Optimal Control of Inhomogeneous Ensembles

Justin Ruths and Jr-Shin Li, *Member, IEEE*

Abstract—**Inhomogeneity, in its many forms, appears frequently in practical physical systems. Readily apparent in quantum systems, inhomogeneity is caused by hardware imperfections, measurement inaccuracies, and environmental variations, and subsequently limits the performance and efficiency achievable in current experiments. In this paper, we provide a systematic methodology to mathematically characterize and optimally manipulate inhomogeneous ensembles with concepts taken from ensemble control. In particular, we develop a computational method to solve practical quantum pulse design problems cast as optimal ensemble control problems, based on multidimensional pseudospectral approximations. We motivate the utility of this method by designing pulses for both standard and novel applications. We also show the convergence of the pseudospectral method for optimal ensemble control. The concepts developed here are applicable beyond quantum control, such as to neuron systems, and furthermore to systems with by parameter uncertainty, which pervade all areas of science and engineering.**

Index Terms—**Convergence of numerical methods, optimal control, robust control.**

I. INTRODUCTION

RECENT advancements in quantum research have enabled breakthroughs in biology, chemistry, physics, engineering, and medicine including better methods to understand the structure of macromolecules used in biochemical signaling and drug delivery, to facilitate the fast and efficient storage of information, and to yield higher resolution medical images for diagnosis and treatment of early stage cancer [1]–[3]. Most, if not all, measurements and manipulations of quantum systems are achieved through the appropriate design of externally applied time-varying electromagnetic pulses, or controls [1]. These pulses guide the system to produce a desired time-evolution or a specific terminal state. The design of such pulses is made significantly more difficult by inherent variations within the systems of interest. Inhomogeneity is one of the fundamental obstacles for the practical implementation and physical realization of quantum science and quantum technology. In classical systems the dispersions resulting from inhomogeneity is often compensated for by feedback control. Significant research effort has been employed in the area of quantum feedback control with several promising theoretical

J. Ruths is with the Pillar of Engineering Systems and Design, Singapore University of Technology and Design, Singapore, 138682 (e-mail: justin ruths@sutd.edu.sg).

J.-S. Li is with the Department of Electrical and Systems Engineering, Washington University, Saint Louis, MO 63112 USA (e-mail: jsli@seas.wustl.edu).

and practical discoveries in recent years [4], [5]. There is still a large portion of quantum systems for which state feedback is either impractical or difficult to achieve due to the short timescales and large state-space of quantum phenomena. These limitations motivate us to consider the open-loop synthesis of optimal pulses that achieve a desired goal while compensating for the inhomogeneity present in the quantum ensemble.

The behavior of a bulk quantum *system* is the aggregate behavior of a large ensemble of individual quantum *systems*. Although in isolation these individual systems, e.g., atoms, spins, qubits, etc., are fundamentally identical, in a physical system they are distinct due to different chemical and electromagnetic environments. This variation across the ensemble exhibits itself at the macroscopic level as variation in the values of parameters that characterize the dynamics of the bulk quantum system [6]. For example, adjacent atoms within the same macromolecule shield the full strength of the applied external magnetic fields. Varied levels of shielding create a dispersion in the frequency of the quantum spins, which is observed as inhomogeneity in the value of the natural frequency of the bulk sample. In addition, hardware imperfection causes attenuation in the applied electromagnetic pulse over the ensemble and can be represented as variation in a scaling factor multiplying the applied pulse. Often several pulses are applied in sequence in order to achieve an intricate time-evolution of the system [7]. Each pulse is designed assuming an exact (usually uniform) initial system state, however, in practice, the previous pulses only prepare the system to within a neighborhood of the assumed initial state. The additive error in such a pulse sequence can cause significant performance degradation.

Guiding the evolution of inhomogeneous ensembles is a central idea in the design and implementation of quantum experiments. As such, there is a rich literature of methods addressing this class of challenging problems. Initially these were intuitive or *ad-hoc* methods motivated by the symmetry of the state space [7], [8], which were then augmented with various heuristic and specifically designed techniques [9], [10]. More recently pulse design problems have been cast as optimal control problems [11]–[15]. Here we present a methodology that addresses the difficulties of the current methods and is easily generalizable to any inhomogeneous ensemble or uncertain system. The proposed method has both theoretical, such as convergence rates and computational complexity, and practical, such as ease of implementation and computation time, amd advantages.

In this paper, we describe a framework to pose robust quantum pulse design problems in the language of mathematical control theory with support from new theoretical concepts in ensemble control [16]–[18] and computational advances in multidimensional pseudospectral methods adapted for ensemble systems [6], [19]. In a larger context, we provide a rigorous methodology to study and control inhomogeneous

ensembles or systems with parameter uncertainty from any area or application. In the following section, we introduce the problem statement as well as our theoretical and computational tools. In Section III, we take several examples from nuclear magnetic resonance (NMR) in liquids modeled by the bilinear Bloch equations, including broadband excitation in the presence of inhomogeneity, a sequence of broadband pulses robust to variation in the initial conditions, and systems with a time-varying frequency. We then provide empirical and theoretical justifications that the solutions computed using the pseudospectral method converge to solution of the original optimal control problem.

II. MOTIVATION & THEORY

In this section, we review the underlying concepts involved in our approach to design robust quantum pulses as well as the broader mathematical formulations necessary to characterize and solve such design problems. In what follows, we present a highly general model of quantum dynamics, which demonstrates the abundance of inhomogeneity in these problems and motivates studying the control of a family of parameterized systems. We then show how the notion of ensemble control is well suited for dealing with the inherent variation and uncertainty in practical quantum systems and formulate a new type of optimal control problem based on ensemble control.

