Designing robust gate implementations for quantum information processing

Jamus Wesenberg
QUANTOP, Danish Research Foundation Center for Quantum Optics,
Department of Physics and Astronomy, University of Aarhus, DK-8000 Århus C, Denmark
(Dated: April 1, 2022)

Quantum information processing systems are often operated through time dependent controls; choosing these controls in a way that makes the resulting operation insensitive to variations in unknown or uncontrollable system parameters is an important prerequisite for obtaining high-fidelity gate operations. In this article we present a numerical method for constructing such robust control sequences for a quite general class of quantum information processing systems. As an application of the method we have designed a robust implementation of a phase-shift operation central to rare earth quantum computing, an ensemble quantum computing system proposed by Ohlsson et. al. [1]. In this case the method has been used to obtain a high degree of insensitivity with respect to differences between ensemble members, but it is equally well suited for quantum computing with a single physical system.

PACS numbers: 02.30.Yy, 03.67.Pp, 32.80.Qk

Introduction

Many potential quantum information processing systems are controlled by means of a set of time-dependent parameters, such as quasi-static electromagnetic fields [2], radio-frequency [3, 4], or optical fields [1, 5, 6, 7]. For most such systems, it is relatively simple to device a set of controls that implement a given evolution in an ideal situation. Often, however, a more careful choice of controls can lead to an implementation that is less sensitive to variations in unknown or uncontrollable system parameters. Examples of such robust implementations include system specific solutions such as the hot gate for ion trap quantum computing which is insensitive to vibrational excitations [8], as well as more general techniques such as composite pulses, a technique originating in NMR spectroscopy [9].

In this article we describe a numerical method for designing robust controls for systems where the evolution is adequately described by a possibly non-unitary evolution operator $U(t)$. This form does not allow a general master equation formulation, but it is sufficient to establish worst case behavior in many quantum computing settings where the worst case effects of decoherence and loss can be adequately modeled by a Schrödinger equation with a non-Hermitian Hamiltonian.

As an application of the method, we will consider the construction of a robust phase shift operation for the rare earth quantum computing (REQC) system [1, 10], which is based on rare-earth ions embedded in a cryogenic crystal. In each ion, two metastable ground-state hyperfine levels, labeled $|0\rangle$ and $|1\rangle$, serve as a qubit register which is manipulated via optical transitions from both states to an inhomogeneously shifted excited state $|e\rangle$. The REQC system is an ensemble quantum computing system and macroscopic numbers of ions are manipulated in parallel, addressed by the value of the inhomogeneous shift of their $|e\rangle$-state. To obtain a sufficient number of ions within each frequency channel, it is necessary to operate on all ions within a finite range of inhomogeneous shifts, and the main difficulty in operating the REQC system is to achieve the same evolution for each of these ions independent of their particular inhomogeneous shift.

This article is divided into two sections: in section I we describe the method we have used to design robust gate implementations; these results should be applicable to a variety of quantum information processing systems. In section II we present the results of applying the method to a specific problem relating to the REQC system, and show that it is indeed possible to obtain very high degrees of robustness.

I. DESIGNING ROBUST GATE IMPLEMENTATIONS

We consider a collection of quantum systems which evolve according to a set of time-dependent controls $\varepsilon(t)$. In addition to $\varepsilon$, the single system Hamiltonian $H(\varepsilon, \varepsilon(t))$ depends on a system specific set of uncontrollable or unknown parameters $\xi$, such as field strength or quantum numbers corresponding to unused degrees of freedom.

The evolution of each system is governed by the Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} U(t) = H(\xi, \varepsilon(t)) U(t), \quad U(0) = 1,$$

where we will allow the Hamiltonian $H$ to include non-Hermitian terms describing loss and decoherence.

Our goal is to choose a set of controls that lead to an evolution $U(T)$ which is as close as possible to a given desired evolution $U_0$ over a range of $\varepsilon$-values. To quantify this, we introduce an objective functional $J(\xi, \varepsilon)$ which describes the performance of a set of controls $\varepsilon$ for a given value of $\xi$. By convention we take a low value of $J$ to indicate a good performance, and the problem of finding a
A robust set of controls thus corresponds to minimizing

$$J_X(\varepsilon) = \max_{\xi \in X} J(\xi, \varepsilon),$$  \hfill (2)

where $X$ is the set of $\xi$-values for which we want the implementation to perform well. The conceptually simple approach we have taken to this problem is to replace $X$ with a discrete subset $X' \subset X$, so that the minimization of $J_X$ has the form of a standard minimax problem, which can be solved efficiently provided that we are able to calculate $\partial J_X/\partial \varepsilon$. Below we show how to achieve this by methods from optimal control theory.

