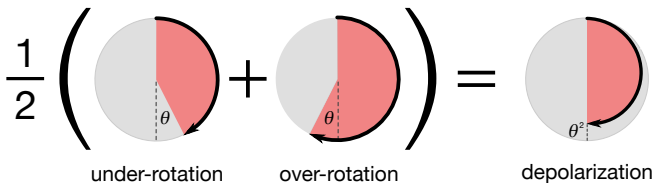


FIG. 1. An example of a balanced control solution. Using optimal control, two implementations of a $Z_\pi$ gate are designed to have equal and opposite sensitivity to errors (if one implementation over-rotates by angle $\theta$, then the other *under*-rotates by $\theta$). Each time the gate is used, one of these implementations is chosen at random. The resulting quantum channel is equivalent to a perfect implementation of the gate followed by dephasing of $\mathcal{O}(\theta^2)$.

Campbell considered the very important problem of minimizing the diamond norm of the resulting error channel. [Mention the mixing lemma of Campbell and how the Hamiltonians should form a convex set containing the origin []]. This is a particualrly appealing target, because the diamond norm provides useful error metrics on quantum circuits. However, it is not the only optimization target that can be chosen. If the ultimate goal is to produce a channel whose effect can be easily simulated, then it may be useful instead to construct a channel whose errors are Pauli stochastic.

## III. EXAMPLE MIXED UNITARY PROCESSES

There are often many possibly ways of implementing any given target quantum gate. Hastings and Campbell, for instance, consider gates compiled using the Solovey-Kitaev algorithm, for which many approximate gate compilations are possible. By selecting from these various implementations at random, they show that the resulting quantum channel can be made to have significantly reduced coherent error. As a simple example of how this occurs, consider a scenario in which we have a single-qubit and four possible implementations of a $\pi$-pulse about the $\sigma_x$ axis. The error channels for these four implementations are themselves unitary rotations about the $\sigma_x$ axis with rotation angles of $\{-2\epsilon, -\epsilon, \epsilon, 2\epsilon\}$. Such a situation could appear, for instance, if there were amplitude errors on the fields used to affect the gates, and if the control could be implemented by a rotation about the positive or negative $\sigma_x$ axis.

| Gate | $H_{\rm eff}$ | AGI | $\|\|\cdot\|\|_\diamond$ |
|---|---|---|---|
| $U_{+2\epsilon}$ | $2\epsilon\sigma_x$ | $4\epsilon^2$ | $2\epsilon$ |
| $U_{+\epsilon}$ | $\epsilon\sigma_x$ | $\epsilon^2$ | $\epsilon$ |
| $U_{-\epsilon}$ | $-\epsilon\sigma_x$ | $\epsilon^2$ | $\epsilon$ |
| $U_{-2\epsilon}$ | $-2\epsilon\sigma_x$ | $4\epsilon^2$ | $2\epsilon$ |

Make a table of these error channels.

