Improving frequency selection of driven pulses using derivative-based transition suppression

F. Motzoi\textsuperscript{1,2} and F. K. Wilhelm\textsuperscript{1,3}
\textsuperscript{1}Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
\textsuperscript{2}Department of Chemistry, University of California, Berkeley, CA 94720 USA
\textsuperscript{3}Theoretical Physics, Saarland University, 66123 Saarbrücken, Germany
(Dated: May 20, 2014)

Many techniques in quantum control rely on frequency separation as a means for suppressing unwanted couplings. In its simplest form, the mechanism relies on the low bandwidth of control pulses of long duration. Here we perform a higher-order quantum-mechanical treatment that allows for higher precision and shorter times. In particular, we identify three kinds of off-resonant effects: i) simultaneous unwanted driven couplings (e.g. due to drive crosstalk), ii) additional (initially undriven) transitions such as those in an infinite ladder system, and iii) sideband frequencies of the driving waveform such as we find in corrections to the rotating wave approximation. With a framework that is applicable to all three cases, in addition to the known adiabatic error responsible for a shift of the energy levels we typically see in the spectroscopy of such systems, we derive error terms in a controlled expansion corresponding to higher order adiabatic effects and diabatic excitations. We show, by also expanding the driving waveform in a basis of different order derivatives of a trial function (typically a Gaussian) these different error terms can be corrected for in a systematic way hence strongly improving quantum control of systems with dense spectra.

PACS numbers: 03.67.Lx, 02.30.Yy, 82.56.Jn, 85.25.Hv

\textbf{I. INTRODUCTION}

Spectroscopy is arguably the most commonly used experimental technique in physics [1–3]. It relies on resonance - the object of study is exposed to monochromatic radiation and responds if the radiation frequency matches a frequency of that system. In quantum systems, that frequency is the difference of two of the systems’ energies.

Complex systems usually contain a wealth of these frequencies. The ability to selectively address these frequencies defines the spectral resolution. The limitation of spectral resolution can be twofold: on the one hand, the frequencies forming spectral lines are intrinsically broadened by decoherence. On the other hand, an ideally monochromatic external excitation is only a convenient fiction - in reality, the bandwidth of that external signal is limited by a scale proportional to $1/T$ where $T$ is the duration of the experimental pulse. In magnetic resonance, e.g., certain spectral lines can only be reached through complex pulse sequences that all need to be executed within the relaxation time of the system. Consequently, a wealth of techniques has been developed that reaches fine spectral selectivity with pulses of limited duration, including 2D-spectroscopy[3–5].

Quantum technologies such as quantum computing are often based on spectroscopic ideas [6–9]. In fact, the already mentioned spin resonance is a primary candidate for the implementation of quantum computing [10–12]. This means that the quantum mechanical transitions corresponding to certain quantum logic operations are typically addressed through their transition frequency. This can occur on the level of single qubits, when the two states representing the qubit are taken out of a complex spectrum with low anharmonicity such as it is the case in superconducting qubits [13–17]. It can occur when multiple qubits are in close spatial proximity, much closer than the spatial resolution of the external field, as it is the case in spin resonance [1, 3, 4, 8]. It can also occur if single elements are multifunctional, e.g., when a single qubit contains transitions pertinent to local rotations as well as to coupling elements [18–23]. Examples of gate operations that contain transitions on single elements are the NOT gate for single qubits and the controlled NOT and iSWAP gates as multi-qubit operations [8]. Note that seriously scalable quantum computing implementation candidates typically do not rely on spectroscopic resolution alone and at least contain some element of local addressability. Yet, clearly, a crowding of the frequency spectrum will be detrimental to both spectroscopy and coherent quantum control.

In quantum information, it is a key requirement to perform a large number of highly accurate operations well within the coherence time of the system. Thus, the challenge of reaching good enough spectral addressability in short times is of particular significance. Now a key difference between spectroscopy and quantum control is the pursuit of selective excitation: Spectroscopy is an analytic technique to find energy levels through transition frequencies, hence we want to guarantee that beyond a narrow band around the desired transition, excitation profiles are suppressed. In quantum control, the spectrum is well characterized and the positions of undesired transitions are known, hence, it is sufficient to suppress the excitation profile at those frequencies. This paper aims primarily at the second approach.

Having a non-vanishing gap between energy levels is
also the precondition for applying the quantum adiabatic theorem. It turns out, as will be made explicit later, that there is an equivalence between the spectral excitation at an undesired transition and the inability to stay adiabatic in the trajectory through parameter space taken by the controls. Several studies have been undertaken to use the predictions of the adiabatic theorem to avoid or to cancel the unwanted excitation \cite{24–29, 31–33}. In particular, Ref. \cite{26} shows that including a control operator to counter the diabatic error can emulate adiabatic dynamics, and demonstrates how adding this (Lorentzian) control can improve population transfer using chirped Adiabatic Passage techniques. Furthermore, Ref. \cite{27} considers driving rotations on a qubit whilst another transition nearby in frequency constitutes leakage out of the qubit subspace. The result shows that one can simultaneously rotate one transition while avoiding the other by using an off-phase derivative of the driving waveform to cancel the diabatic error, allowing for an adiabatic expansion of the joint dynamics, and was first verified experimentally in Refs. \cite{34, 35}. Ref. \cite{33} considers removing the diabatic error when multiple homogenous transitions are avoided for an Ising lattice. Ref. \cite{29} considers the general case when multiple inhomogeneous transitions are present. By using a Schrieffer-Wolff transformation the authors show how (in principle) each order in the expansion can be optimized numerically to minimize the aggregate error, in particular when using a constrained set of controls.

In this paper, we further expand on these methods by constructing analytical protocols for removing multiple unwanted transitions or higher order errors. This is accomplished using a pseudo-adiabatic expansion in a way that properly tracks the order of different types of terms in the expansion. In effect, the technique generalizes the Derivative Removal by Adiabatic Gate (DRAG) protocol of Refs. \cite{27, 29} by including a set of higher-order derivatives. In the lowest order of perturbation theory these constitute a basis with which a linear set of equations approximating the differential equations giving the effective spectrum of the waveform can be solved. Moreover, higher order effects such as couplings to higher states in an anharmonic ladder can be similarly removed using extra derivatives to satisfy the additional constraints introduced by the higher-order effects.

The paper is organized as follows: in Sec. II, we introduce the problem of selectivity and in Sec. III discuss it as an application of semi-classical sideband suppression, deriving an asymptotic upper bound on off-resonant excitation related to higher derivatives; in Sec. IV, we show how the selectivity criteria can be derived for a quantum algebra and define different ways to generalize it to multiple transitions; in Sec. V we apply the formalism to a set of frequency-separated qubits and show how to use it to suppress crosstalk between them when using a common drive. In Sec. VI, we treat higher-orders in the problem of selective driving by considering a ladder of connected transitions and show using higher derivatives can prevent the (adiabatic) expansion from diverging. Sec. VII discusses frequency selectivity and gives the example of very short pulses where precise Rabi-like rotations can be maintained using the same selectivity criteria.

II. QUANTUM SELECTIVITY CRITERIA

The controls that are used to manipulate quantum systems, typically external AC fields, can often neither spatially nor by selection rule distinguish between the quantum transition that is being controlled and other quantum transitions. This can be mitigated if all these transitions have distinct transition frequencies \( \omega_{j,k} = E_k - E_j \) where \( E_j \) is the energy eigenvalue of state \( j \) and here and hereafter we use natural units with \( \hbar = 1 \). If we now drive the system control indexed by \( l \) with a drive frequency \( \omega^l_j \) that is much closer to a specific transition frequency labelled by \( j(l), k(l) \) than to any other, and if this control has an appreciable matrix element \( \hat{\Gamma}^{l}_{jk} \) for this transition, only it will be driven, and no other transition. We will quantify this statement below and outline its limitations.

We start by assuming a Hamiltonian \( \hat{H}_{\text{sys}} = \hat{H}_0 + \hat{H}_{\text{control}} \) and work in the basis of eigenstates of \( \hat{H}_0 \). We can formalize the statement about spectral selectivity by assuming that the drive Hamiltonian has some appreciable matrix elements for multiple quantum elements in the system, that is we have the control Hamiltonian

\[
\hat{H}_{\text{control}}(t) = \sum_{l=0}^{p-1} \Omega_l(t)e^{-i\phi_l} \sum_{j,k} \hat{\Gamma}^{l}_{jk} + \text{h.c.},
\]

where there are \( n \) matrix elements (transitions) in the system and \( p \) drives to control them.