A. Quantum Dynamics and Pulse Design

The dynamics of a quantum system is given by the time-evolution of its density matrix. We consider here general dynamics in which the system may have interaction with the environment that leads to dissipation in the system state. Under the Markovian approximation, where the environment is modeled as an infinite thermostat which has constant state, the evolution of the density matrix can be written in Lindblad form in terms of the system Hamiltonian $H(t)$ and superoperator $L(\cdot)$ which model the unitary and nonunitary dynamics [20], respectively,

$$\frac{d}{dt}\rho = -i[H(t),\rho] - L(\rho) \quad (\hbar = 1).$$

The expression of the Hamiltonian has components corresponding to free evolution Hamiltonian, H_f, and the control Hamiltonians H_i,

$$H(t) = H_f + \sum_{i=1}^{m} u_i(t)H_i$$

where $u_i(t)$ are externally applied electromagnetic pulses that can be used to manipulate, or guide, the evolution of the system state. Typical pulse design problems involve designing these pulses, or controls, to bring the final state of the density matrix $\rho(T)$ as close as possible to a target operator. This problem can be transformed, by taking the expectation values of the operators involved in the state transfer, to the vector-valued, bilinear control problem, $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ given by

$$\frac{d}{dt}x = \left[\mathcal{H}_d + \sum_{i=1}^{m} u_i(t)\mathcal{H}_i\right]x \quad (1)$$

where $\mathcal{H}_d \in \mathbb{R}^{n \times n}$ corresponds to the drift evolution representing H_f and L, $\mathcal{H}_i \in \mathbb{R}^{n \times n}$ corresponds to the controlled evolution representing H_i, and $t \in [0, T]$ [21]. While (1) accurately represents the classical interaction of magnetic fields, in practice the effective fields—and, therefore, the matrices representing the Hamiltonians $\mathcal{H}_d$ and $\mathcal{H}_i$—show variation in magnitude due to different chemical environments and equipment errors. The system can no longer be described by a single equation but rather by a family of equations with variation in the parameters that characterize the motion, which motivates us to consider the dispersion in the dynamics as a continuum parameterized by the system values

$$\frac{d}{dt}x(t,s) = \left[\mathcal{H}_d(s) + \sum_{i=1}^{m} u_i(t)\mathcal{H}_i(s)\right]x(t,s) \quad (2)$$

where $s \in D \subset \mathbb{R}^d$ is a d-dimensional interval representing the d parameters exhibiting variation [19]. In a more general formulation the matrices representing the Hamiltonians can be time-dependent, $\mathcal{H}_d = \mathcal{H}_d(t,s)$ and $\mathcal{H}_i = \mathcal{H}_i(t,s)$, as in the case of random fluctuations.

Designing a single set of controls (pulses) $u_i(t)$ that simultaneously steer an ensemble of dispersive systems in (2) from an initial state to a desired final state is a fundamental problem in the control of quantum systems. Moreover, similar parameterized structures can be found across all areas of science and engineering, such as in neuroscience where a single stimulus is used to trigger a simultaneous firing of neuron oscillators with distinct oscillation frequencies [22]. In these applications full state feedback, which is required in most current methods to compensate for system uncertainty, is impractical to obtain due to the sheer number of members (and states) within the ensemble. Averaged measurement is possible in some applications; however, this type of measurement restricts the forms of available feedback. It is, then, of particular importance to consider the corresponding open-loop control problem.

B. Optimal Ensemble Control

Systems as in (2) motivate the study of a new class of inhomogeneous control systems. Ensemble control [17] is a mathematical framework to characterize parameterized systems of the form

$$\frac{d}{dt}x(t,s) = F\big(t,s,x(t,s),u(t)\big), \quad x(0,s) = x_0(s) \quad (3)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $s \in D \subset \mathbb{R}^d$, with F, and $x_0(s)$ smooth functions of their respective arguments. The significant challenge of this class of control problems originates from requiring the same open-loop control, $u(t)$ to guide the continuum of systems from an initial distribution, $x_0(s)$, to a desired final distribution, over the corresponding function space. Fundamental properties of these systems, such as controllability, are of particular interest—specifically addressing what types of inhomogeneities can be compensated for robustly. For example, it has been shown that the controllability of an ensemble of bilinear Bloch equations, used as a sample system in this paper (see Section III), corresponds to the synthesis of appropriate polynomials [16] and controllability conditions for

an ensemble of time-varying linear systems are related to the Picard criterion of Fredholm integral equations of the first kind [17]. These concepts have also been approached with functional analysis techniques from the study of partial differential equations and have investigated, in particular, the exact and approximate ensemble controllability of the Bloch equations with bounded and unbounded controls, finite and infinite time, as well as finite and infinite parameter domains [23].

Subsequently, given dynamics and initial and final distributions, we seek methods to construct controls for such steering problems. As with any control problem, in general there may be many possible solutions that satisfy a state-to-state ensemble control problem. In addition there are often benefits, penalties, and limitations that are associated with the physical system, which can be used to rank the different solutions. Such practical considerations lead to considering an optimal control problem based on the ensemble dynamics in (3) which includes a cost functional (with terminal, φ, and running, $\mathcal{L}$, cost terms) to be minimized as well as possible endpoint and path constraints (e and g, respectively),

$$\min \quad \int_D \varphi(T, x(T, s)) + \int_0^T \mathcal{L}(x(t, s), u(t)) dt \, ds$$
$$\text{s.t.} \quad \frac{d}{dt} x(t, s) = F(t, s, x(t, s), u(t))$$
$$e(x(0, s), x(T, s)) = 0$$
$$g(x(t, s), u(t)) \leq 0. \tag{4}$$

A rich literature exists on parametric optimal control which has a similar formulation to that of optimal ensemble control (4). Parametric optimal control focuses on sensitivity analysis, with respect to system parameters, of a perturbed optimal control problem [24], while ensemble control synthesizes robust compensating controls for possibly large parameter dispersions.

An optimal nonlinear control problem of this form is, in general, analytically intractable. Computational methods are then required to solve such exceedingly complex optimal ensemble control problems. The idea from our previous work that constructing appropriate polynomials is a key tool in characterizing the controllability of ensemble systems of interest motivates the use of polynomials within the computational framework [16]. Below, we review the main ideas of the previously established pseudospectral method for optimal control to lay the foundation for our developed extension to optimal ensemble control problems. In Section IV, we complete this framework with a proof of convergence of this numerical method.

Without loss of generality, we consider a general continuous-time optimal control problem defined on the time interval $\Omega = [-1, 1]$, which can be achieved by a simple affine transformation.