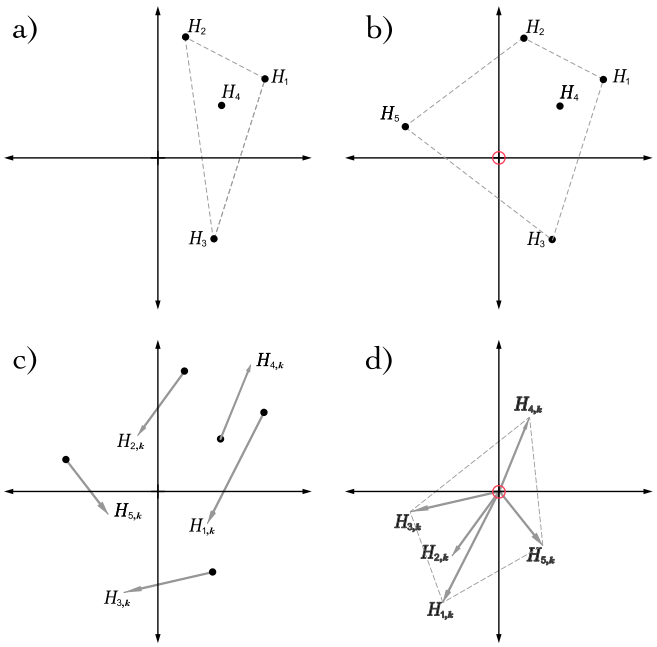


FIG. 2. A target unitary gate can be implemented a number of ways, each with a different effective Hamiltonian error. These error Hamiltonians lie in a vector space. a) Four effective Hamiltonians. The origin is not contained in their convex hull, so there are no balanced control solutions. b) The origin is contained in the covex hull after adding an additional control solution. Because there are more than $n+1$ implementations, there exist an infinite number of balanced control solutions. c) The error Hamiltonians shown with their derivative with respect to a control parameter. As this parameter drifts, a $0^{\text{th}}$-order balanced control solution may drift, leading to a first-order error. d) The derivatives also lie in a vector space. If the origin lies in their convex hull, then it may be possible to construct a $1^{\text{st}}$-order robust balanced control solution.

** This example is really useful and should be expanded. We should discuss the various optimization problems, a mixed unitary strategy, the fidelity of the resulting channel.

## IV. ROBUSTLY MIXED UNITARY PROCESSES

While mixed unitary

## V. NUMERICAL RESULTS

In the following numerical results, we explore using the methods in Section III to build MUPs. We consider the following model for a single tunable qubit:

$$H(\delta,\epsilon,t) = \epsilon\sigma_z + (1+\delta)(c_x(t)\sigma_x + c_y(t)\sigma_y) \tag{11}$$

We use the GRAPE algorithm as discussed in **??** with N=25 steps and total evolution time of $\pi$ to generate 100 candidate controls. The performance function used to terminate the gradient descent was Equation (30) in [**?** ] – the square of the Hilbert–Schmidt inner product between the generated unitary and the target unitary. The gradient, however, is the gaussian quadrature gradient described in **??**. In particular, using degree two (which should be accurate for polynomials up to degree 3) Gauss-Hermite quadrature we evaluated the gaussian weighted average gradient for standard deviation $\sigma = .001$, on both the qubit drive, and the qubit frequency. We assume that the errors on $\sigma_x$ and $\sigma_y$ are perfecly correlated, as in superconducting qubit systems that share a common RF drive.

Solving the optimization problem defined in IV with the MOSEK solver in `cvxpy` yields a 0MUP with non-trivial (nearly uniform 1%) support on all the members of the control family. However, by imposing the sparsity constraint described in IV the number of controls can be constrained to be just five. When generating a 1MUP we impose an $\ell 1$ penalty so that the algorithm favors smaller error Hamiltonians.

In our two-qubit example we consider the following model for two tunable qubits coupled by a resonant exchange interaction, similar to that in [**?** ]:

$$H(\vec{\delta},\vec{\epsilon},t) = \sum_{j=1}^{2}(\epsilon_j\sigma_z^j + (1+\delta_j)(c_x^j x(t)\sigma_x^j + c_y^j(t)\sigma_y^j)) + \frac{1}{10}(XX+YY) \tag{12}$$

In this example it was infeasible to use GRAPE to return non-trivial solutions. Instead we manually selected piecewise constant echoing sequences with 500 steps and total evolution time of $\frac{5\pi}{2}$. In particular, we considered $RX(\pi)$, $RX(-\pi)$, $RY(\pi)$ and $RY(-\pi)$ pulses used to decouple detuning errors as taking only one time step, to avoid significant errors arising from non-commutativity. The specific collection of bang-bang sequences [**?** ] considered were all combinations of simultaneous $\pi$ pulses activated at multiples of 8 steps from the first time step, and the same multiple of 8 steps prior to the last time step. To give the control family a variety of RF errors, we added on uniformly distributed errors to each $\pi$ pulse, between $-.25\%$ and $.25\%$. This resulted in 1024 controls, over which we performed the convex optimization described in IV, with neither sparisty, nor $\ell 1$ constraints.