As a typical example, this can arise if we consider \( n \) qubits and a collective drive composed of \( p \) frequencies, each of which is meant to address a particular qubit but has additional, unintended crosstalk on the rest of the \( n \) qubits, as it e.g. occurs in NMR. Then the full Hamiltonian will read

\[
\hat{H}_{\text{control}} = \sum_{l=1}^{p} 2\Omega_l(t)e^{-i\phi_l} \cos \left( \int_0^t \omega^l_j(t')dt' \right) \sum_{m=1}^{n} \hat{\sigma}_m^+ + \text{h.c.}
\]

\[
\hat{H}_0 = \sum_{m=1}^{n} \frac{\omega^1_m(t)}{2} \hat{\sigma}_m^z,
\]

where we have left all terms time-dependent for generality. Other examples, specifically where the transitions are not disjoint (i.e., they cannot be described using a tensor sum), will be discussed in Secs. VI and VII and will have similar forms. We can better appreciate the selectivity condition by moving to the standard interaction picture, \( \hat{H}_I \), and applying the rotating wave approximation, whereupon

\[
\hat{H}_{\text{control}} = \sum_{l=1}^{p} 2\Omega_l(t)e^{-i\phi_l} \cos \left( \int_0^t \omega^l_j(t')dt' \right) \sum_{m=1}^{n} \hat{\sigma}_m^+ + \text{h.c.}
\]

\[
\hat{H}_0 = \sum_{m=1}^{n} \frac{\omega^1_m(t)}{2} \hat{\sigma}_m^z,
\]

where we have left all terms time-dependent for generality. Other examples, specifically where the transitions are not disjoint (i.e., they cannot be described using a tensor sum), will be discussed in Secs. VI and VII and will have similar forms. We can better appreciate the selectivity condition by moving to the standard interaction picture, \( \hat{H}_I \), and applying the rotating wave approximation, whereupon
\[ \hat{H}_I = e^{-i \int H_0 dt} \hat{H}_{\text{control}} e^{i \int H_0 dt} = \sum_l \sum_{j \neq k} \Gamma_{j,k}^l \]

\[ \Gamma_{j,k}^l = \lambda_{j,k}^l e^{-i \int \Delta_{j,k}(t') dt'} |j\rangle \langle k|, \]

and the offsets \( \Delta_{j,k}(t) = \omega^\theta_j(t) - \omega_{j,k}(t) \) define the distance from resonance of the transitions. The \( \lambda_{j,k}^l \) weigh the relative strengths of the different transitions, letting \( \lambda_{j,(l,k)}^l = 1 \). In the disjoint qubits example, we have \(|j\rangle = |0\rangle_m, |k\rangle = |1\rangle_m\). The evolution of a system under the interaction Hamiltonian is then given by

\[
U(0,T) = \mathcal{T} \exp \left( \sum_{j,k,t} \int_0^T \Omega_l(t)e^{-i\phi_t \Gamma_{j,k}^l(t)} dt + \text{h.c.} \right),
\]

(2.3)

where \( T \) is the evolution time and \( \mathcal{T} \) enforces time-ordering. Here, it is tacitly assumed that the envelope \( \Omega_l = \text{Re} \Omega_l + i \text{Im} \Omega_l \) is complex-valued. For the implementation of simple drive pulses, the phase is typically assumed to be constant which with appropriate choice of reference means \( \text{Im} \Omega_l = 0 \); however, later we will explicitly use the ability to control both terms independently.

Without loss of generality, we assume that to each driving field indexed by \( l \) we match a transition \( j, k \) to which it is almost resonant, identified as \( j(l) \) and \( k(l) \). In this interaction frame representation, we can formulate the sufficient conditions for selectivity: in order for the drive element Eq. 2.3 to be effective, it must oscillate more frequently than the time scale of the transition \( |\lambda_{j,k}^l \Omega_l|^{-1} \) but less than \( |\lambda_{j,k}^l \Omega_l|^{-1} \), specifically

\[
\Delta_{j,k,l} \gg |\lambda_{j,k}^l \Omega_l| \quad \forall j, k \neq j(l), k(l)
\]

(2.4)

\[
\Delta_{j(l),k(l),l} \ll |\lambda_{j(l),k(l)}^l \Omega_l|
\]

(2.5)

As \( \omega_{j,k} \) are given by the quantum system under consideration, the choice of driving frequencies \( \omega_l \) can only maximize the left hand side of Eq. 2.4 to a certain limit set by the need to obey Eq. 2.5. Thus, obeying these conditions requires to keep the control amplitudes \( \Omega_l \) low enough, which increases the duration of the control pulse, but makes the transition vulnerable to decoherence and relaxation. Thus, we practically demand that \( |\lambda_{j,(l,k)}^l \Omega_l| \gg \gamma \) where \( \gamma \) represents typical incoherent rates of the system. This constraint on addressability is a result of spectral crowding and the loss of fidelity due to the need for long pulses degrades spectroscopic techniques as well as the implementation of coherent gates in a quantum computer. Thus, the spectrum sets a speed and fidelity limit on quantum control. We will derive these conditions and bound them using the Fourier transform in the next section.

III. ASYMPTOTIC LIMIT

A. Fourier transform

It is well established\cite{4, 36–38} that for a system of qubits or spins 1/2 driven by a weak external field, i.e., for small \( \Omega/\Delta \), an accurate measure of off-resonant excitation is the Fourier transform

\[
S(\Omega, \Delta) = \int_0^T \Omega(t)e^{-i\Delta t} dt
\]

(3.1)

That is, at long times and large frequency separation \( \Delta \), the time-ordering terms in Eq. 2.3 will commute and the time-ordering operator can be dropped \cite{38}. Note that this is a limited-interval Fourier transform that can be consolidated with the regular, infinite-time Fourier transform by assuming the pulse envelope \( \Omega(t) \) vanishes outside the integration interval. The conditions for selectivity (Eqs. 2.4 and 2.5) then imply that, for large times,

\[
S(\Omega, \Delta, j, l) = \theta \delta_{j,l} \delta_{k,l}
\]

(3.2)

where \( \delta_{a,b} \) is the Kronecker delta. When the time-ordering can be dropped \cite{39}, Eq. 2.3 then gives back trivially a \( \theta \) rotation on (only) the desired transition \( j \rightarrow k \).

A caveat to this approximation is that off-resonant levels will induce additional phase errors (coming from, e.g., AC Stark shifts) for shorter times or multiple drives, coming from enforcing time-ordering. In practice, these can be corrected by some combination of adjusting resonance conditions, applying compensatory gates to undo the accumulated phase at the end of the operation, or by inserting frame transformations between operations (see Appendix A). The derivation of these phase terms will be discussed in detail in later sections when we consider the full dynamics of concrete examples (see also \cite{29}).

One well-established way to compensate spectral weight off-resonance while still maintaining a pulse of limited length is to use pulse shaping \cite{4}. For this purpose, it is customary \cite{40, 41} to use Gaussian profiles, which are well-confined Gaussians both in the frequency and time domains. In this case, the Gauss function describing \( \Omega(t) \) must be suitably chosen to start and end at zero amplitude and takes the form

\[
\Omega_G(t) = A \left( \exp \left[ -\frac{(t - T/2)^2}{2\sigma^2} \right] - \exp \left[ -\frac{(T/2)^2}{2\sigma^2} \right] \right)^m.
\]

(3.3)

Here, \( \sigma \) is the standard deviation, \( m \) is chosen such that \( m - 1 \) derivatives of the function start and end at 0, and \( A \) is chosen such that the correct amount of rotation is implemented (e.g. \( A = \pi/\sqrt{2\pi \sigma^2} \text{erf}[T/\sqrt{8\sigma}] \) for an area \( \pi \) pulse). We will follow a different strategy that, rather than reducing off-resonant excitations for a full band of
energies, eliminates excitation for one or more discrete frequencies.

B. Order counting

For this purpose, we wish to be able to quantify and remove the effect of unwanted off-resonant error. To be able to remove multiple such errors, we will want to find equivalent formulations of the error, which we will see below will be found using different orders of differentiation of the driving waveform. Finally, we will want to see the effect polynomial functions of these derivatives as these will be needed if we want to Taylor expand the dynamics (Sec. IV) in terms of these different orders. Having established the role of the Fourier transform, we can now adapt an idea from classical calculus and signal processing. We start from the excitation profile for detuning $\Delta$, then integrate by parts (IBP), assume that the envelope and its lowest $n$ derivatives vanish in both the beginning and the end of the pulse, and find

$$S(\Omega, \Delta) = \int_0^T \Omega(t) e^{-i\Delta t} dt$$

$$= -i \int_0^T \frac{d\Omega(t)}{\Delta} e^{-i\Delta t} dt$$

$$= (-i)^n \int_0^T \frac{d^n\Omega(t)}{\Delta^n} e^{-i\Delta t} dt$$

This result tells us that the spectral weight of the $n$-th derivative of the control signal will be amplified by a factor $\Delta^n$ relative to the original waveform, or asymptotically will be $\Theta(S(\Delta^n, \Omega(t), \Delta))$. Moreover, this equivalence will hold over infinitesimal intervals $[t, t + dt]$ of the full evolution as well (neglecting the boundary terms, which will cancel between intervals). Likewise it is easy to see that derivatives of polynomial functions of the waveform will obey the same formula

$$S \left( \frac{a_m}{\Delta^n} \left( \prod_{k} \left( \frac{d^k \Omega(t)}{dt^k} \right)^{n_k} \right), \Delta \right) = \Theta(S(\Delta^n, \prod_{k} \left( \frac{d^k \Omega(t)}{dt^k} \right)^{n_k}))$$

More generally, it can also be verified numerically that, asymptotically in $\Delta$,

$$S \left( \frac{\Omega(t)}{\Delta^n}, \Delta \right) = \mathcal{O} \left( S \left( \prod_{k} \left( \frac{d^k \Omega(t)}{dt^k} \right)^{n_k}, \Delta \right) \right)$$

for given $n_k$.

In the quantum limit, when in the adiabatic regime ($\Omega < \Delta$), we will see in the next sections that adiabatic expansions around a small parameter $\epsilon = \frac{\Omega}{\Delta}$ will obey the same infinitesimal-time asymptotics, as commutators of terms in $S(\epsilon, \Delta)$ will be of the next or higher order in $\epsilon$ and hence not contribute. In addition, in the extreme limit ($\Omega \ll \Delta$), the full integral over $T$ will commute with other small terms and thereby accurately predict off-resonant excitation (since the Fourier transform is a good measure). We will see in Sec. IV how these terms can be accounted for to precisely calculate gate errors for shorter times as well.