1) Problem 1 (Continuous-Time Optimal Control):

$$\min \quad J(x, u) = \varphi(x(1)) + \int_{-1}^{1} \mathcal{L}(x(t), u(t)) dt \tag{5}$$

$$\text{s.t.} \quad \frac{d}{dt} x(t) = f(t, x(t), u(t)) \tag{6}$$
$$e(x(-1), x(1)) = 0 \tag{7}$$
$$g(x(t), u(t)) \leq 0 \tag{8}$$
$$\|u\|_\infty \leq A, \quad u \in H_m^\alpha(\Omega), \ \alpha > 2 \tag{9}$$

where $\varphi \in C^0$ is the terminal cost; the running cost, $\mathcal{L} \in C^\alpha$, where C^α is the space of continuous functions with α classical derivatives, and dynamics, $f \in C_n^{\alpha-1}$, where $C_n^{\alpha-1}$ is the space of n-vector valued $C^{\alpha-1}$ functions, with respect to its arguments, the state $x(t) \in \mathbb{R}^n$ and control $u(t) \in \mathbb{R}^m$; e and g are terminal and path constraints, respectively; $H_m^\alpha(\Omega)$ is the m-vector valued Sobolev space. The norm associated with the Sobolev space with $m = 1$, $H^\alpha(\Omega)$, is given with respect to the $L^2(\Omega)$ norm [25]

$$\|h\|_{(\alpha)} = \left(\sum_{k=0}^{\alpha} \|h^{(k)}\|_2^2 \right)^{1/2}.$$

C. Pseudospectral Method

The pseudospectral method was originally developed to solve problems in fluid dynamics and since then has been successfully used for optimal control [26]–[28] and applied to various areas [21], [22]. Pseudospectral discretization methods use expansions of orthogonal polynomials to approximate the states of the system and thereby inherit the spectral accuracy characteristic of orthogonal polynomial expansions (the kth coefficient of the expansion decreases faster than any inverse power of k) [25]. Through special properties, derivatives of these orthogonal polynomials can be expressed in terms of the polynomials themselves, making it possible to accurately approximate the differential equation that describes the dynamics with an algebraic relation imposed at a small number of discretization points. An appropriate choice of these discretization points, or nodes, facilitates the approximation of the states as well as ensuring accurate numerical integration through Gaussian quadrature.

As a collocation (or interpolation) method, the pseudospectral method uses Lagrange polynomials to approximate the states and controls of the optimal control problem

$$x(t) \approx \mathcal{I}_N x(t) = \sum_{k=0}^{N} \bar{x}_k \ell_k(t) \tag{10}$$

$$u(t) \approx \mathcal{I}_N u(t) = \sum_{k=0}^{N} \bar{u}_k \ell_k(t) \tag{11}$$

where $x(t_k) = \mathcal{I}_N x(t_k) = \bar{x}_k$ and $u(t_k) = \mathcal{I}_N u(t_k) = \bar{u}_k$ because the Lagrange polynomials have the property $\ell_k(t_i) = \delta_{ki}$, where δ_{ki} is the Kronecker delta function and t_k are the interpolation nodes [29]. Therefore, the coefficients $\bar{x}_k$ and $\bar{u}_k$ are the discretized values of the original problem and become the decision variables of the subsequent discrete problem.

Although the interpolation with Lagrange polynomials discretizes the original problem, we require a means to ensure that both the integral in the cost functional is computed accurately and the dynamics are obeyed. The integral can be approximated through Gauss quadrature; here we use Legendre polynomials as the orthogonal basis for the pseudospectral method. The Legendre–Gauss–Lobatto (LGL) quadrature approximation

$$\int_{-1}^{1} f(t) dt \approx \sum_{i=1}^{N} f(t_i) w_i, \quad w_i = \int_{-1}^{1} \ell_i(t) dt \tag{12}$$
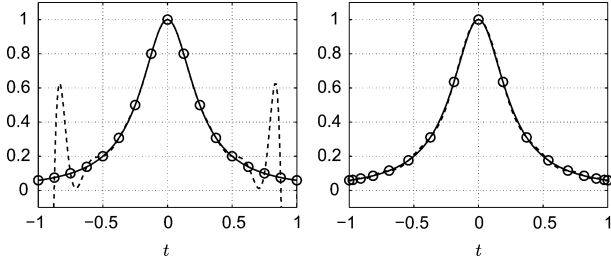
Fig. 1. Runge phenomena is aptly demonstrated by approximating $f(t) = 1/(16t^2 + 1)$ with an $N = 16$ order interpolating polynomial using uniform (left) and LGL (right) nodes.

is exact if the integrand $f \in \mathbb{P}_{2N-1}$ and the nodes $t_i \in \Gamma^{LGL}$, where $\mathbb{P}_{2N-1}$ denotes the set of polynomials of degree at most $2N-1$ and where $\Gamma^{LGL} = \{t_i : \dot{L}_N(t)|_{t_i} = 0, \, i = 1, \ldots, N-1\} \bigcup \{-1, 1\}$ are the $N+1$ LGL nodes determined by the derivative of the Nth order Legendre polynomial, $\dot{L}_N(t)$, and the interval endpoints [25]. To set the notation, the $\mathcal{I}_N$ operator will perform a collocation approximation using the LGL nodes, unless stated otherwise. The effect of this node choice can be seen in Fig. 1.

Using the LGL nodes, we can rewrite the Lagrange polynomials in terms of the orthogonal Legendre polynomials, which is critical to inherit the special derivative and spectral accuracy properties of the orthogonal polynomials despite using Lagrange interpolating polynomials. Given $t_k \in \Gamma^{LGL}$, we can express the Lagrange polynomials as [30]

$$\ell_k(t) = \frac{1}{N(N+1)L_N(t_k)} \frac{(t^2-1)\dot{L}_N(t)}{t - t_k}.$$

The derivative of (10) at $t_j \in \Gamma^{LGL}$ is then

$$\frac{d}{dt}\mathcal{I}_N x(t_j) = \sum_{k=0}^{N} \bar{x}_k \dot{\ell}_k(t_j)$$
$$= \sum_{k=0}^{N} D_{jk}\bar{x}_k \doteq (\mathcal{D}_N x)(t_j) \tag{13}$$

where D is the constant differentiation matrix [31].