## VI. EXPERIMENTAL RESULTS

Here we present experimental results from implementing our routine on a fixed-frequency superconducting transmon qubit. In particular, we used qubit 8 on the Rigetti 19Q-Acorn chip, whose characterization can be found in [**?** ]. To implement a MUP on this qubit, four incorrectly calibrated Gaussian pulses were produced by

scaling the pulseshape amplitude for a calibrated 10 sample 50ns $RX(\frac{\pi}{2})$ pulse by 106.4%, 103.9%, 93.7% and 91.2%.

Using the COBYLA minimizer in `scipy.optimize`[**?** ] to perform constrained minimization via Sequential Quadratic Programming[**?** ], we minimized the off–diagonal elements of the resulting mixed process. [Explain why this an okay thing to do, or reference the section where we explain it.] To benchmark the quality of the new MUP, we then performed six randomized benchmarking experiments[**?** ]: one for each over– and under–calibrated pulse, one for the calibrated pulse, and one for the mixed process. We used 1000 shots per experiment, 10 sequences per sequence length, for sequence lengths of $2, 4, 8, 16, 32$ and 64. In each case, our Clifford operations were decomposed into RX($\frac{\pi}{2}$)) and RY($\frac{\pi}{2}$)) pulses. In our implementation, these gates are implemented using the same pulse envelope definitions and control electronics, phase shifted by $\frac{\pi}{2}$ radians, and are therefore subject to identical miscalibration errors. The results are shown in Figure 7 for sequence lengths $L = 64$. The complete dataset can be found at [**?** ].

Using `scipy.optimize.curve_fit` to fit the randomized benchmarking data to Equation 6 in [**?** ] with $B_0$ set to .5, and bounds between (0, 1), we find one-qubit gate fidelities of 99.3% for the calibrated pulse, 98.9% for Pulse1, 99.1% for Pulse2, 98.9% for Pulse3, 98.5% for Pulse4, and 99.2% for the MUP. In the cases of each miscalibrated pulse, the second moment of the RB number[**?** ] for the sampled sequences extends significantly, which is in agreement with the results shown in [**?** ]. Specifically, for non-Markovian error models noise will manifest as gamma distributed points for each sequence length. On the other hand, Markovian noise, such as depolarizing noise, will result in Gaussian distributed fidelity estimates for each randomized benchmarking sequence length. We see that the coherently miscalibrated controls have long tails, consistent with gamma distributed random variables, while the calibrated and randomized implementations both have much shorter tails, consistent with Gaussian distributed random variables.

## VII. ACKNOWLEDGEMENTS

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

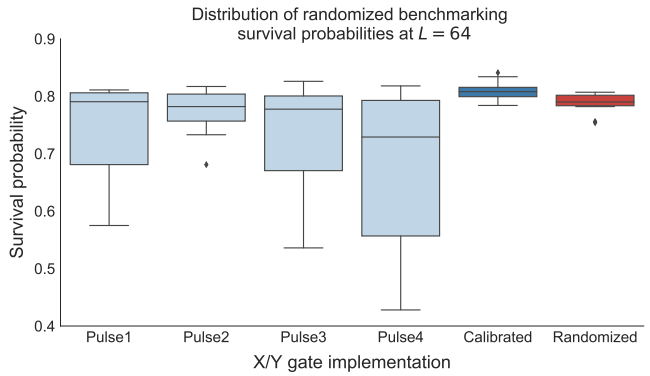


FIG. 3. Randomized benchmarking experiments ran using different pulse definitions. The four plots on the left are from the incorrectly calibrated pulse, while the top right is the calibrated pulse, and the bottom right is the BCS.

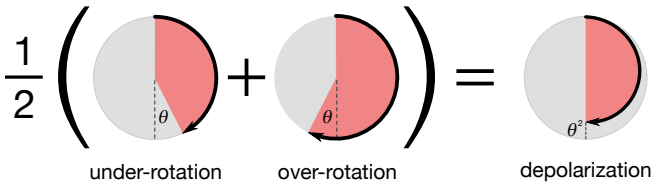


FIG. 4. An example of a balanced control solution. Using optimal control, two implementations of a $Z_\pi$ gate are designed to have equal and opposite sensitivity to errors (if one implementation over-rotates by angle $\theta$, then the other *under*-rotates by $\theta$). Each time the gate is used, one of these implementations is chosen at random. The resulting quantum channel is equivalent to a perfect implementation of the gate followed by dephasing of $\mathcal{O}\left(\theta^2\right)$.