C. Engineering of the instantaneous spectrum

For now, we describe a semi-classical strategy for utilizing the correspondence between the frequency spectra of the waveform and its derivatives to suppress off-resonant excitation. The strategy is to satisfy Eqs. 3.2 by supplementing the waveform with some small auxiliary controls proportional to the derivatives,

$$\Omega(t) = \text{Re}\Omega(t) + i\text{Im}\Omega(t)$$

$$= \Omega_0(t) + \sum_{r=1}^{n/2} a_{2r} \frac{d^{2r} \Omega_0(t)}{dt^{2r}} + i \sum_{r=1}^{n/2} b_{2r-1} \frac{d^{2r-1} \Omega_0(t)}{dt^{2r-1}}$$

where $a_i$ and $b_i$ will be chosen to satisfy the selectivity constraints. For example, using a Gaussian as our base waveform $\Omega_0(t) = \Omega_G(t)$, we can engineer a hole in the spectrum at a frequency offset $\Delta$ from the driven transition. The simplest way to do this is by choosing

$$\text{Im}\Omega(t) = -\frac{\text{Re}\Omega(t)}{\Delta},$$

i.e., $b_1 = -\frac{1}{\Delta^2}$. The spectrum of this control shape is illustrated in Fig. (1)A. The zero of the frequency axis is set to the wanted transition, the undesired transition is placed at $\Delta = -3/\sigma$ ($\sigma$ being the standard deviation of the Gaussian). The Gaussian definitely has appreciable spectral weight at the unwanted transition. The derivative also has spectral weight there, so the difference with the appropriate weight ($\frac{1}{\Delta^2}$) will be zero. By construction, the derivative of the Gaussian has no spectral weight at the working transition (it is anti-symmetric) hence does not alter the spectral profile of the working transition. Note that the perturbation caused by the auxiliary control is small both in the time and frequency domains as it suppressed by a factor $\Delta^{-1}$. Thus, we see the derivatives of the function have two effects: on the one hand they result in a disproportionately large error off-resonance (Eq. 3.4); but on the other, with the introduction of a small perturbation we are able to completely
cancel out the undesired excitation from the principle waveform (Eq. 3.6).

We can apply the same technique for higher-order derivatives. For example, the second-derivative solution
\[ a_2 = \frac{1}{\Delta^2}, \]
will satisfy Eqs. 3.2 provided \( \int_0^T \Omega(t)\,dt = \theta \). The effect is demonstrated in Fig. 1B. Off resonance, unwanted transitions are cancelled at the chosen \( \Delta = \pm 2/\sigma \) while maintaining full rotation on resonance. We choose \( \Omega_0(t) = \Omega_G(t) \) with \( m = 2 \) to ensure the IBP formula is valid twice over in Eq. 3.4. The strategy may be preferable to the first derivative solution in certain cases. Since the first derivative is anti-symmetric it increases excitation at \( \Delta = +1 \), thus the second derivative may be more useful when transitions are not wanted on both sides of resonance, as would happen for a spectrum with a Liouvillian degeneracy, i.e., with distinct transitions having equal frequencies. Moreover, we can see the overall bandwidth (above a given signal-to-noise threshold—here 0.001) is decreased compared to a traditional Gaussian by about 25%, where some of the energy has been moved from the selective region to the tails where it instead falls below threshold. This could be useful when a continuum of excitations needs to be avoided, as in resonance spectroscopy. Lastly, using only controls in phase with each other implies they commute (they obey an area theorem, integrating to \( \theta \)) avoiding higher order effects such as phase shifts and rotation errors on the working transition.

Finally, let us notice again that since Eq. 3.4 holds for infinitesimal times as well, the spectrum engineering is far more effective than simply obeying Eq. 3.2. For example, the pulse shape \( \Omega(t) = \Omega_0(t) + \frac{1}{\Delta^2} e^{i(\Delta t - T)} \Omega_0(t) \) will also have the same average spectrum at the critical frequencies and benefit from being only a small perturbation; however, this solution is not valid over intervals smaller than \( T \) (\( S(\Omega(t), \Delta) \) does not vanish for small time intervals), and hence the time-ordering operator in Eq. 2.3 cannot be easily accounted for. Thus, in what follows, we will only consider instantaneous-time solutions such as we have found above, which will allow for an instantaneous-time expansion of the dynamics in Sec. IV.

\[ 1 + \sum_{r=1}^{n/2} (-1)^r (\Delta_j)^{2r} a_{2r} - \sum_{r=1}^{n/2} (-1)^r (\Delta_j)^{2r-1} b_{2r-1} = 0, \]

where \( n \) is the number of undesired transitions. Such a system of linear equations can easily be solved. For example, for \( n = 2 \), the structure of the solution is
\[ \Omega(t) = \Omega_0(t) - i \left( \frac{1}{\Delta_1} + \frac{1}{\Delta_2} \right) \Omega_0(t) + \frac{\dot{\Omega}_0(t)}{\Delta_1 \Delta_2}, \]
or alternatively
\[ \Omega(t) = \Omega_0(t) + \frac{\Delta_1^2 + \Delta_2^2}{\Delta_1 \Delta_2} i \dot{\Omega}_0(t) + \frac{i \ddot{\Omega}_0(t)}{\Delta_1 \Delta_2}. \]

For \( n = 3 \) we can use
\[ \Omega(t) = \Omega_0(t) - i \left( \frac{1}{\Delta_1} + \frac{1}{\Delta_2} + \frac{1}{\Delta_3} \right) \Omega_0(t) + \frac{\Delta_1 + \Delta_2 + \Delta_3}{\Delta_1 \Delta_2 \Delta_3} i \dot{\Omega}_0(t) + \frac{1}{\Delta_1 \Delta_2 \Delta_3} i \ddot{\Omega}_0(t). \]

While these semi-classical solutions will be valid in the limit of \( \Omega \ll \Delta \), they will become less accurate as higher-order derivatives are used, as is typical for an asymptotic expansion. It will be crucial to supplement them with corrections to rotation angle and resonance/phase errors and other higher-order quantum effects.

D. Multiplet engineering

We can generalize the semi-classical solution to suppressing multiple unwanted excitations. Specifically, for crowded spectra and high precision requirements, it may not be sufficient to only put one or two holes in the spectrum and rely on bandwidth constraints for the rest. Instead, we must now solve Eq. 3.2 for multiple offsets, \( \{ \Delta_j \} \). Plugging Eq. 3.6 into Eq. 3.2 and applying the IBP formulae Eq 3.4, we null the integrand to obtain

\[ \text{IV. FREQUENCY-SELECTIVE ADIABATIC EXPANSION} \]

In the limit of \( \Omega \ll \Delta \), the primary effect of the driving field is to time-dependently change the energies in the system. This can be understood as an application of the adiabatic theorem, or to higher order, the superadiabatic expansion [42]. On the other hand, we have seen with Eq. 3.4 that in exactly the same regime, the excitation of unwanted transitions will occur proportionally to derivatives of the waveform, and we will refer to this error as diabatic error. Our goal will be to remove both the diabatic and the adiabatic errors, which inhibit perfect rotation of the working qubit, in particular when multiple unwanted transitions exist in the system. In this section, we will show how the dynamics can be expanded in terms of a small parameter to compute and suppress these errors, in particular in extension to previous works by considering the expansion in terms of the derivatives (which we will see naturally arise in the adiabatic expansion). We will be able to characterize and suppress the order-of-magnitude diabatic effect of terms involving derivatives relative to the small parameter by using the asymptotic scaling found in Sec. III B.
We choose to work in a (computational) frame where the time-independent part of the Hamiltonian has been diagonalized (into, in general, dressed eigenstates) so that energy transitions are clearly defined by the difference in diagonal entries in the Hamiltonian matrix. We then perform a sequence of time-dependent transformations that allow us to obtain instantaneous-time control operators for which gate synthesis is trivial.

A. Rotating frame

We start by moving to a frame where all transitions between adjacent energy levels are rotating at the frequency of the drive. For clarity, we choose indices so that adjacent levels have minimal energy difference between them. Our goal, once the frame transformation has been performed, is that the matrix elements corresponding to the drive for adjacent levels will contain a term that does not oscillate (and whose only time-dependence comes from that of the drive waveform envelope). These elements will constitute the primary error with respect to selectively driving a particular transition. Other less significant terms, including counter-rotating terms, matrix elements between non-adjacent levels, and extra drive terms used to simultaneously drive other transitions, will oscillate at their sideband frequency relative to the rotating frame. Sec. VII gives an example of how these errors can be suppressed using the same formalism.

The rotating frame transformation is defined by

$$R = \exp \left( -i \sum_{k=1}^{N} \int_{0}^{T} \Delta_{k-1,k,l}(t) dt |k\rangle\langle k| \right)$$

and the transformed frame with respect to the interaction picture given by

$$\hat{H}_R = \hat{R} \hat{H} \hat{R}^\dagger + i \dot{\hat{R}} \hat{R}^\dagger$$

$$= \sum_{l=0}^{L-1} \Omega_l(t) e^{-i\phi_l} \sum_{j,k} \chi_{j,k}^l e^{-i \int_{0}^{T} \omega_{j,k-1}(t) dt} |j\rangle\langle k| + \text{h.c.}$$

$$+ \sum_{k=1}^{N} \Delta_{k-1,k,l}(t) |k\rangle\langle k|$$

(4.1)

where \( \Delta_{j,k,l} = \omega_{j,l}^k - \omega_{j,k} \) is defined as before. For the simple case outlined above then we have in particular \( \omega_{j,k-1} = 0 \) and the only time-dependence comes from \( \Omega_l(t) \). As a final preparation step, we separate the rotating Hamiltonian into wanted, unwanted, and irrelevant (e.g. diagonal) terms. Thus, we can write equivalently

$$\hat{H}_R = \hat{H}_w + \hat{H}_{uw} + \hat{H}_{diag}$$

where

$$\hat{H}_w = \Omega_i(t) e^{-i\phi_l} |j(l)\rangle\langle k(l)|$$

$$\hat{H}_{uw} \approx \Omega_i(t) e^{-i\phi_l} \sum_{j,k} \chi^l_{j,k} |j\rangle\langle k| + \text{h.c.} \quad \forall j, k \neq j(l), k(l)$$

$$\hat{H}_{diag} = \sum_{k=1}^{N} \Delta_{k-1,k,l}(t) |k\rangle\langle k|.$$ 

Note that we have neglected oscillating terms for clarity in this final form.