We are now able to write the discretized optimal control problem using (10)–(13). We transform the continuous-time problem to a constrained optimization

$$\min \quad \varphi(\bar{x}_N) + \sum_{i=0}^{N} \mathcal{L}(\bar{x}_i, \bar{u}_i)w_i^N$$
$$\text{s.t.} \quad \sum_{k=0}^{N} D_{jk}\bar{x}_k = f(\bar{x}_j, \bar{u}_j)$$
$$e(\bar{x}_0, \bar{x}_N) = 0$$
$$g(\bar{x}_j, \bar{u}_j) \leq 0, \quad \forall j \in \{0, 1, \ldots, N\}. \tag{14}$$

D. Multidimensional Pseudospectral Method

The pseudospectral method lends itself to a natural extension to consider the ensemble case, which we develop here. This is most readily apparent for a single parameter variation, i.e., $s \in$

$[a, b] \subset \mathbb{R}$, however, is easily scaled to an arbitrary parameter dimension. In this basic case, the ensemble extension of (10) is

$$x(t, s) \approx \mathcal{I}_{N \times N_s} x(t, s) = \sum_{k=0}^{N} \bar{x}_k(s)\ell_k(t)$$
$$\approx \sum_{k=0}^{N} \left(\sum_{r=0}^{N_s} \bar{x}_{kr}\ell_r(s) \right) \ell_k(t).$$

The approximate derivative from (13) at the LGL nodes in the respective t and s domains, $t_i \in \Gamma^{LGL}$ and $s_j \in \Gamma_{N_s}^{LGL}$, is

$$\frac{d}{dt}\mathcal{I}_{N \times N_s} x(t_i, s_j) = \sum_{k=0}^{N} D_{ik} \left(\sum_{r=0}^{N_s} \bar{x}_{kr}\ell_r(s_j) \right)$$
$$= \sum_{k=0}^{N} D_{ik}\bar{x}_{kj}$$

where $\bar{x}_{kj} = x(t_k, s_j)$. In these equations we use a two-dimensional interpolating grid at the $N+1$ and N_s+1 LGL nodes in time and the parameter, respectively. For a general number of parameters, $\mathbf{s} = (s_1, s_2, \ldots, s_d)' \in D \subset \mathbb{R}^d, d > 1$

$$x(t, \mathbf{s}) \approx \mathcal{I}_{N \times N_{s_1} \times \cdots \times N_{s_d}} x(t, \mathbf{s}) = \sum_{k=0}^{N} \bar{x}_k(\mathbf{s})\ell_k(t)$$
$$= \sum_{k=0}^{N} \sum_{r_1=0}^{N_{s_1}} \cdots \sum_{r_d=0}^{N_{s_d}} \bar{x}_{kr_1 \ldots r_d}\ell_{r_d}(s_d) \cdots$$
$$\times \ell_{r_1}(s_1)\ell_k(t). \tag{15}$$

and the derivative is, correspondingly, with $\mathbf{j} = (j_1, j_2, \ldots, j_d)'$

$$\frac{d}{dt}\mathcal{I}_{N \times N_{s_1} \times \cdots \times N_{s_d}} x(t, \mathbf{s_j}) = \sum_{k=0}^{N} D_{ik}\bar{x}_{kj_1 \ldots j_d}. \tag{16}$$

The simplification from (15) to (16) illustrates why the pseudospectral approximations are effective methods for ensemble control, as they mimic the lack of information in the parameter dimension. This aspect will also make the extension of the convergence proof for ensemble systems straightforward, as will be discussed in Section IV.

III. EXAMPLES

In this paper, we consider several examples based on the prototypical quantum control system described by the Bloch equations [32]. The Bloch equations have been found to model a range of quantum phenomena from protein spectroscopy in nuclear magnetic resonance (NMR) [1] and medical scans in magnetic resonance imaging (MRI) [33] to Rabbi oscillations in quantum optics [34]. In the following discussion, we will consider the specific application and terminology for NMR spectroscopy; however, the methods and results are easily transferred to these other areas of interest. In NMR spectroscopy, when the duration of the pulse design problem is small compared with the relaxation times ($T \ll T_1, T_2$, the characteristic longitudinal and transverse relaxation times, respectively), the evolution of spins can be well approximated as sequences of unitary rotations driven by the static magnetic field and the applied electromagnetic controls. In practice, the effective fields generating
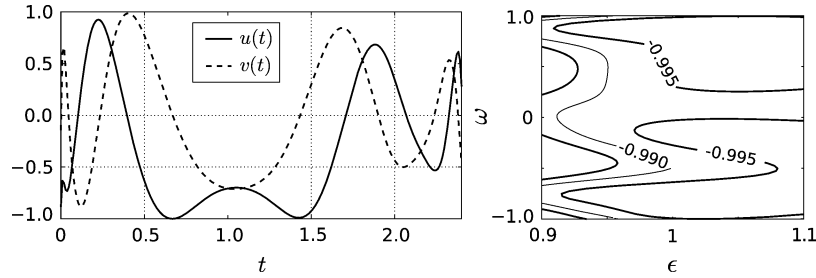
Fig. 2.   The "inversion" control pulse designed by the multidimensional pseudospectral method to make the state transfer $M(0, \omega, \epsilon) = (0\ 0\ 1)' \rightarrow M(T, \omega, \epsilon) = (0\ 0\ -1)'$ while compensating for $\omega \in [-1, 1]$ and $\epsilon \in [0.9, 1.1]$, i.e., $B = 1$ and $\delta = 0.1$. The final states $M_z(T, \omega, \epsilon)$ shown have an average value less than -0.99, achieving a highly uniform transfer across the ensemble.

these rotations show variation across the quantum sample due to hardware imperfection and chemical shielding, which leads us to consider a range of magnetic field variations. The corresponding dimensionless Bloch equations in the rotating frame (see Appendix A) are

$$\frac{d}{dt} M(t, \omega, \epsilon) = [\omega\Omega_z + \epsilon u(t)\Omega_y + \epsilon v(t)\Omega_x] M(t, \omega, \epsilon) \quad (17)$$

where $M(t, \omega, \epsilon) = (M_x(t, \omega, \epsilon), M_y(t, \omega, \epsilon), M_z(t, \omega, \epsilon))^T$ is the Cartesian magnetization vector for the parameter values $s = (\omega, \epsilon)$, $\omega \in [-B, B] \subset \mathbb{R}$, is the dispersion of natural frequencies, $\epsilon \in [1 - \delta, 1 + \delta]$, $0 < \delta < 1$, is the amplitude attenuation factor, and $\Omega_\alpha \in SO(3)$ is the generator of rotation around the α axis. A pulse that compensates for the dispersion in frequency and is insensitive to the scaling of the applied controls is called a broadband pulse robust to inhomogeneity. In this section, we consider several examples based on this model, including pulses robust not only to frequency dispersion and inhomogeneity, but also robust to uncertainty in initial conditions and time-varying frequencies. General implementation details can be found in Appendix B.

A. Robust π Pulse

Variation and dispersion in system dynamics pervade all physical experiments. In quantum systems, these inhomogeneities are often large enough to cause significant reduction in performance. The systematic framework we present here provides a rigorous way to frame any general pulse design problem for quantum control, as well as other areas of parameterized and uncertain systems.