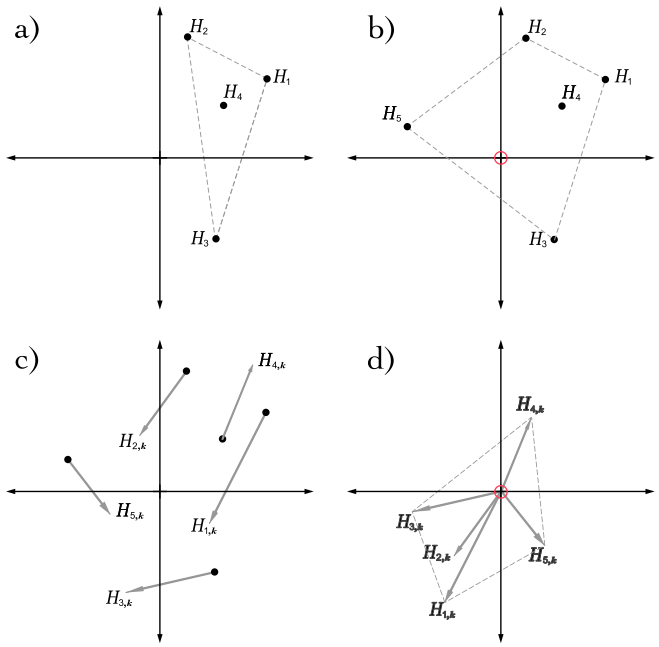


FIG. 5. A target unitary gate can be implemented a number of ways, each with a different effective Hamiltonian error. These error Hamiltonians lie in a vector space. a) Four effective Hamiltonians. The origin is not contained in their convex hull, so there are no balanced control solutions. b) The origin is contained in the covex hull after adding an additional control solution. Because there are more than $n+1$ implementations, there exist an infinite number of balanced control solutions. c) The error Hamiltonians shown with their derivative with respect to a control parameter. As this parameter drifts, a $0^{\text{th}}$-order balanced control solution may drift, leading to a first-order error. d) The derivatives also lie in a vector space. If the origin lies in their convex hull, then it may be possible to construct a $1^{\text{st}}$-order robust balanced control solution.

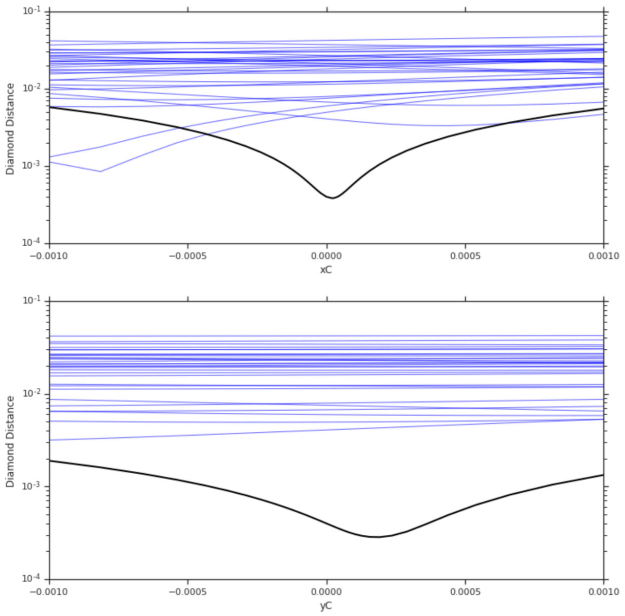


FIG. 6. Placeholder image until we figure out what to plot. This currently shows that for one of the control the diamond norm decreased by over an order of magnitude. The labels need to be bigger.