B. Block diagonal frame

From Eq. 4.1, we now want to move to another frame where the remaining time-dependence can be trivially calculated. To do this we will find a diagonal representation for the unwanted transitions so that only an (easily correctable) phase shift will result on these levels. The problem lies with being able to find such a diagonalizing transformation. A general transformation of this kind will be of the form

$$\hat{H}_D = \hat{D} \hat{H} \hat{D}^\dagger + i \dot{\hat{D}} \hat{D}^\dagger.$$ 

In general, for fast pulses, the second term on the right hand side can have a larger contribution than the terms that were being diagonalized, as we have seen in Sec. III B. Instead, we will use intuition from Sec. III C to define an interaction frame with respect to an auxiliary (off-phase, derivative) control operator, within which the diagonalization will be well-defined. That is

$$\hat{H}_D = \hat{D} \hat{H} \hat{D}^\dagger + i \dot{\hat{D}} \hat{D}^\dagger.$$ 

Thus, \( \hat{H}_{aux} \) is chosen such that \( D \) diagonalizes \( \hat{H}_{uw} \) provided \( \hat{H}_{aux} \) is also chosen such that it averages to zero over the time \( T \) and that it commutes with itself at different times. The final form can be written more explicitly as

$$\hat{H}_D = \hat{D} \hat{H} \hat{D}^\dagger + i \dot{\hat{D}} \hat{D}^\dagger.$$ 

$$\hat{H}_{tot} = (\hat{H}_w + \hat{H}_{uw} + \hat{H}_{diag}) + \hat{H}_{aux} \quad (4.2)$$

$$D = \exp \left( i \int_{0}^{T} \hat{H}_{aux}(t) dt \right)$$

$$\tilde{H} = D \hat{H} \hat{D}^\dagger + i \dot{\hat{D}} \hat{D}^\dagger$$

where

$$\hat{H}_{tot} = \hat{H}_w + \hat{H}_{uw} \quad (4.3)$$

$$\int_{0}^{T} \hat{\Omega}(t) dt = \theta,$$

which is the quantum mechanical equivalent of Eq. 3.2. This technique can be used to exactly solve certain transition selection/avoidance problems, such as the two-qubit crosstalk problem in Sec. V A.
In general, it is not possible to analytically diagonalize \( \hat{H}_R \) exactly and instead one will have to use perturbation theory with respect to a small parameter, here chosen as

\[
\epsilon = \max_{jk} \left( \frac{|\lambda_{jk}| |\Omega(t)|}{\Delta_{jkl}} \right), \tag{4.4}
\]

In effect, if more than one unwanted transition elements exist in the system then one will need more than the single auxiliary control waveform to cancel out the undesired dynamics. That is, one will need to solve a system of differential equations relating to the diagonalization transformation and time derivative of the transformation for each of the unwanted off-diagonal elements (Eq. 4.3). Once again, we can use the intuition from Sec. III D to solve the system using a basis of higher order derivatives as an ansatz. Let the controls be defined analogously to Eq. 3.6 by

\[
\Omega(t) = \text{Re} \Omega + i \text{Im} \Omega = \Omega_0 + \sum_{r=1}^{n} \Omega_r^Q + i \sum_{r=1}^{n} \Omega_r^l
\]

\[
\hat{H}_{\text{aux}} = \sum_r \hat{H}_r^Q + \sum_r \hat{H}_r^l
\]

\[
\hat{H}_r^Q = \Omega_r^Q(t) e^{-i\phi + i\pi} \sum_{j,k} \lambda_{j,k}^l |j\rangle \langle k| + \text{h.c.} \tag{4.5}
\]

\[
\hat{H}_r^l = \Omega_r^l(t) e^{-i\phi} \sum_{j,k} \lambda_{j,k}^l |j\rangle \langle k| + \text{h.c.}
\]

Each of the \( n \) unwanted transitions will be approximately diagonalized by a combination of a real and imaginary operator, thus there are \( 2n \) operators that define the (self-commuting for different times) transformations

\[
D(t) = \exp \left( -i \hat{Y} \right) = \exp \left( i \sum_{r} \hat{Y}_r \right) \tag{4.6}
\]

\[
E(t) = \exp(-i \hat{X}) = \exp \left( i \sum_{r} \hat{X}_r \right)
\]

where the real and imaginary operators have been applied in separate transformations and contain higher-derivative contributions. Applying the transformations in sequence gives the effective Hamiltonian

\[
\tilde{H} = D \left( E \tilde{H}_{\text{tot}} E^\dagger \right) D^\dagger + i D \dot{E} E^\dagger D^\dagger + i \dot{D} D^\dagger
\]

\[
= D \left( E (\hat{H}_{\text{diag}} + \hat{H}_{\text{aux}}) E^\dagger + \hat{H}_w + \hat{H}_{uw} \right) D^\dagger
\]

\[
+ \dot{D} \dot{X} D^\dagger + \dot{Y}
\]

\[
= \tilde{H}_w + \bigoplus_{r=1}^{n} \tilde{H}_{\text{diag}} (r)
\]

The last line will hold only if the transformations combine to time-independently diagonalize the unwanted transitions \((j - k)\) of the Hamiltonian, that is

\[
\tilde{H}_{jk} = 0, \tag{4.8}
\]

for each \( j, k \neq j(l), k(l) \). The simultaneous diagonalization of these unwanted transitions sets up a system of equations. Remarkably, the linearization of this system (Eqs. 4.6 - 4.7) in terms of \( \dot{X}_r \) and \( \dot{Y}_r \) is identical to the semiclassical system, Eq. 3.8. More concretely, we have the linear approximation

\[
\left( \hat{H}_{uw} + \sum_r ([i \hat{Y}_r, \hat{H}_{\text{diag}}] + [i \hat{X}_r, \hat{H}_{\text{diag}}] \right. \tag{4.9}
\]

\[
+ \hat{H}_w^Q + \hat{H}_w^l + \dot{\hat{X}}_r + \dot{\hat{Y}}_r \right) j_k \approx 0,
\]

which must hold for each pair \((j, k)\) not amongst the desired transition(s). Thus, starting with the unwanted transition element \((\hat{H}_{uw})_{jk}\), each successive order of transformation diagonalizes the unwanted off-diagonal transition element (first line), and leaves in its place a derivative that is only partially canceled by the auxiliary derivative control \( \hat{H}_w^Q(t) \) of that order (second line). This process is iterated with each order leaving a higher-order, off-diagonal derivative with a smaller prefactor. By the \( n \)-th order, the off-diagonal contribution has been fully removed in aggregate by all the auxiliary controls, and ensuring this happens for all unwanted pairs \((j, k)\) solves the system of equations.

Moreover, Eq. 4.7 allows \( \hat{H}_w \) and \( \hat{H}_{\text{diag}} \) to, be computed, unlike the semiclassical approach where these are assumed to stay at their bare values. Variations in \( \hat{H}_w \) from the ideal Eq. 4.3 will result in resonance and rotation errors on the working transition. Variations in \( \hat{H}_{\text{diag}} \) will be a little more subtle and will primarily result in time-dependently changing values of \( \{\Delta_j\} \) in Eq. 3.8. Last but not least, higher order in \( \epsilon \) will result not only in corrections to Eq. 4.9 but also in new (unwanted) transitions \( j - k \) not amongst the \( n \) transitions initially driven in the bare frame, Eq. 4.1.

Note that the system of linear differential equations, Eq. 4.7, has been replaced by a system of (linear) algebraic equations, Eq. 4.9, where only the multiplicative factor in front of the derivatives need be computed. That is, the derivatives form a basis for the evolution of the populations, and thus the prefactors in front of them that solve the algebraic equations also solve the differential equation at all time. Thus, the result is an instantaneous solution which solves the (unwanted excitation) problem exactly at all times in the evolution. This is similar to the conclusion at the end of Sec. III D, but now it holds to higher orders, including the fact that resonance and rotation errors that are only uncovered in the quantum mechanical treatment and do not occur in the semiclassical order can also be computed and suppressed. An example to suppressing multiple undesired transitions is given in Sec. V, which demonstrates solutions to the multi-qubit crosstalk problem.