### A. Calculating $\partial J/\partial \varepsilon$

In this section we show how $\partial J/\partial \varepsilon$ may be calculated for a quite general class of objective functionals. To keep the notation simple and avoid unnecessary restrictions, we will consider the following generalization of the Schrödinger equation $\mathbf{U}$,

$$\dot{\mathbf{U}}(t) = \mathbf{f}(\mathbf{U}(t), \varepsilon(t)), \quad \mathbf{U}(0) = \mathbf{1},$$  \hfill (3)

determining the evolution of a complex-valued, time-dependent matrix $\mathbf{U}$ due to a set of real-valued, time-dependent controls, $\varepsilon(t)$. We will consider objective functionals of the form

$$J(\varepsilon) = \phi(\mathbf{U}(T)) + \int_0^T l(\mathbf{U}(t), \varepsilon(t))dt,$$  \hfill (4)

where $\phi(\mathbf{U}(T))$ is a real-valued function quantifying how close the final state $\mathbf{U}(T)$ is to our goal, and the real-valued function $l(\mathbf{U}(t), \varepsilon(t))$, referred to as a penalty function, can be chosen to discourage the use of certain control values. Our goal is to calculate $\partial J/\partial \varepsilon$ subject to the constraint that $\varepsilon$ and $\mathbf{U}$ obey Eq. (3). To achieve this we introduce the modified objective functional:

$$J' = J - \int_0^T \operatorname{Tr} \left( \Lambda^\dagger \left[ \dot{\mathbf{U}} - \mathbf{f}(\mathbf{U}, \varepsilon) \right] + \text{h.c.} \right) dt,$$  \hfill (5)

where the complex time-dependent adjoint state matrix $\Lambda$ is in effect a continuous set of Lagrange multipliers leaving $J'$ identical to $J$, provided that $\mathbf{U}$ and $\varepsilon$ obey Eq. (3). If we require $\Lambda$ to obey the adjoint equations

$$\dot{\Lambda}^\dagger = -\frac{\partial h}{\partial \mathbf{U}}, \quad \Lambda^\dagger(T) = \frac{\partial \phi}{\partial \mathbf{U}} \bigg|_T,$$  \hfill (6)

where $\mathbf{U}$ and $\mathbf{U}^\dagger$ should be considered as independent with respect to the partial derivative and $h$ is defined as

$$h(\varepsilon, \mathbf{U}, \Lambda) = l(\mathbf{U}, \varepsilon) + \operatorname{Tr} \left( \Lambda^\dagger \mathbf{f}(\mathbf{U}, \varepsilon) + \text{h.c.} \right),$$  \hfill (7)

we find by integration by parts that the differential of $J'$ is given by

$$dJ' = \int_0^T \frac{\partial h}{\partial \varepsilon} \delta \varepsilon(t) dt,$$  \hfill (8)

from which the derivatives of $J$ with respect to the parameters used to parametrize $\varepsilon$ can be calculated by the chain rule.

We now return to the case of a quantum system governed by the Schrödinger equation $\mathbf{U}$. If we assume the penalty function $l$ to be independent of $\mathbf{U}$, the adjoint equations in this case are

$$ih_0 \frac{\partial}{\partial t} \Lambda(t) = \mathbf{H}\Lambda(t), \quad \Lambda(T) = \frac{\partial \phi}{\partial \mathbf{U}} \bigg|_T,$$  \hfill (9)

and $dJ$ is given by (8) with

$$\frac{\partial h}{\partial \varepsilon} = \frac{\partial l}{\partial \varepsilon} + \frac{2}{\hbar} \operatorname{Im} \operatorname{Tr} \left( \Lambda^\dagger \frac{\partial \mathbf{H}}{\partial \varepsilon} \mathbf{U} \right).$$  \hfill (10)

The role of the adjoint state and the adjoint equations is often described as back-propagating the errors in achieving the desired final state. If $\mathbf{H}$ is Hermitian, the boundary value for $\Lambda$ can be optimized for numerical computation as shown in Appendix A.