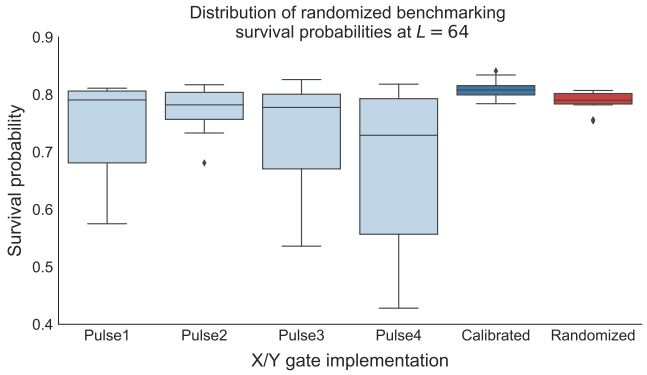


FIG. 7. Randomized benchmarking experiments ran using different pulse definitions. The four plots on the left are from the incorrectly calibrated pulse, while the top right is the calibrated pulse, and the bottom right is the BCS.

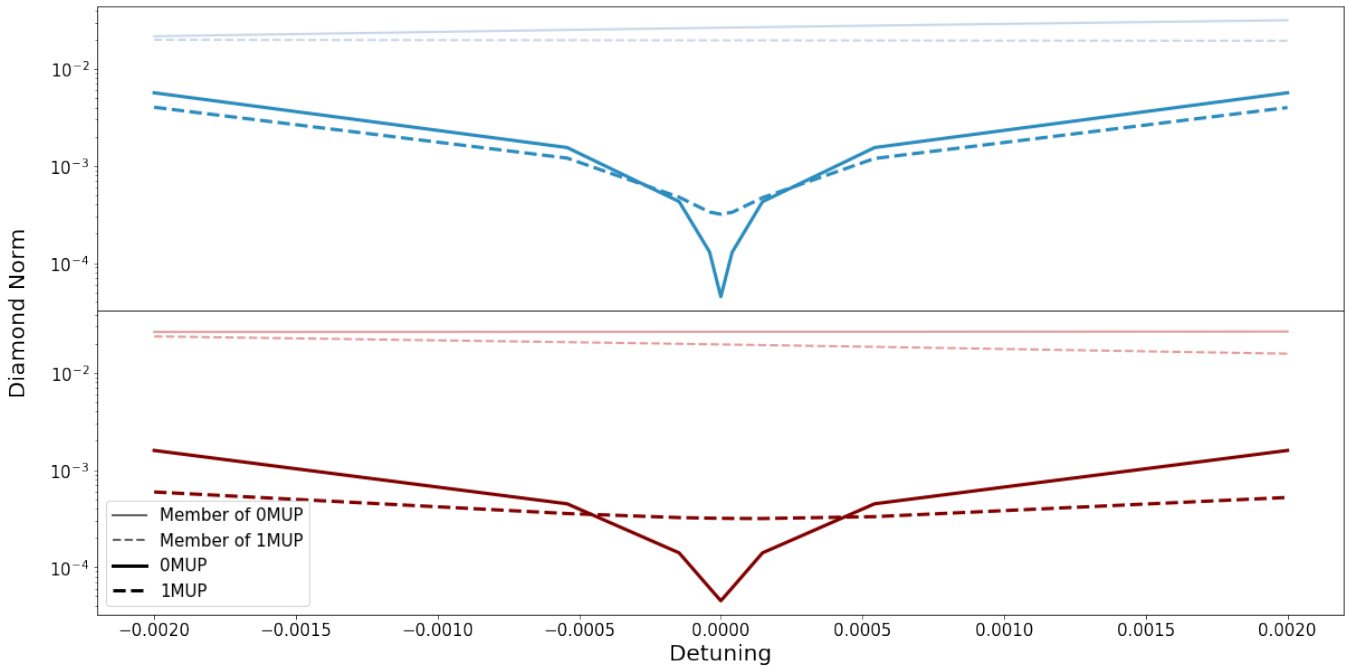


FIG. 8. Numerical results comparing a 0MUP to a 1MUP for a single tunable qubit. Shown with lower alpha values are example members of the control families, with diamond norms of around $10^{-2}$. The 0MUP can be seen to outperform the members of the control family by two orders of magnitude at the origin, and the 1MUP by an order of magnitude. However, the 1MUP can be seen to be flatter. In paritcular, for small detunings, the 1MUP performs better than the 0MUP.