When we further include the non-linear (higher-order) effects in the small parameter \( \epsilon \) (defined in Eq. 4.4) of the evolution, we will no longer be able to completely can-
cell terms containing derivatives, as cross-terms will arise that contain more than a single order of differentiation of the trial function (e.g. $\Omega^2 \hat{\Omega}$). Nonetheless, going to higher orders may be necessary for high-precision control or to apply further error-correcting protocols. Qualifying the effect of these terms will be important as the number of error terms will grow exponentially with each order in the expansion (though in practice very few orders will be needed to suppress the error). Moreover, expansions that calculate the effect of these terms such as adiabatic or super-adiabatic expansions will typically diverge [42, 43] and so knowing which terms cause divergence will be crucial to controlling the expansion and avoiding divergences. These systematic effect of these terms can be gauged (and reconciled with the previous paragraph) by considering that the $k$-th order in $\epsilon$ will contain at most $\frac{k^2}{2}$ different cross-terms, and these will form a basis for the evolution to that order. To do this we expand around the semi-classical, large $T$ limit (that is, the adiabatic limit) given in Sec.IIIB so that we collect terms up to each order $O(\epsilon^k)$ in the expansion. Instead of Eq. 4.7, the transformation is identified recursively up to each order by

$$\hat{H}(t) = \hat{H}^{(0)}(t) = \hat{H}_R + \sum_g \hat{H}_g^{(aux)}$$

$$\tilde{\hat{H}}^{(h)}(t) = \hat{D}_h(t)\hat{H}^{(h-1)}(t)\hat{D}_h^\dagger(t) + i\hat{D}_h(t)\hat{D}_h^\dagger(t)$$

$$= \tilde{\hat{H}}_w^{(h)} \oplus \tilde{\hat{H}}_{\text{diag}}^{(h)} + O(\epsilon^k)$$

$$\tilde{\hat{H}}_{g,h}^{(aux)}(t) = \hat{D}_h(t)\hat{H}_{g,h}^{(aux)}\hat{D}_h^\dagger(t)$$

where $h$ indexes the order of the frame transformation, $g$ gives the (derivative) order of the auxiliary controls, and $k$ indexes the order of the error for the given frame. Since certain transitions will not correspond to controlled transitions (amongst the original $n$), not all transformations $D_h$ will correspond to an interaction frame and hence $h \geq k$. For the ones that do, we once again have

$$D_h(t) = \exp \left( i \int_0^t \tilde{\hat{H}}_{k,h}^{(aux)}(t')dt' \right)$$

In other cases, one simply has the diagonalization

$$D_h(t) = \exp \left( i \tilde{\hat{H}}_h(t) \right).$$

Then, plugging these into Eq. 4.10 and taking the k-th order expansion of the Baker-Campbell-Hausdorff lemma, the effective $h-$th order Hamiltonians are defined as

$$\tilde{\hat{H}}^{(h)} = \hat{H}^{(h-1)} + i[\hat{H}_{k,h}^T, \hat{H}^{(h-1)}]$$

$$+ \frac{1}{2}[\hat{H}_{k,h}^T, [\hat{H}_{k,h}^T, \hat{H}^{(h-1)}]] + O(\epsilon^k),$$

where $\hat{H}_{k,h}^T = \int_0^t \tilde{\hat{H}}_{k,h}^{(aux)}(t')dt'$, and

$$\tilde{\hat{H}}^{(h)} = \hat{H}^{(h-1)} + i[\tilde{\hat{S}}_h, \hat{H}^{(h-1)}],$$

$$+ \frac{1}{2}[\tilde{\hat{S}}_h, [\tilde{\hat{S}}_h, \hat{H}^{(h-1)}]] + \hat{S}_h + O(\epsilon^k),$$

respectively. The goal is to pick auxiliary controls $\hat{H}_{g,0}^{(aux)}$ such that in the higher-order transformed frame they cancel out with unwanted excitation error (note the lack of derivative in Eq. 4.11, $i\hat{D}_h(t)\hat{D}_h^\dagger(t) = -\tilde{\hat{H}}_{k,h}^{(aux)}$). Similarly, $\tilde{\hat{S}}_h$ can be calculated as is typically done with the Schrieffer-Wolf transformation [44], essentially the diagonalizing operator to the next order (but introducing a higher derivative of the same order of error). As in the other cases, the transformation corresponds to a time-dependent, adiabatic one (now an adiabatic expansion) within the interaction frame given by the auxiliary controls. Note that outside the interaction frame, the evolution is not strictly adiabatic as the states followed by the system are not eigenstates of the complete Hamiltonian as is typically true of the adiabatic theorem. In fact, we may not even be in the adiabatic regime at all as the derivatives may be large. In the following three sections we go through examples for suppressing multiple unwanted transitions, higher order errors, and sideband frequency errors.

V. DISJOINT TRANSITIONS

We demonstrate the formalism for the multiple qubit problem, with the simplification that couplings between the qubits are neglected for our purposes. In this situation, the primary error when driving single qubit rotations will be unwanted crosstalk to other qubits from the external driving field. The lab frame Hamiltonian is given by Eq. 2.2, which we transform to the rotating frame as shown in Sec. IV A to give

$$\hat{H}_R = \Omega(t)e^{-t\phi} \sum_{m=1}^n \lambda_m \hat{\sigma}_m^+ + \text{h.c.}$$

$$+ \sum_{m=1}^n \Delta_m(t) |1\rangle \langle 1|_m$$

with $\Delta_m = \omega_m(t) - \omega_{0,1}^{(j)}$.

A. Exact doublet solution

The simplest scenario is for two qubits for which we can use Eq. 4.2 to remove crosstalk exactly. We put the drive on resonance with the first qubit (detuning away from $\Delta_1(t) = 0$ will be used later to cancel phase error) and let $\Delta_2(t) = \Delta_1(t) + \Delta(t)$. Using the transformation
error term in the adiabatic theorem, which would be

$$D = \left( \cos \left( \int_0^t e^{-i\phi} \text{Im}\Omega(t) \lambda_1 dt \right) \right) \hat{1} + \sin \left( \int_0^t e^{-i\phi} \text{Im}\Omega(t) \lambda_1 dt \right) \hat{\sigma}_1^+ + \text{h.c.} \right) \right) \hat{1}$$

(5.2)

$$\otimes \left( \cos \left( \int_0^t e^{-i\phi} \text{Im}\Omega(t) \lambda_2 dt \right) \right) \hat{1}$$

(5.3)

$$+ \sin \left( \int_0^t e^{-i\phi} \text{Im}\Omega(t) \lambda_2 dt \right) \hat{\sigma}_2^+ + \text{h.c.} \right)$$

(5.4)

and solving Eq. 4.3 for qubit 2 subspace then gives the solution

$$e^{-i\phi} \lambda_2 \int_0^t \text{Im}\Omega(t) dt = \frac{1}{2} \tan^{-1} \left( \frac{2e^{-i\phi} \lambda_2 \text{Re}\Omega(t)}{\Delta} \right)$$

(5.5)

to avoid crosstalk, or equivalently

$$\text{Im}\Omega(t) = \frac{\Delta \text{Re}\Omega(t)}{\Delta^2 + (2\lambda_2 \text{Re}\Omega(t))^2}.$$  

(5.6)

This is the quantum mechanical version of the IBP formula, Eq. 3.7. In particular its first order Taylor expansion is the same, and can often be easier to work with. This solution also bears close resemblance to the diabatic error term in the adiabatic theorem, which would be

$$H_{\text{diab}} = \frac{\lambda_2 e^{-i\phi} \Omega(t)}{\sqrt{\Delta^2 + (2\Omega(t))^2}} \sigma_z^+ + \text{h.c.}$$

in the absence of a perturbation (the denominator is the instantaneous-time energy). The discrepancy occurs on account of the perturbation also introducing diabatic error, and hence needing itself to be corrected.

Introducing an imaginary part of the control will also affect the dynamics of the first qubit. Solving again Eq. 4.3 we get

$$\Delta_1(t) = -2 \text{Re}\Omega(t) \tan \left( 2 \int_0^t \text{Im}\Omega(t') dt' \right)$$

(5.7)

$$\int_0^T \text{Re}\Omega(t) \sec \left( 2 \int_0^t \text{Im}\Omega(t') dt' \right) dt = \theta$$

For Eq. 5.6 with $\lambda_2 = 1$, these take on the simple form

$$\Delta_1(t) = \frac{1}{2} \left( -\Delta + \sqrt{\Delta^2 - (4\text{Re}\Omega(t))^2} \right)$$

(5.8)

$$\int_0^T \frac{\Omega(t)|\Delta|}{\sqrt{\Delta^2(t)^2 + 4(\text{Re}\Omega(t))^2}} dt = \theta$$

To quantify the selection error, we use the phase-insensitive quantum fidelity for a unitary map, given by

$$F = \frac{1}{2^n} \text{Tr} |U^1 V|,$$

(5.9)

where $U$ is the evolution given by the chosen set of controls, and $V$ is the desired evolution, here $\hat{\sigma}_z \otimes^{n-1} \hat{1}$ (with $n = 2$). The selection error $(1 - F)$ is plotted as a function of gate time in Fig. 2. The error for a simple Gaussian of correct area is plotted in dotted blue, while the DRAG solution is in solid red. The Gaussian does suppress the selection error at certain times, where Eqs. 4.3 are all satisfied (or approximately Eqs. 3.2, see Ref. [38]), though clearly the pulses are susceptible to pulse-time miscalibrations. On the other hand, the Gaussian pulse with DRAG correction is an exact solution at all times beyond about $2.5\pi/\Delta$, that is, when $4\Omega(t) < \Delta_2$ is met. At shorter times the detuning (Eq. 5.8) has no real solution, so the first order solution is plotted instead, still outperforming the Gaussian.

**B. General class of solutions**

For more qubits, Eqs. 4.7 must be solved. For two-level systems, the transformation Eq. 4.6 can be parametrized with $X = \sum_j X_j \sigma_j^+ + Y = \sum_j Y_j \sigma_j^+$ such that
formula Eq. 3.6 comes from consolidating the frequency offset with the time-dependent energies of the unper-
turbed qubits, \( \sqrt{\Delta_j^2 + (2\lambda_j\Omega(t))^2} \). Eq. 5.6 gives the exact relation. In the subsections that follow, we solve for explicit forms of the solutions to the systems for 2, 3, and 4 qubits. The result is straightforwardly general-
ized to larger systems, but the results are cumbersome to display.