A canonical problem in the control of quantum systems modeled by the Bloch equations is to design pulses that will accomplish a state-to-state transfer of the system. Such pulses, e.g., $\pi/2$ and π pulses (accomplishing $\pi/2$ and π rotations, respectively), are the fundamental building blocks of the pulse sequences used in many quantum experiments. Here, consider the inversion, or π, pulse that rotates the net magnetization from the equilibrium position $(+z)$ to the $-z$ axis, i.e., $M(0) = (0\ 0\ 1)' \rightarrow M(T) = (0\ 0\ -1)'$. In the ensemble case, this goal corresponds to a uniform inversion of the spin vector across all

choices of frequency and inhomogeneity. Specifically we consider the optimal ensemble control problem

$$\min \int_{1-\delta}^{1+\delta} \int_{-B}^{B} M_z(T, \omega, \epsilon)\, d\omega\, d\epsilon + \int_0^T u^2(t) + v^2(t)\, dt,$$

$$\text{s.t.}\ \frac{d}{dt} M(t, \omega, \epsilon) = \left[\omega\Omega_z + \epsilon u(t)\Omega_y + \epsilon v(t)\Omega_x\right] M(t, \omega, \epsilon),$$

$$M(0, \omega, \epsilon) = (0\ 0\ 1)',$$

$$\sqrt{u^2(t) + v^2(t)} \leq A,\ \forall\, t \in [0, T]$$

where A is the maximum allowable amplitude and the cost functional serves to equally minimize the z-component of the spin vector (integrated across the ensemble) and the energy of the designed pulse.

Fig. 2 displays a pulse that compensates for $B = 1$ and $\delta = 0.1$ (10%) as well as the corresponding inversion profile. In physical units for a normalizing amplitude of 20 kHz, the maximum amplitude is $A = 20$ kHz with bandwidth $\omega \in [-20, 20]$ kHz and duration $T = 120\ \mu$s. Pulses developed in this manner have been implemented experimentally in true protein NMR experiments to yield significant improvement in signal recovery [6]. Although designing individual pulses is of great importance and benefit, there are a myriad of other variations and uncertainties within typical quantum experiments, which calls for an approach that can address such new inhomogeneities and their corresponding challenges.

B. Uncertainty in Initial Conditions

In most experiments, individual pulses, such as the one in Fig. 2, are combined into a longer pulse sequence, which performs a more complicated manipulation of the system state with intermediate steps and goals. Even in the case of highly optimized individual pulses, as shown in the prior example, there is an error between the desired and actual final states. Moreover, pulses depend upon an exact (and usually uniform) initial condition in order to achieve their expected levels of performance. These effects combine to create a magnified accumulated error at the termination of the pulse sequence. The variation of the initial conditions of these pulses, therefore, causes significant degradation in achievable performance.

A representative example of such a pulse sequence is to perform a three step pulse sequence, which rotates the magnetization of the ensemble 1) from equilibrium $(+z)$ to a point on the

transverse plane (e.g., $+y$); 2) to the opposite point on the transverse plane (e.g., $-y$); 3) back to the equilibrium position ($+z$). Such pulses generally include "phase locking" pulses before and after the second pulse during which the magnetization dissipates. This dissipation is the portion of the experiment that is important to recover accurately and reflects a quantity to be measured, for example, a metabolic rate [35], [36]. If, in addition, there is accumulated error due to uncertainty in the initial conditions of the individual pulses, this leads directly to measurement inaccuracy. Here, by removing the "phase locking" pulses, we can abstract this pulse sequence to a unitary process and directly address any losses due to error. The controllability of the Bloch equations is shown constructing parameter-dependent (e.g., frequency, RF inhomogeneity) rotations of the spin vectors [18]. This, therefore, ensures that the problem with variation in initial conditions can be solved provided that the initial conditions can be parameterized by the frequency and RF-inhomogeneity.

Fig. 3 displays a three-stage optimized pulse designed by the multidimensional pseudospectral method which is robust to frequency dispersion and variation in the initial conditions of the three stages. This pulse was run as three concurrent optimizations, with the final states of one pulse fed in as the initial conditions of the next. This optimized pulse is compared with the combination of three separately optimized pulses; these combined pulses were designed with equal total duration. The terminal profiles at each intermediate goal quickly show the evidence of accumulated error in the case of the individually optimized pulses (each individual pulse has an average performance greater than 0.98). Most importantly, the uniformity of the inversion is lost in the additive error, with dips in performance down to 0.91.

C. Time-Varying Frequency

Until now, we have considered that the dispersion and uncertainty of the system are stationary. However, addressing time-varying fluctuations in parameters is also of particular theoretical and practical importance. For example, in the formulation of quantum control problems given in (2) we noted that the Hamiltonians can be time-varying, motivated by such phenomena as random telegraph noise [37]. The first step to addressing stochastic variations in such physical systems is to demonstrate control of time-varying systems, such as given by the expectation value of the corresponding random process.

Fig. 4 presents a series of optimizations designing $\pi/2$ pulses providing a state transfer $+z$ to $+x$, while compensating for a time-varying frequency, $\omega(t) = \sin(t)$. Various choices of cost functional yield different results. The arbitrary control pulse profile corresponding to the terminal cost $\varphi(T) = M_x(T)$ (Fig. 4, left) motivates studying optimal control methods that provide the capacity for hybrid objectives resulting in more physically meaningful controls, e.g., minimizing energy (middle) and time (right).

IV. CONVERGENCE

By accepting and implementing a numerical method, we implicitly assume that the transformations and discretization used to prepare the problem for computational work does not fundamentally alter the nature of the problem. It is then critically
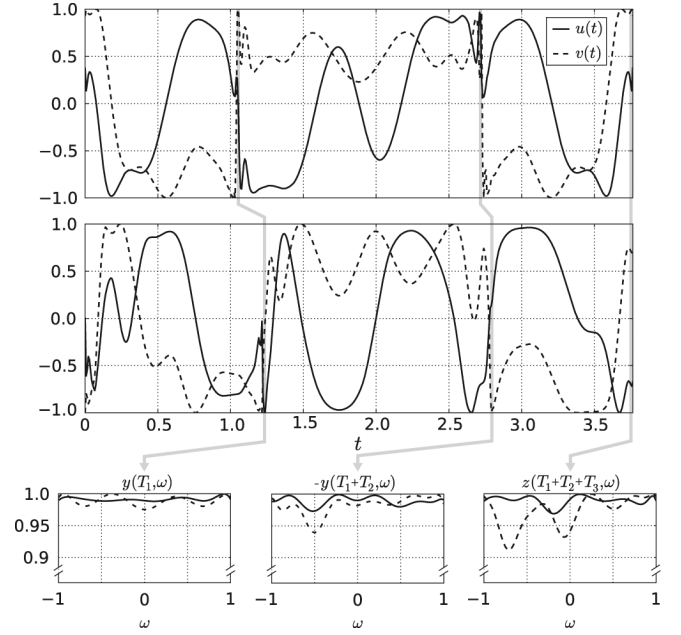

