

I. CHOOSE YOUR OWN ADVENTURE

A. 1Q

In the numerics and experiments run in this paper, a variety of techniques were employed to guide the convex solver to desired solutions. In the one qubit examples we enforced sparsity using the convex optimization problem defined in [?] so that the 0RBCs generated would have non-trivial support on only 10 of the 500 controls used. In addition we added a hyperparameter to penalize for the l1 norm of the constituent Hamiltonians. This was used in the 1RBCs for the one qubit controls so that we could trade off some of the insensitivity of the 1RBC for better performance at the origin.

Both the 0RBC and the 1RBC in the one qubit examples were constructed by solving the following equation using MOSEK:

$$\begin{aligned} & \underset{\omega_j \geq 0, |\omega|_1=1}{\text{minimize :}} \quad ||D_\ell^T \omega|| \\ & \text{subject to:} \quad \forall i < \ell, \sum \omega_j D_j^i = 0 \end{aligned} \quad (1)$$

The elements of the one qubit control families were generated using the GRAPE algorithm with N=25 steps and total evolution time of π . The performance function used to terminate the gradient descent was equation (30) in Khaneja's paper - the square of the Hilbert-Schmidt inner product between the GRAPE unitary and the target unitary. The gradient, however, is the gaussian quadrature gradient described in the main text. 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 a $\sigma = .001$, on both the qubit drive, and the qubit frequency.

B. 2Q

In the two-qubit case the situation was more difficult - GRAPE would not return non-trivial solutions so we manually selected piecewise constant sequences with 500 steps and total evolution time of $\frac{5\pi}{2}$. We then prepared 1024 controls. In this case, we treated the $RX(\pi)$, $RX(-\pi)$, $RY(\pi)$ and $RY(-\pi)$ pulses used to decouple the tuning errors as taking only one time step, to avoid significant errors arising from mixing the eigenspaces of the interaction term. The particular collection of bang-bang sequences considered were all combinations of of simultaneous π 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 different errors, we added on uniformly distributed errors to each π pulse, between $-.25\%$ and $.25\%$. We did not apply the sparsity constraint, or any $\ell 1$ constraints to the solutions, we only applied the minimization problem defined earlier in this document.

C. Experimental

For the experimental data the four controls generated were generated by scaling an approximately calibrated $50\mu s$ Gaussian π pulse by 106.4%, 103.9%, 93.7% and 91.2%. Then, using a naive driven qubit model (no noise), we predicted the unitaries that these pulses would generate, calculated their error maps, and used COBYLA in scipy minimize to minimize the off-diagonal elements of the mixed channel. Using the weights returned from that, we ran a randomized benchmarking experiment, with the details listed in the main document.