1. Doublet solutions

To demonstrate this class of solutions with the general form Eq. 5.12, consider again the two-qubit system (Eq.
5.1 with \( n = 2 \)). Since the system contains only one transition we want to cancel, we can solve for it directly.
In addition to the first derivative solution given by Eq. 5.6, a different solution exists for each higher derivative.
For instance, the second derivative solution, for which by construction \( \text{Im}\Omega = 0 \), can be found by setting \( \Omega_j^2 = X \)
in Eq. 5.12, from which

\[
\Omega(t) = \Omega_0 + \frac{d}{dt} \frac{\dot{\Omega}_0}{\Delta^2 + (2\lambda_2 \Omega(t))^2}. \tag{5.14}
\]

The selection error (Eq. 5.9) for this pulse sequence is plotted in dot-dashed orange in Fig. 2, clearly out-
performing the Gaussian result as well as the exact, first-derivative solution in the very short time regime
where the adiabaticity of the eigenstates breaks down, \( \bar{\Omega} > \Delta(t) \). To general order, choosing the recurrence rela-
tion \( X_i = \frac{-1}{\Delta + 4(\text{Re}\Omega(t))^2} \bar{\Delta} i_{i-1} \), gives the (approximate) general real and imaginary solutions

\[
\Omega^R = \frac{d}{dt} \prod_{q=1}^{r/2} \left( \frac{1}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2} \frac{d^2}{dt^2} \right) \frac{\dot{\Omega}_0(t)}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2}, \tag{5.15}
\]

The fourth derivative real solution is

\[
\Omega(t) = \Omega_0(t) + \Omega_0^R(t) \tag{5.15}
\]

\[
= \Omega_0(t) + \frac{d}{dt} \frac{1}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2} \frac{d^2}{dt^2} \frac{\dot{\Omega}_0(t)}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2}. \tag{5.15}
\]

The third derivative complex solution is given by

\[
\Omega_3^C(t) = \frac{\Delta}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2} \frac{d}{dt} \frac{1}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2} \frac{d}{dt} \frac{\dot{\Omega}_0(t)}{\Delta^2 + (2\lambda_2 \Omega_0(t))^2}. \tag{5.16}
\]
The controls being complex, the working transition will also be affected but we can unwind the detuning and rotation error exactly using Eq. 5.8. The selection errors for the two pulse sequences are plotted in Fig. 2 as the dashed green line for the third derivative and the dot-dot-dashed black line for the fourth derivative, still outperforming the Gaussian result in the long-time (adiabatic) limit. The same methodology can be followed to obtain solutions involving even higher order derivatives. It is important to emphasize that in order for the effective frame to be equivalent to the original bare frame we must choose a pulse shape $\Omega_0(t)$ whose derivatives are zero at the endpoints of the pulse (e.g. Eq. 3.3).

2. Triplet solutions

If more than 2 qubits are in the system, with only a single amplitude control it becomes increasingly difficult (at the cost of larger $T$) to find a gate where crosstalk is avoided on all other qubits using a Gaussian pulse. The recurrences of low errors in Fig. 2 roughly every 2 units of time become suppressed or disappear when additional crosstalk qubits are included because the underlying Bohr frequencies become incommensurate. For example, adding a qubit at $\Delta_3 = 1.7\Delta_2$, we see in Fig. 3 that the Gaussian pulse (in dotted blue) no longer performs nearly as well as it did for two qubits (Fig. 2).

Thus, it is all the more useful to apply the higher-derivative formalism to suppress crosstalk for more qubits. The various solutions are plotted in Fig. 3 for the 3 qubit problem. Eqs. 5.13 can be straightforwardly solved by hand for small $n$, while the task is aided by a computer algebra system as the formulae will become cumbersome. For the first and second derivative solution, we get

$$\Omega(t) = \Omega_0(t) + \frac{d}{dt}a_2(t)\dot{\Omega}_0(t),$$

$$= \Omega_0 - i\frac{\Delta_2\Delta_3 (E_2^2 - E_3^2)}{E_2^2 E_3^2 (\Delta_2 - \Delta_3)} \dot{\Omega}_0 + \frac{d}{dt} \frac{\Delta_2 E_2^2 - \Delta_3 E_3^2}{E_2^2 E_3^2 (\Delta_2 - \Delta_3)} \dot{\Omega}_0,$$

with $E_2^2(t) = \Delta_2(t)^2 + 4\Omega_0(t)^2\lambda_2^2$ and $E_3^2(t) = \Delta_3(t)^2 + 4\Omega_0(t)^2\lambda_3^2$, and where the time-dependence has been left implicit after the first line. The selection error (Eq. 5.9) for this pulse is plotted in Fig. 3 as the solid red line. For the first and third derivative we obtain

$$\Omega(t) = \Omega_0 + i\frac{\Delta_2 E_3^4 - E_2^2 E_3^4}{E_2^2 E_3^2 (E_2^2 - E_3^2)} \dot{\Omega}_0 + i\frac{\Delta_2 E_2^2 - \Delta_3 E_3^2}{E_2^2 E_3^2 (E_2^2 - E_3^2)} \ddot{\Omega}_0,$$

which is plotted in as the dot-dashed orange line. Using the second and third derivative the solution is

$$\Omega(t) = \Omega_0 + i\frac{\Delta_2\Delta_3 (E_2^2 - E_3^2)}{E_2^2 E_3^2 (\Delta_2 - \Delta_3)} \dot{\Omega}_0 + \frac{d}{dt} \frac{E_3^4 - E_3^2 \Delta_3}{E_2^2 E_3^2 \Delta_2 - E_2^2 E_3^4 \Delta_3} \dot{\Omega}_0,$$

which is plotted in dashed. Finally, for the second and fourth derivative (real) solution we get

$$\Omega(t) = \Omega_0 + \frac{d}{dt} \frac{1}{E_2^2} \frac{1}{E_3^2} \dot{\Omega}_0$$

$$+ \frac{d}{dt} \frac{1}{E_2^2 E_3^2} \ddot{\Omega}_0,$$

which does not require compensating the driving qubit (using Eq. 5.7) as for the previous pulse. The error for this pulse is plotted in Fig. 3 as the dot-dot-dashed black line. Other such pulses can also be found for other derivative combinations.

3. Quadruplet solutions

For four qubits ($n = 4$), one driven and three affected by crosstalk, Eqs. 5.13 can once be solved using one main control $\Omega_0(t)$ and three auxiliary controls. Here we show the solution to the equations using the first, second, and third derivatives.
\[
\Omega(t) = \Omega_0(t) \pm \frac{E_2^3 (E_3^2 - E_1^2) \Delta_4 \Delta_1 + E_1^4 (E_3^2 - E_1^2) \Delta_2 \Delta_3 + E_3^4 (E_2^2 - E_1^2) \Delta_2 \Delta_4}{E_2^2 E_3^3 E_4^2 ((E_4^2 - E_3^2) \Delta_2 + (E_4^2 - E_3^2) \Delta_3 + (E_4^2 - E_3^2) \Delta_4)} \dot{\Omega}_0(t)
\]
\[+
\frac{d}{dt} \frac{E_2^3 (E_3^2 - E_1^2) \Delta_2 \Delta_3 + E_2^4 (E_3^2 - E_1^2) \Delta_2 \Delta_3 + (E_3^2 E_1^2 - E_2^2 E_1^2) \Delta_1}{E_2^2 E_3^3 E_4^2 ((E_4^2 - E_3^2) \Delta_2 + (E_4^2 - E_3^2) \Delta_3 + (E_4^2 - E_3^2) \Delta_4)} \dot{\Omega}_0(t)\] (5.21)
\[+
\frac{E_4^2 (E_4^2 - E_1^2) \Delta_2 \Delta_4 + E_2^4 (E_3^2 - E_1^2) \Delta_2 \Delta_3 + E_3^4 (E_2^2 - E_1^2) \Delta_2 \Delta_4}{E_2^2 E_3^3 E_4^2 ((E_4^2 - E_3^2) \Delta_2 + (E_4^2 - E_3^2) \Delta_3 + (E_4^2 - E_3^2) \Delta_4)} \dot{\Omega}_0(t).\] (5.22)

Figure 4: Selection error (Eq. 5.9) for four uncoupled qubits for energy differences \( \Delta_2, \Delta_3 = 1.7\Delta_2, \) and \( \Delta_4 = -1.3\Delta_2 \) from the driven qubit. The dotted blue line shows the error using standard Gaussian shaping, Eq. 3.3, while the solid red line is for the pulse shape given by the first-order DRAG solution, Eq. 5.21.

The selection error vs. gate time for this pulse sequence (in solid red) and for the Gaussian (in dotted blue) is plotted in Fig. 4.

C. Discussion

The advantage of using the expansion which incorporates higher order derivatives is that each additional order of derivative allows the removal of (first-order) diabatic error from an unwanted transition, effectively allowing the adiabatic expansion to be taken into account. Thus, what we see in Eq. 5.13 is that each derivative removes some portion of the diabatic error so that in aggregate it is fully removed for each unwanted transition. This is made possible by the fact that a discrete set of conditions is removed using a discrete number of control variables. It is also interesting to go back to look at the case of the continuous excitation spectrum, or one where the exact position of transitions is not known. As in the semi-classical case, the effect of adding in derivative terms is to put a hole at some point in the spectrum, however holes can exist for other reasons, such as the short time window of the pulse leading to a convolution of the Gaussian spectrum with a Sinc function. For the second derivative, the excitation spectrum is shown in Fig. 5. As before (Fig. 1B), adding in the second deriva-

tive can have an effect on the bandwidth, here chosen at the cut-off of 0.1% excitation. Placing the holes appropriately, the DRAG solution (in red) has a bandwidth ~25% narrower than the narrowest Gaussian pulse (of the same duration, starting at 0) with the same cut-off. This is consistent with Ref. [38] where second-order Hermite polynomials are seen to have a similar excitation profile, though here the location of the holes is engineered for the second derivative of the Gaussian with the prefactor \( \frac{\dot{\Omega}}{\Omega} \). We see the main reason for the decreased bandwidth is that the area under the curve is the same as the Gaussian (by conservation of energy), with weight being moved from the high-excitation region to the tails in the low-excitation region. Other results for continuous spectra can be found for other odd or even derivatives, with holes either one or both sides of the centre frequency, respectively. However, using higher derivatives comes at a cost, which is that the trial function (here Eq. 3.3) must have its derivatives begin and end at 0, which effectively increases the bandwidth of the pulse, and so adding more holes is not necessarily beneficial to an engineered continuous spectrum.