Fig. 3. Pulses are optimized to produce a desired $z \rightarrow y \rightarrow -y \rightarrow z$ evolution of the Bloch equations. The upper plot displays the concatenation of individually optimized $z \rightarrow y$ and $y \rightarrow -y$ pulses, which achieves the dashed terminal profiles shown below, with respective average performances: 0.99, 0.98, 0.97 (0.91 minimum). The middle plot displays a three-part simultaneously-optimized pulse robust to variation in the initial condition and achieves the solid terminal profiles shown below, with respective average performances: 0.99, 0.99, 0.99 (0.97 minimum). The noticeable enhancement in performance and uniformity is due to compensating for the inhomogeneity in the initial condition of the individual pulses.

important to show that this assumption is justified. Here we do so by both empirical and theoretical means. More specifically, we show that as the number of discretizations in the pseudospectral method (and samples in the multidimensional pseudospectral method) increases the solution of the algebraic nonlinear programming problem converges to the solution of the original continuous-time optimal control problem. For this argument, we consider a modified nonlinear programming problem statement.

1) Problem 2 (Algebraic Nonlinear Programming):

$$\min \quad \bar{J}(\bar{x}, \bar{u}) = \varphi(\bar{x}_N) + \sum_{k=0}^{N} \mathcal{L}(\bar{x}_k, \bar{u}_k) w_k \qquad (18)$$

$$\text{s.t.} \quad \left\| f(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x \right\|_N \leq c_d N^{1-\alpha} \qquad (19)$$

$$e(\bar{x}_0, \bar{x}_N) = 0 \qquad (20)$$

$$g(\bar{x}_k, \bar{u}_k) \leq 0 \qquad (21)$$

$$\|u_k\| \leq A \quad \forall k = 0, 1, \dots, N \qquad (22)$$

where c_d is a positive constant; we define the discrete $L_n^2(\Omega)$ norm $\|h\|_N = \sqrt{\langle h, h \rangle_N}$, for $h, h_1, h_2 \in L_n^2(\Omega)$, $\Omega = [-1, 1]$, with

$$\langle h_1, h_2 \rangle_N = \sum_{k=0}^{N} h_1'(t_k) h_2(t_k) w_k$$

where $'$ denotes the transpose and w_k is the Gauss quadrature weight from (12).

Remark 1: The dynamics in (19) have been relaxed from the equality in (14) to ensure the feasibility of the discrete problem,
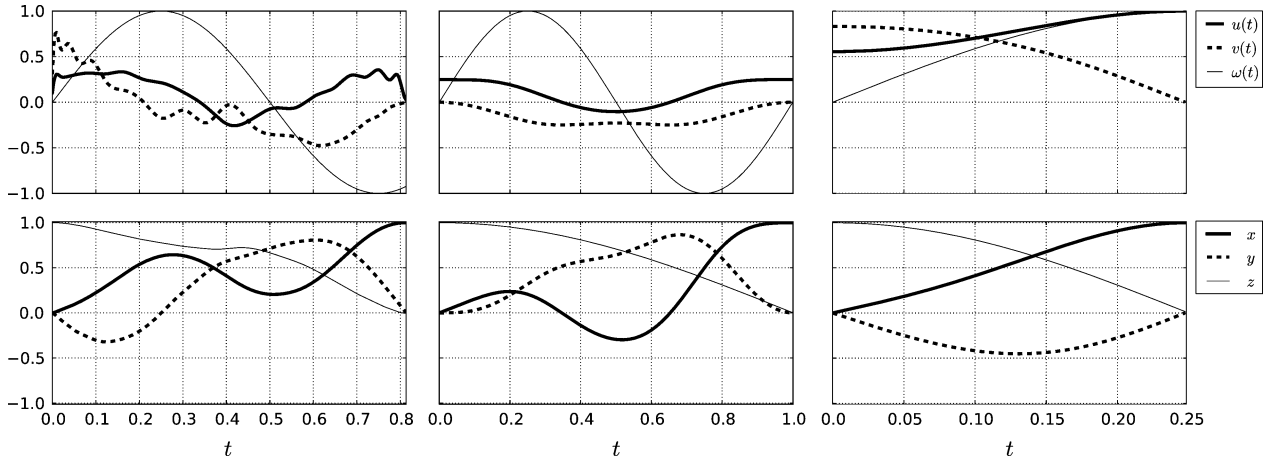
Fig. 4. Control pulses (top) and state trajectories (bottom) corresponding to different objectives and designed to compensate for the time-varying frequency $\omega(t) = \sin(t)$. A single-system state transfer $M(0) = (0\ 0\ 1)' \rightarrow M(T) = (1\ 0\ 0)'$ is designed using the terminal cost $\varphi(T) = M_x(T)$ and running costs $\mathcal{L}(t) = 0$ (left), $\mathcal{L}(t) = 0.1(u(t)^2 + v(t)^2)$ (middle), $\mathcal{L}(t) = 0.1$ (right). The terminal time was free in all cases, bounded by $T_{\max} = 1$.

which is used in Proposition 1. It is trivial to show that in the limit, as $N \rightarrow \infty$, these two conditions coincide.

We seek to address three questions related to solving the continuous-time optimal control (Problem 1) by solving the pseudospectral discretized constrained optimization (Problem 2). Suppose a feasible solution (x, u) exists to Problem 1. Under what conditions:

1) *Feasibility*: For a given order of approximation, N, does Problem 2 have a feasible solution, $(\bar{x}, \bar{u})$, which are the interpolation coefficients given in (10) and (11)?

2) *Convergence*: As N increases, does the sequence of optimal solutions, $\{(\bar{x}^\dagger, \bar{u}^\dagger)\}$, to Problem 2 have a corresponding sequence of interpolating polynomials which converges to a feasible solution of Problem 1? Namely,

$$\lim_{N \rightarrow \infty} (\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = (x, u)$$

3) *Consistency*: As N increases, does the convergent sequence of interpolating polynomials corresponding to the optimal solutions of Problem 2 converge to an optimal solution of Problem 1? Namely,

$$\lim_{N \rightarrow \infty} (\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = (x^*, u^*)$$

Remark 2: It is possible that Problem 1 has more than one optimal solution, i.e., there is more than one solution with the same optimal cost $J(x^*, u^*) = J^*$. Therefore, to show that the sequence of discrete solutions converges to *an* optimal solution, we can instead show that the cost of the discrete solution, $\bar{J}$, converges to the optimal cost J^*.