### B. Fidelity of quantum evolution

We will now discuss the choice of the function $\phi$, quantifying how well the obtained evolution $\mathbf{U}(T)$ approximates $\mathbf{U}_0$. As we are concerned with quantum information processing we will assume that all operations start out with an unknown state in the qubit subspace $\mathcal{H}_Q$ of the full system Hilbert space $\mathcal{H}$, and that this subspace is invariant under the ideal evolution $\mathbf{U}_0$. The function $\phi$ should not depend on the evolution of states outside $\mathcal{H}_Q$, nor on collective phases on the states originating in $\mathcal{H}_Q$. A cautious choice of $\phi$ fulfilling these conditions could be based on the worst case overlap fidelity $\mathcal{F}$,

$$\mathcal{F} = \min_{|\psi\rangle \in \mathcal{H}_Q} \left| \left| \langle \psi | \mathbf{U}_0 \mathbf{U}(T) | \psi \rangle \right| \right|_2,$$  \hfill (11)

which measures the least possible overlap between the obtained output state $\mathbf{U}(T) |\psi\rangle$ and the ideal output $\mathbf{U}_0 |\psi\rangle$ for initial states in $\mathcal{H}_Q$. This fidelity measure has the desirable quality that both population transfer from $\mathcal{H}_Q$ to $\mathcal{H}_Q^\perp$ and population transfer completely out of $\mathcal{H}$, as described by a non-unitary evolution, is counted as loss of fidelity.

From the point of view of optimal control, a significant drawback of the worst case overlap fidelity, $\mathcal{F}$, is that it is computationally complicated $\mathcal{F}_Q$. A computationally accessible fidelity measure which share many appealing features with $\mathcal{F}$ is the trace fidelity $\mathcal{T}$,

$$\mathcal{T} = \frac{1}{n} \left| \operatorname{Tr} \left( \mathbf{U}_0^\dagger \mathbf{U} \right) \right|,$$  \hfill (12)

where $n = \dim(\mathcal{H}_Q)$. As shown in Appendix A $\mathcal{T}$ is related to $\mathcal{F}$ by the strict bound

$$1 - \mathcal{F} \leq n(1 - \mathcal{T}),$$  \hfill (13)
indicating that we can safely replace $F$ by $T$ for numerical computations on a few qubits at high fidelity.

For numerical calculations it is beneficial to use $T^2$ rather than $T$ \[13\]; in the calculations presented in the next section we have used $\phi = 1 - T^2$, leading to an adjoint state boundary condition of

$$
\Lambda(T) = \left. \frac{\partial \phi}{\partial U^\dagger} \right|_T = -\frac{1}{n^2} U_0 \text{Tr}(U_0^\dagger U(T)),
$$

which can be directly computed.

C. Minimization algorithms

One approach to minimizing $J$ is to directly solve the extremum condition $\partial J/\partial \epsilon = 0$ for $\epsilon$. This task is significantly simplified if a penalty function proportional to $\epsilon^2$ is introduced, $h = \lambda \epsilon^2 + h_0$, so that the extremum condition according to \[5\] reads,

$$
\epsilon(t) = -\frac{1}{2\lambda} \frac{\partial h_0}{\partial \epsilon},
$$

which may be used as an iterative formula for calculating $\epsilon$. Variations over this iterative approach give rise to the Krotov \[16\] and Zhu-Rabitz \[17, 18\] algorithms which have been shown to have excellent convergence properties, and have been successfully applied to optimal control of unitary transformations for one set of parameters by the group of R. Kosloff \[12, 16\]. A unifying view of these direct methods can be found in Ref. \[17\].

In the work presented here we have chosen to use an indirect minimization algorithm: rather than trying to solve the extremum condition directly, we have used the gradient information obtained through Eq. \[8\] as input for a general sequential quadratic programming procedure based on a constrained quasi-Newton method \[20, 21, 22\]. The primary advantage of this approach is that we have total freedom to choose the parametrization of the controls, and can place arbitrary bounds on these. This allows us to more accurately model the fact that the experimental limitations most often only distinguish between possible and impossible controls: no possible controls are significantly harder than others. An explicit field strength limit also serves to introduce an absolute scale on which to introduce decay strengths etc.