VI. CONNECTED TRANSITIONS

Going to higher orders in an expansion of the small parameter may be necessary for high-precision control as needed, e.g., in implementing error-correcting protocols. In particular, when transitions are not disjoint but form a
connected graph between energy levels, then the higher-order effects can be particularly detrimental. The first reason is that errors on the unwanted transitions will no longer commute with the working qubit, and errors generated on the working qubit will typically be more pronounced than were for the other qubits (the largest energy scale is smaller, \( \Omega < \Delta \)). The second reason is that certain resonances may appear between harmonics or sidebands of the pulse frequency and energy differences between non-nearest-neighbour energy levels. Finally, as we move to shorter gate times, the errors will become larger and going to higher orders will be unavoidable.

To demonstrate these detrimental off-resonant effects and their removal we will consider an anharmonic ladder system. These systems are quite common and rather ubiquitous in superconducting qubit systems. In this section we will rederive and then build on the single-transition removal strategies in [27, 29] which were subsequently (first) verified in [34, 35]. We show how higher order effects can be removed in a systematic manner. Note also that in Ref. [45] a combination of disjoint and ladder transitions was studied with an extension of our approach.

The general form of the anharmonic ladder in the rotating frame is the following

\[
\hat{H}^R(t) = \sum_{j=1}^{d-1} (j\delta(t) + \Delta_j) \hat{\Pi}_j + \sum_{j=1}^{d-1} \lambda_{j-1} \Omega(t) \hat{\sigma}_{j-1,j}^+ + \text{h.c.},
\]

with \( \hat{\sigma}_{j-1,j}^+ = e^{-i\phi}|j-1\rangle \langle j| \) and where we assume only one drive at frequency \( \omega_d \approx \omega_{01} \) (using multiple frequencies allows the system to be solved exactly [30]). Next we apply the DRAG formalism, Eqs. 4.10. Going to an interaction frame with respect to the first-order out-of-phase control (here \( \Omega_1^I(t) = -\frac{\text{Re}\Omega(t)}{\Delta} \)) with the transformation

\[
D_1 = \exp\left( -\sum_{j=1}^{d-1} \lambda_{j-1} \int_0^t \Omega_1^I(t') dt' \hat{\sigma}_{j-1,j}^+ \right) - \text{h.c.}
\]

gives the interaction Hamiltonian

\[
\hat{H}^{(1)} = \text{Re}\Omega(t) \left( 1 + \frac{(4 - \lambda^2)(\text{Re}\Omega(t))^2}{2\Delta^2} \right) \hat{\sigma}_{0,1}^+ + \text{h.c.}
+ \left( \delta(t) + \frac{(4 - \lambda^2)(\text{Re}\Omega(t))^2}{\Delta} \right) \hat{\Pi}_1
+ \left( \Delta + 2\delta(t) + \frac{(\lambda^2 + 2)(\text{Re}\Omega(t))^2}{\Delta} \right) \hat{\Pi}_2
+ \frac{\lambda(\text{Re}\Omega(t))^2}{2\Delta} e^{-2i\phi} \hat{\sigma}_{0,2}^+ + \text{h.c.} + O(\epsilon^3),
\]

where \( \hat{\sigma}_{0,2}^+ = e^{-i\phi}|0\rangle \langle 2| \). In this frame, it is easy to see there are three errors associated with the qubit subspace. The selection error is corrected with the off-phase derivative control \( \text{Im}\Omega(t) = -\frac{\dot{\Omega}_1^I(t)}{\Delta} \), which in this first order is again exactly the semiclassical result (Eq. 3.7). The resonance error can be corrected with \( \delta(t) = \frac{(\lambda^2 - 4i\Omega_0(t)^2)}{2\Delta} \) by either shifting the eigen-energies of the system or by a combination of phase ramping and frame compensation (see Appendix A). Finally, this method changes the rotation angle \( \theta \) about the rotation axis \((e^{-i\phi} \hat{\sigma}_{0,1}^+ + \text{h.c.})\), which has to be compensated by enforcing the area law \( \int_{0}^{T} (\Omega_0(t) + \frac{(\lambda^2 - 2i\text{Re}\Omega(t))^2}{\Delta^2}) dt = \theta \).

To avoid higher order commutator errors, it is even better to satisfy the condition at all times by renormalizing \( \text{Re}\Omega(t) = \Omega_0(t) - \frac{(\lambda^2 - 2i\text{Re}\Omega(t))^2}{\Delta^2} \).

In the next order we see that the error comes from the \( \lambda(\text{Re}\Omega(t)\hat{\sigma}_{0,2}^+) \) term. Now a direct control in the \( 0 \rightarrow 2 \) transition is assumed not present in our system. Therefore we cannot remove this excitation error with an interaction frame. Instead we must use a composite transformation. First, we apply the simple next-order diagonalizing transformation

\[
D_2 = \exp\left( \frac{i\lambda\Omega_0(t)}{2\Delta} \hat{\sigma}_{0,2}^+ / \Delta - \text{h.c.} \right)
\]

with the transformed frame now being

\[
\hat{H}^{(2)} = \text{Re}\Omega(t) \left( 1 + \frac{(4 - \lambda^2)(\text{Re}\Omega(t))^2}{2\Delta^2} \right) \hat{\sigma}_{0,1}^+ + \text{h.c.}
+ \left( \lambda(\text{Re}\Omega(t))^2 - \frac{\lambda\Omega_0(t)}{\Delta} - \frac{\lambda\Omega_0(t)}{2\Delta} \right) \hat{\sigma}_{0,2}^+ + \text{h.c.}
+ \left( \delta(t) + \frac{(4 - \lambda^2)(\text{Re}\Omega(t))^2}{\Delta} \right) \hat{\Pi}_1
+ \left( \Delta + 2\delta(t) + \frac{(\lambda^2 + 2)(\text{Re}\Omega(t))^2}{\Delta} \right) \hat{\Pi}_2 + O(\epsilon^3).
\]

but the transformation leaves in the term \( +iD_2(t)\hat{D}_2^I(t) \) term which is of the same order in \( \epsilon \) as the term it diagonalizes. However, in the next order we can obtain a cancelation between the error terms and the next order auxiliary control. Using the transformation

\[
D_3 = \exp\left( \frac{i\lambda\Omega_0(t)}{\Delta^3} e^{-2i\phi} \hat{\sigma}_{0,2}^+ / \Delta - \text{h.c.} \right)
\]

we obtain
Figure 6: For a 3-level system driven by a $T = 4\pi/\Delta$ with pi-pulse, gate error is plotted vs. leakage transition strength $\lambda$ for solutions to the adiabatic expansion to different orders $H^{(k)}$ correcting for errors to order $k$. The solid blue line shows the error for a standard Gaussian, Eq. 3.3. Each line under that gives the error when correcting for the next order in the adiabatic expansion of the control operators. Frame A gives the solutions to different orders when no derivative controls are used. Frame B shows when the 0−1 transition is removed using a first derivative. Frame C is the error plotted when an additional perturbative control is used that includes the second derivative, which enables the removal of the 0−2 transition. How the different lines for each order are calculated is discussed in the text.

\[ \tilde{H}^{(3)} = \text{Re}\Omega(t) \left( 1 + \frac{(4 - \lambda^2)(\text{Re}\Omega(t))^2}{2\Delta^2} \right) \tilde{\sigma}_{0,1} + \text{h.c.} \]

\[ + \left( \frac{\lambda(\text{Re}\Omega(t))^2}{2\Delta} - \frac{\lambda(\tilde{\Omega}_0^3 + \tilde{\Omega}_0\tilde{\Omega}_0)(t)}{\Delta^2} - \frac{\lambda(\tilde{\Omega}_0^2(t))}{2\Delta} \right) \tilde{\sigma}_{0,2} \]

\[ + \text{h.c.} + \left( \delta(t) + \frac{(4 - \lambda^2)(\text{Re}\Omega(t))^2}{\Delta} \right) \tilde{\Pi}_1 \]

\[ + \left( \Delta + 2\delta(t) + \frac{(\lambda^2 + 2)(\text{Re}\Omega(t))^2}{\Delta} \right) \tilde{\Pi}_2 + O(e^3). \]