Previous work has been done in the area of convergence of the pseudospectral method and we aim to augment this literature with several key insights that make convergence results applicable to a wider class of systems and relax the conditions on which the current proofs are based. Rather complete analysis has been done for the class of nonlinear systems which can be feedback linearized, including convergence rates [38], [39]. We show below that ensemble quantum systems of interest do not fall within the class of feedback linearizable systems. Work has

also included general nonlinear systems, but with the assumption that the solutions of the algebraic nonlinear programming problem have a limit point (i.e., have a convergent subsequence) [40]. In the language used above, this is very close to assuming "Convergence," which in this presentation we relax and prove feasibility, convergence, and consistency directly. Finally, we examine the convergence of the multidimensional pseudospectral method as applied to ensemble optimal control problems. In what follows, we consider first the convergence of the standard pseudospectral method and then discuss the convergence of the ensemble case.

We first observe that ensemble control systems of interest are not feedback linearizable [41], which motivates a need for a more general convergence proof. Consider the bilinear Bloch equations in (17) without variation in RF inhomogeneity (i.e., $\epsilon = 1$). The ability to feedback linearize a general nonlinear system is given by the Lie algebra generated by the drift and control vector fields (the conditions on this algebra must hold for each control term individually; here we consider the case for u). In particular, the terms $\mathrm{ad}_{\omega\Omega_z}^0 \Omega_y = \Omega_y$, $\mathrm{ad}_{\omega\Omega_z}^1 \Omega_y = -\omega\Omega_x$, $\mathrm{ad}_{\omega\Omega_z}^2 \Omega_y = -\omega^2\Omega_y$, ..., and

$$\mathrm{ad}_{\omega\Omega_z}^{2k-1}\Omega_y = (-1)^k \omega^{2k-1}\Omega_x$$
$$\mathrm{ad}_{\omega\Omega_z}^{2k}\Omega_y = (-1)^k \omega^{2k}\Omega_y$$

where $k = 1, 2, \ldots$, and ω is any value in the interval $D \subset \mathbb{R}$. It is clear that this Lie algebra, with increasing powers of the parameter ω, is never closed. Therefore, the span of the appropriate Lie brackets is not involutive, which indicates that such a system is not feedback linearizable.

A. Empirical Convergence

The orthogonal polynomials of the pseudospectral method provide spectral convergence rates similar to Fourier series approximations for periodic functions, which can easily be seen in practice. Fig. 5 shows the rapid convergence of the method in both the discretization (time) and sampling (parameter) dimensions for a broadband $\pi/2$ pulse maximizing the terminal x value across the ensemble. As the order of discretization (N)
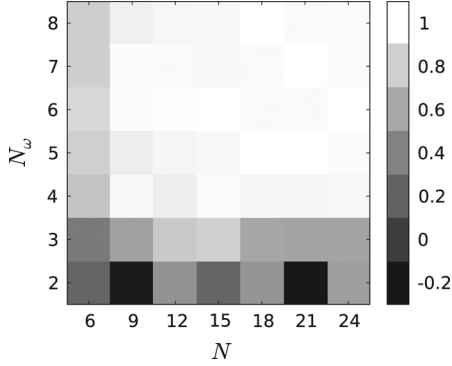
Fig. 5. Characteristic rapid convergence of the multidimensional pseudospectral method for a $\pi/2$ pulse designed to perform the state transfer $M(0, \omega, 1) = (0\ 0\ 1)' \rightarrow M(T, \omega, 1) = (1\ 0\ 0)'$, with $B = 1$ and $T = 1$. Average terminal values of $M_x(T, \omega, 1)$ are shown for various choices of N and $N\omega$.

and/or sampling (N_s) increase, the method yields an objective $(\varphi(T) = M_x(T, \omega, 1))$ that converges to the maximum value of unity. The low order of approximation is a characteristic of the orthogonal approximations at the heart of the numerical method. Although such empirical figures are convincing, we now show this convergence in a more rigorous fashion.

B. Theoretical Preliminaries

The results in this section will provide the foundation on which we can analyze the feasibility, convergence, and consistency of the pseudospectral approximation method for optimal control problems. We begin by presenting several key established results in polynomial approximation theory and the natural vector extensions. With these identities, we are able to then prove feasibility and convergence. We define an optimal solution to Problem 1 as any feasible solution that achieves the optimal cost $J(x^*, u^*) = J^*$. We use this definition of an optimal solution within the subsequent preliminaries and the main result.

Remark 2: Note that $x \in H_n^\alpha(\Omega)$. Since $x(t)$ exists and $f \in C_n^{\alpha-1}$, all the derivatives $x^{(k)} \in C_n^0, \forall k = 0, 1, \ldots, \alpha$ exist and are square integrable on the compact domain Ω, $x^{(k)} \in L_n^2(\Omega)$. Therefore, $x \in H_n^\alpha(\Omega)$.

Lemma 1 (Interpolation Error Bounds [25], p. 289): If $h \in H^\alpha(\Omega)$, $\Omega = [-1, 1]$, the following hold for an interpolation approximation at the LGL nodes, with $c_1, , c_2, , c_3, , c > 0$.

a) The interpolation error is bounded,

$$\|h - \mathcal{I}_N h\|_2 \le c_1 N^{-\alpha}\|h\|_{(\alpha)}.$$

b) The error between the exact derivative and the derivative of the interpolation is bounded

$$\|\dot{h} - \mathcal{D}_N h\|_2 \le c_2 N^{1-\alpha}\|h\|_{(\alpha)}.$$

The same bound holds for the discrete $L^2(\Omega)$ norm

$$\|\dot{h} - \mathcal{D}_N h\|_N \le c_3 N^{1-\alpha}\|h\|_{(\alpha)}.$$

c) The error due to quadrature integration is bounded

$$\left| \int_{-1}^{1} h(t)dt - \sum_{k=0}^{N} h(t_k)w_k \right| \le cN^{-\alpha}\|h\|_{(\alpha)}$$

where t_k is the kth LGL node and w_k is the corresponding k^{th} weight for LGL quadrature as defined in (12).

Remark 4: The convergence of interpolation approximations are highly dependent on the choice of collocation nodes. Fig. 1 depicts the interpolation error possible using different types of node choices.