II. APPLICATION TO RARE EARTH QUANTUM COMPUTING

The motivation for the work presented in this article has been the design of robust gate implementations for the REQC system mentioned in the introduction. As an example, we will consider the construction of a robust implementation of the single qubit operation

$$
U_0 = |1\rangle \langle 1| - |0\rangle \langle 0|,
$$

which could simply be implemented by a single $2\pi$ pulse on the $|0\rangle$-$|e\rangle$ transition if we were not concerned with robustness. Our primary concern will be to make the implementation robust with respect to variations in the inhomogeneous shift $\delta$ of the $|e\rangle$-state in order to allow the use of finite with channels. Since it is experimentally difficult to obtain a homogeneous field strength over the crystal, we would also prefer the implementation to be insensitive to variations in the relative field strength, which we will denote $\gamma$.

In addition to requiring the implementation of $U_0$ to be robust with respect to variations in $\delta$ and $\gamma$, we will add the requirement that ions outside the channels should not be affected, as this allows us to use the obtained implementation of $U_0$ as a part of a controlled phase shift operation \[11, 12\]. If the $|e\rangle$-states of the qubit ion and a controlling ion are coupled sufficiently strongly by static dipole interaction, an excitation of the controlling ion will effectively shift the qubit ion out of the channel thus conditioning the evolution of the qubit on the state of the controlling ion.

Even though the simplest implementation of $U_0$ for an ideal ion with $\delta = 0$ and $\gamma = 1$ would involve only the $|0\rangle$ and $|e\rangle$ levels, a robust implementation must also involve the $|1\rangle$-state since the coupling of $|0\rangle$ to $|e\rangle$ will result in a $\delta$-dependent phase on $|0\rangle$ which can only be compensated by introducing the same phase on the $|1\rangle$-state, e.g. through phase compensating rotations \[13\]. A highly successful example of this approach is the sech-pulse sequence suggested by Roos and Molmer \[23\], which as illustrated by Fig. \[1a\] yields a very robust implementation, achieving high fidelities over a wide range of parameter values.

We model the REQC system by the single ion Hamiltonian

$$
H = -\delta |e\rangle \langle e| + \frac{\gamma}{2} \sum_{i=0,1} (\Omega_i(t) |e\rangle \langle i| + \text{h.c.}),
$$

which does not include any effects of decay or decoherence. In the notation introduced in section \[11\] the system parameters are $\xi = (\gamma, \delta)$, the controls are $\epsilon = (\Omega_0, \Omega_1)$, and the qubit subspace is $H_Q = \{|0\rangle, |1\rangle\}$. No penalty function is used: we use $J = \phi = 1 - T^2$, and limit the field by strict bounds on $\Omega_i(t)$, as this is the relevant limiting parameter in the REQC system. Inspired by the success of the sech-based solution and the hat-like Fourier spectrum of the sech-pulse, $\Omega_i(t)$ is parametrized in terms of a truncated Fourier basis. Based on trial and error we have arrived at 49 $\xi$-values to constitute $X'$, some within the neighborhood of $\delta = 0$, $\gamma = 1$, and some at large detunings where the ions should not be disturbed. The result of the optimization with this choice of $X'$ is shown in Fig. \[1b\], where the circles indicate the members of $X'$. It is evident from the plot that the optimization has achieved a high fidelity over an even larger range of parameters than the sech-pulse sequence illustrated in Fig. \[1a\].

With respect to not disturbing the detuned ions, both
FIG. 1: The worst case overlap fidelity of $U_0$ as given by Eq. (16) implemented (a) by a series of sech pulses as suggested by Roos and Mølmer [23], and (b) by an optimized pulse. The fidelity is plotted as a function of relative field strength $\gamma$ and inhomogeneous shift $\delta$ relative to the maximal resonant Rabi frequency $\Omega_0$. The duration of both pulses is $24\pi/\Omega_0$, and the circles in (b) indicates the parameter set $X'$ used during the optimization.