The most straightforward way to cancel the error is to choose

\[ \text{Re}\Omega(t) = \Omega_R(t) - \frac{(\lambda^2 - 4)(\Omega_R(t))^3}{2\Delta^2} \]

\[ \Omega_R(t) = \sqrt{\frac{\Omega_0^2(t)}{2} + \frac{(\tilde{\Omega}_0^2 + \tilde{\Omega}_0\tilde{\Omega}_0)(t)}{\Delta}} \]

\[ \text{Im}\Omega(t) = \frac{-\Omega_R(t)}{\Delta} \]

\[ \delta(t) = \frac{(\lambda^2 - 4)\Omega_R^2}{\Delta} \]

such that

\[ \tilde{H}^{(3)} = \Omega_R(t) \tilde{\sigma}_{0,1} + \text{h.c.} \]

\[ + \tilde{\Delta}(t)\tilde{\Pi}_2 + O(e^3) \] (6.4)

as required, with $\int_0^{\pi} \Omega_R(t)\,dt = \theta$ enforcing the rotation angle. Note that including the second derivative has allowed us to remove a second undesired transition. The fourth order can be calculated in a similar way by adding additional perturbations to the waveform. However, in the fifth order, we will require using third derivative, which comes from removing the 1−2 transition to the next order (which contains a factor in $\tilde{\Omega}(t)$). Going to higher and higher order, eventually the 0−3 and 1−2 transitions will need to be taken into consideration. To demonstrate the asymptotic bounds derived in Sec. 3.4, we have plotted the result of Schrieffer-Wolff diagonalization to multiple orders when using the standard adiabatic expansion, the interaction frame for the 1−2 transition, and the interaction picture for both the 0−2 and 1−2 transitions. Fig. (6) A, B and C show these, respectively. The gate fidelity is calculated by

\[ F = \frac{1}{2n}|\text{Tr}(U^\dagger V)|^2, \] (6.5)

which is the phase-sensitive version of Eq. 5.9. For each graph, the top blue sensitive line is the zeroth order and corresponds to a Gaussian function with area $\pi$. The dotted red line under it corrects the phase ($\sigma_z$) error on the
qubit and (for B and C) the selection error via the derivative. The dot-dashed yellow line under that corrects the second-order rotation angle (1) error on the qubit. The dashed green line under that corrects error coming from the 0 – 2 transition. The black dashed and solid orange under those correct the next set of errors. It is clear that using the “interaction frame” with respect to the auxiliary control(s) prevents the error from asymptoting as a result of the diabatic error being undiagonalizable as a direct consequence of the IBP formula, Eq. 3.4. In Fig. (6)C, the proper ordering of derivatives is used, and we see that indeed each order qualitatively improves the fidelity compared to the last one.

VII. SELECTIVITY WITH FREQUENCY SIDEBANDS

As a final example of unwanted off-resonant excitation, we consider additional frequency components that can be present in a drive and which we want to suppress. As already mentioned, matrix elements between non-adjacent levels and simultaneous drives to rotate more than one transition will have such frequency sidebands. Perhaps the most ubiquitous occurrence is when applying the rotating-wave approximation to go from Eq. 2.2 to Eq. 2.3 and dropping terms rotating at twice the frequency of the original drive. That is there are also drive elements of the form

\[ \hat{\Gamma}_{j,k}^t = \lambda_{j,k} e^{-i \int_0^t 2\omega_d(t)dt} \Delta_{j,k}(t) \langle j | k \rangle. \]  

(7.1)

To characterize the effect of these and similar terms, we move to the rotating frame, as in Eq. 4.1, only rotating in the opposite direction such that

\[ \hat{H}_R = \Omega(t) e^{-i \phi_i} \left( e^{+i \int_0^t 2\omega_d(t)dt} + 1 \right) \hat{\sigma}^+_{0,1} + \text{h.c.} \]

\[ + 2\omega_d(t) \langle 1 | 1 \rangle \]

where for clarity we only consider one qubit (on resonance). Note the fast-oscillating term is actually the term that is on resonance with the transition and responsible for rotations. The effect of the off-resonant term can be seen by diagonalizing it with

\[ D_1 = \exp \left( \int_0^t \Omega_1^2(t)dt \right) \left( 1 + ce^{+i \int_0^t 2\omega_d(t)dt} \right) \hat{\sigma}^+_{0,1} \text{-h.c.} \]

(with c an arbitrary constant) and choosing \( \Omega_1(t) = -c \frac{\Omega_0}{2\omega_d} \) we get

\[ \hat{H}^{(1)} = \text{Re} \Omega(t) \left( 1 + e^{+i \int_0^t 2\omega_d(t)dt} \right) \hat{\sigma}^+_{0,1} + \right. \]

\[ \text{Re} \Omega(t) \left( \frac{1}{2} - c \right) \left( \Omega_0(t) \right)^2 \left( 1 + e^{+i \int_0^t 2\omega_d(t)dt} \right) \hat{\sigma}^+_{0,1} \]

\[ + i \left( 1 - c \right) \text{Im} \Omega(t) \hat{\sigma}^+_{0,1} + \text{h.c.} \]

\[ + \left( 2\omega_d(t) + \frac{1}{2} - c \right) \left( \Omega_0(t) \right)^2 \left[ \langle 1 | 1 \rangle + O(c^3) \right] \]

(7.2)

This gives the second order solution (with \( c = 1 \))

\[ \Delta_{0,1}(t) = \frac{\Omega_0^2(t)}{4\omega_d} \]

\[ \text{Im} \Omega(t) = \frac{\Omega_0(t)}{2\omega_d} \]

\[ \int_0^T \left( \Omega_0(t) \right)^2 \langle 1 | 1 \rangle + O(c^3) \]

(7.3)

Fig. 7 demonstrates the performance of this strategy. The dotted blue line shows gate error vs. gate time when using a Gaussian pulse. The dot-dashed yellow line shows the improvement when optimizing using only the derivative control and constant drive frequency (\( c = 1/2 \), see Ref. [29]). The solid red line shows the gate error when both the derivative and detuning are applied, Eq. (7.3). In both cases we see the RWA errors are completely suppressed even at short times.

VIII. CONCLUSIONS

We have shown how derivatives of a pulse shape driving the evolution of a quantum system can be used to suppress undesired transitions by introducing auxiliary perturbative controls. In general, multiple off-resonant transition elements can be removed by using higher derivatives. The only caveat to these analytic solutions is phase...
errors need to be corrected (subsequently) and derivatives need to start and end at 0 (effectively increasing bandwidth). To find these solutions to higher orders a pseudo-adiabatic expansion is performed, in terms of time-instantaneous basis functions formed from the derivatives of the trial function. The time-instantaneous nature of the solutions means they can be expressed easily, even at higher orders, and that their form does not change with time, only through the time-dependence of the trial function. In addition to the computational advantage of being independent of time, the solutions to higher order are also advantageous in accuracy in comparison to other analytical techniques for which only low order solutions can be computed, often because the expansions involved are asymptotic. Here, we have shown that the pseudo-adiabatic expansion does not suffer from asymptotic behaviour provided higher order derivatives are used to cancel higher-order diabatic errors.

The analytic pulses also motivate numerical and experimental ansatz solutions to other unwanted off-resonant terms in other (e.g. not fully characterized) physical systems; including higher derivative terms in addition to amplitude and phase modulation of the shaping function, which ensures smooth control pulses, may prove to also be a computationally efficient way to remove undesired terms in a Hamiltonian. The derivative-based approach may also be useful in conjunction with other analytical techniques, such as dynamical decoupling. In cases where both dynamical decoupling (DD) and smooth pulse solutions exist, smooth pulses offer the advantage of compatibility with strongly filtered control lines which would distort hard DD pulses. Also, the total amount of energy transferred to the sample is in general lower for smooth than for DD pulses which is important in cryogenic situations. As was pointed out in Ref. [27], the interaction picture which is used to motivate DRAG is also used as a first step to the Magnus expansion for DD. That is, the derivative solutions (Eq. (3.4)) are exactly the solutions to the decoupled average Hamiltonian. While we choose to use a small-parameter expansion due to the use of a small perturbative control, it is possible another expansion such as the Magnus expansion may be more suitable at short times (where it converges) relative to the inverse of the detuning to the unwanted transition (e.g. for broadband pulses, spin echoes, etc.).

Finally, we have considered three particular classes of physical problems for which derivative removal can be a successful strategy for combatting off-resonant errors. These are multiple off-resonant transitions, multiple higher-order transitions, and sideband transitions. Worked examples were given to illustrate the solutions, which were multi-qubit crossstalk, anharmonic ladder transitions, and compensation for the rotating wave approximation, respectively. In all three cases, including derivatives of the trial function with appropriate prefactors was shown to reduce transfer and gate errors by orders of magnitude relative to conventional pulse shaping techniques.

### Acknowledgments

We thank J. Gambetta, S. Merkel, L. Bishop, and S. Girvin for valuable discussions. This research was supported by NSERC through the discovery grants and QuantumWorks. This research was also funded by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Office. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI, or the U.S. Government.

### I. PHASE COMPENSATION

Using a time-changing drive frequency $\omega_d$ such as in Eq. (2.2) impacts on the choice of the phase reference for computational states (or for sequences of non-commuting operations). That is, the average frequency $\bar{\omega}_d = \int_0^T \frac{\omega_d(t)}{T}$ will differ from the reference given by the nearest qubit, $\omega_q$. This relative phase offset can be compensated by applying a $Z$ operation that undoes the accumulated phase, or, if such an operator is not available, by applying a discretely rotating frame with effective Hamiltonian (after $M$ operations each causing phase offset $U^Z_M$) given by

$$\hat{H}_{\text{eff}}(t) = (U^Z_\theta)^M \hat{H} \hat{U}_\theta^Z M^{-1} Y_b(t) Y_a(t) \hat{X}$$

where $\theta = (\bar{\omega}_d - \omega_q)T$. If, in addition, one cannot change the drive frequency time-dependently, one can alternatively satisfy the requirement with (in the frame rotating at $\bar{\omega}_d$)

$$\hat{H}_{\text{eff}}(t) = \exp \left( i \phi_d(t) \hat{Z} \right) (a(t) \hat{X} + b(t) \hat{Y}) \exp \left( -i \phi_d(t) \hat{Z} \right)$$

where $\phi_d(t) = \int_0^t (\omega_d(t') - \bar{\omega}_d) dt'$ is the ramped phase. These two phase compensation techniques commute and can be applied together.

---

[1] C. Slichter, *Principles of magnetic resonance*, no. 1 in Series in Solid State Sciences (Springer, Berlin, 1996), 3rd ed.