FIG. 2: The effect of the sech-pulse sequence (dashed line) and the optimized (solid line) implementation of $U_0$ on far-detuned ions, as described by $1 - F$, where $F$ is calculated with respect to the identity, as these ions should ideally not be disturbed. Both implementations achieve worst case fidelities very close to unity, with the sech-pulse sequence achieving the best results. The curve for the optimized pulse is actually a running maximum, as the actual value of $1 - F$ oscillates with a period of $2\pi/T$.

III. CONCLUSIONS AND OUTLOOK

We have shown that it is possible to construct highly robust gate implementations for quantum information processing by a quite general method. In particular, the method has been used to greatly enhance the performance of a gate implementation for a model REQC system by extending the range of inhomogeneous shifts and relative field strengths over which an acceptable performance is achieved.

The model REQC system used in this article ignores many performance degrading factors, the two most important being decoherence and implementation noise. Decoherence could in the present case be adequately modeled by a non-Hermitian Hamiltonian, for which we expect the method described in this paper to be able to find a robust implementation as in the decoherence-free case. It is not clear, however, how the method could be extended to address the problem of robustness with respect to implementation imperfections.

Acknowledgments

The author would like to thank the people at the Centre for Quantum Computer Technology at the University of Queensland for their hospitality, and Klaus Mølmer for valuable comments on the manuscript. This research was funded by project ESQUIRE of the IST-FET programme of the EC.
APPENDIX A: THE TRACE FIDELITY

In this section we prove the relation (13) between the worst case overlap fidelity, $\mathcal{F}$, and the trace fidelity $T$. Referring to the definition (11), we note that $\mathcal{F}$ is completely determined by the restriction $\mathcal{O}$ of the operator $U_0U(T)$ to $\mathcal{H}_Q$. Since $U(T)$ describes the evolution of a quantum system, it is possible to extend it to a unitary operation on a Hilbert space containing $\mathcal{H}$, and $\mathcal{O}$ is consequently the restriction of a unitary operator to $\mathcal{H}_Q$. In the ideal case, consequently the restriction of a unitary operator to $\mathcal{H}_Q$, perhaps with the exception of a complex phase.

$\mathcal{F}$ is defined as the minimum of the overlap $|\langle \psi | \mathcal{O} | \psi \rangle |$. Since the unit sphere of $\mathbb{C}^n$ is compact, this minimum will be attained for some $|\psi_0\rangle$: $\mathcal{F} = |\langle \psi_0 | \mathcal{O} | \psi_0 \rangle |$. We now extend $\{|\psi_0\rangle\}$ to an orthonormal basis $\{|\psi_k\rangle\}_{k=0,...,n-1}$ by the Gram-Schmidt process. Evaluating the trace fidelity in this basis we find by the triangle inequality:

$$T \leq \sum_{k=0}^{n-1} |\langle \psi_k | \mathcal{O} | \psi_k \rangle|$$

which we have used that $|\langle \psi | \mathcal{O} | \psi \rangle| \leq 1$ for all $|\psi\rangle$ since $\mathcal{O}$ is the restriction of a unitary operator. By rewriting (11) we obtain the desired relation, $1 - \mathcal{F} \leq n(1 - T)$.

We note that the established bound is strict in the sense that for any $0 \leq \mathcal{F}_0 \leq 1$, the operator $\mathcal{O} = 1 - (1 - \mathcal{F}_0) |\psi\rangle \langle \psi|$ will fulfill Eq. (13) with equality for any $|\psi\rangle$.

APPENDIX B: OPTIMIZED ADJOINT STATE BOUNDARY CONDITION

In the case of a Hermitian Hamiltonian and a penalty function $l$, that does not depend on the state $U$, it is possible to modify the adjoint state boundary condition (14) to reduce the required accuracy of the adjoint state propagation.

In this case, we find according to Eqs. (8) and (10) that $dJ$ has the form

$$dJ = \int_0^T \left[ \frac{\partial l}{\partial \varepsilon} + 2 \Im \left( \sum_k \lambda_k \frac{\partial H}{\partial x_k} \right) \right] \delta(t) \, dt,$$

where $x_k$ and $\lambda_k$ denotes the $k$-th columns of $U$ and $A$ respectively. Since $H$ and thus $\partial H/\partial \varepsilon$ are assumed to be Hermitian, $dJ$ as given by (B1) is not affected by adding to $\lambda_k$ a component of $\alpha_k x_k$ for any real $\alpha_k$. Furthermore, since $\lambda_k$ and $x_k$ evolve according to the same Schrödinger equation, this corresponds to modifying the boundary condition for the adjoint state to read:

$$\lambda_k(T) = \frac{\partial \phi}{\partial x_k^T} + \alpha_k x_k(T),$$

for any real $\alpha_k$.

The obvious use of the freedom in the choice of boundary value is to minimize the norm of the adjoint state, in order to relax the requirements of the relative precision of the adjoint state propagation. This minimum is easily calculated from (B2), but we prefer to illustrate the physical background of the result by calculating it in a different way: The freedom in the choice of boundary value (B2) is allowed by the Hermiticity of $H$. The same Hermiticity ensures that $x_k$ is normalized, so that $\phi \circ n(x_k)$, where $n$ is the normalization function $n(v) = v/|v|$, is equal to $\phi(x_k)$. The gradient of $\phi = \phi \circ n$, however, is different from that of $\phi$. In fact we find

$$\frac{\partial \phi}{\partial x_k^T} = \frac{\partial \phi}{\partial x_k} - \text{Re} \left[ x_k^\dagger \frac{\partial \phi}{\partial x_k^T} \right] x_k,$$

where $x_k$ and $x_k^\dagger$ should be considered independent with respect to the derivative. Comparing this expression to Eq. (B2), it is tempting to let $\alpha_k = \text{Re} \left[ x_k^\dagger \frac{\partial \phi}{\partial x_k^T} \right] x_k$, which is indeed the answer found by minimizing $|x_k(T)|$ subject to Eq. (B2).

The modified boundary condition of Eq. (B3) carries the same error information as that of Eq. (14), but the required relative numerical precision when propagating the adjoint state will be significantly reduced.

[1] N. Ohlsen, R. K. Mohan, and S. Kröll, Opt. Comm. 201, 71 (2002).
[2] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, Nature 398, 786 (1999).
[3] D. G. Cory, A. F. Fahmy, and T. F. Havel, Proc. Natl. Acad. Sci. 94, 1634 (1997).
[4] B. E. Kane, Nature 393, 133 (1998).
[5] J. I. Cirac and P. Zoller, Phys. Rev. Lett. 74, 4091 (1995).
[6] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. 83, 4204 (1999).
[7] M. D. Lukin and P. R. Hemmer, Phys. Rev. Lett. 84, 2818 (2000).
[8] A. Sørensen and K. Mølmer, Phys. Rev. Lett. 82, 1971 (1999).
[9] H. K. Cummins, G. Llewellyn, and J. A. Jones, Phys. Rev. A 67 (2003).
[10] J. J. Longdell and M. J. Sellars (2002), quant-ph/0208182.
[11] D. G. Luenberger, Introduction to Dynamic Systems (John Wiley & Sons, 1979).
[12] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2000).
[13] J. Wesenberg and K. Mølmer, Phys. Rev. A 68, 012320 (2003).
[14] J. P. Palao and R. Kosloff, Phys. Rev. Lett. 89, 188301 (2002).
[15] J. P. Palao and R. Kosloff (2003), quant-ph/0309011.
[16] D. J. Tannor, V. A. Kazakov, and V. Orlov, in Time-Dependent Quantum Molecular Dynamics, edited by J. Broeckhove and L. Lathouwers (Plenum, New York, 1992), pp. 347–360.
[17] W. Zhu and H. Rabitz, Phys. Rev. A 58, 4741 (1998).
[18] W. Zhu and H. Rabitz, Jour. Chem. Phys. 109, 385 (1998).
[19] Y. Maday and G. Turinici, J. Chem. Phys. 118, 8191 (2003).
[20] R. Fletcher, Practical Methods of Optimization (John Wiley & Sons, 1987), 2nd ed.
[21] P. E. Gill, W. Murray, and M. H. Wright, Practical Optimization (Academic Press, 1981).
[22] C.-J. Lin and J. J. Moré, SIAM Journal on Optimization 9, 1100 (1999).
[23] I. Roos and K. Mølmer (2003), quant-ph/0305060.