Lemma 2: If $h \in H_n^\alpha(\Omega)$, i.e., an n-vector valued Sobolev space, $h = (h_1\ h_2\ \ldots\ h_n)'$, $h_i \in H^\alpha(\Omega)$, $i = 1, 2, \ldots, n$.

1) The vector-valued extension of Lemma 1a is, by the triangle inequality on the $L_n^2(\Omega)$ norm,

$$\|h - \mathcal{I}_N h\|_2 \le \sum_{i=1}^{n}\|h_i - \mathcal{I}_N h_i\|_2 \le \sum_{i=1}^{n} c_i N^{-\alpha}\|h_i\|_{(\alpha)}.$$

2) Similarly, 1b can be extended:

$$\|\dot{h} - \mathcal{D}_N h\|_2 \le \sum_{i=1}^{n}\|\dot{h} - \mathcal{D}_N h\|_2$$
$$\le \sum_{i=1}^{n} c_i N^{1-\alpha}\|h_i\|_{(\alpha)}$$
$$\le cN^{1-\alpha}$$

which again also holds for the discrete $L_n^2(\Omega)$ norm.

Proposition 1 (Feasibility): Given a solution (x, u) of Problem 1, then Problem 2 has a feasible solution, $(\bar{x}, \bar{u})$, which are the corresponding interpolation coefficients.

Proof: Given the feasible solution (x, u), let $(\mathcal{I}_N x, \mathcal{I}_N u)$ be the polynomial interpolation of this solution at the LGL nodes. Our aim is to show that the coefficients of this interpolation satisfy (19)–(21) of Problem 2. Consider the constraints imposed by the dynamics in (19). Because the discrete norm is evaluated only at the interpolation points

$$\|f(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x\|_N = \|f(x, u) - \mathcal{D}_N x\|_N$$
$$= \|\dot{x} - \mathcal{D}_N x\|_N$$
$$\le c_d N^{1-\alpha}$$

where the last step is given by Lemma 2b. Therefore, the interpolation coefficients $(\bar{x}, \bar{u})$ satisfy the dynamics of Problem 2 in (19). We can easily show that the path constraints are also satisfied because $g(x(t), u(t)) \le 0$ for all $t \in \Omega$ by (8). Since this holds for all $t \in \Omega$, it also holds for all LGL nodes $t_k \in \Gamma^{\text{LGL}}$, i.e.,

$$g(\bar{x}_k, \bar{u}_k) = g(x(t_k), u(t_k)) \le 0$$

which gives (21). The endpoint constraints are trivially satisfied by the definition of interpolation and the presence of interpolation nodes at both endpoints. Therefore, $(\bar{x}, \bar{u})$ is a feasible solution to Problem 2. ∎

Proposition 2 (Convergence): Given the sequence of solutions to Problem 2, $\{(\bar{x}, \bar{u})\}_N$, then the sequence of corresponding interpolation polynomials, $\{(\mathcal{I}_N x, \mathcal{I}_N u)\}$, has a convergent subsequence, such that

$$\lim_{N_j \to \infty} (\mathcal{I}_{N_j} x, \mathcal{I}_{N_j} u) = (\mathcal{I}_\infty x, \mathcal{I}_\infty u)$$

which is a feasible solution to Problem 1.

Proof: Given that $(\bar{x}, \bar{u})$ is a feasible solution of Problem 2, it satisfies (19)–(21). Our goal is to show that the sequence of polynomials, $\{(\mathcal{I}_N x, \mathcal{I}_N u)\}_N$, 1) is bounded, 2) has a convergent subsequence and 3) its limit is a feasible solution of Problem 1, satisfying (6)–(8).

1) Explicitly writing out the discrete norm in (19) gives

$$\left(\sum_{k=0}^{N} \sum_{i=1}^{n} (f_i(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x_i)^2 (t_k) \right)^{1/2} \le c_d N^{1-\alpha}.$$

Because f is continuous, it satisfies

$$\lim_{N \to \infty} \left(f_i(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x_i \right)(t_k)$$
$$= \left(f_i(\lim_{N \to \infty} \mathcal{I}_N x_i, \lim_{N \to \infty} \mathcal{I}_N u) - (\lim_{N \to \infty} \mathcal{I}_N x_i)' \right)(t_k)$$
$$= 0, \qquad (23)$$

which means that the derivative of the interpolating polynomial and the state dynamics match at the interpolation nodes. Moreover, as $N \to \infty$, the LGL nodes $t_k \in \Gamma^{\mathrm{LGL}}$ are dense in Ω, which shows that they match along the entire domain.

2) The sequence $\{\mathcal{I}_N x\}$ is a sequence of polynomials on the compact domain Ω; therefore, for each finite N, $\mathcal{I}_N x \in H_n^\alpha(\Omega)$. In the limit, we showed above in (23) that $(\lim_{N \to \infty} \mathcal{I}_N x)'$ matches the state dynamics $f \in C_n^{\alpha-1}$, so that $\{\mathcal{I}_N x\}$ is bounded, because f is bounded over Ω, and also satisfy $\mathcal{I}_N x \in H_n^\alpha(\Omega)$ for all N. With the boundedness of these interpolating polynomials $\{\mathcal{I}_N x\}$ all supported in Ω, Rellich's Theorem (cf., e.g., [42, p. 272]) gives that there is a subsequence $\{\mathcal{I}_{N_j} x\}$ which converges in $H_n^{\alpha-1}(\Omega)$. The same is true for the control interpolating polynomial. Therefore, there exists at least one limit point of the function sequence $\{(\mathcal{I}_N x, \mathcal{I}_N u)\}$ which we denote $(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$.

3) Because $\{\mathcal{I}_N x\}$ has a convergent subsequence, we can express (23) as

$$\frac{d}{dt}(\mathcal{I}_\infty x)(t_k) = f(\mathcal{I}_\infty x, \mathcal{I}_\infty u)(t_k) \qquad (24)$$

which states that $(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$ satisfies the dynamics in (6) at the interpolation nodes. Again, as $N \to \infty$, the LGL nodes $t_k \in \Gamma^{\mathrm{LGL}}$ are dense in Ω, which further shows that $(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$ satisfies the dynamics of Problem 1 at all points on the interval Ω. Similarly, one can prove that this solution satisfies the path constraints because the LGL nodes become dense in Ω as $N \to \infty$ and $g(\bar{x}_k, \bar{u}_k) = g(x(t_k), u(t_k)) \le 0$ at all LGL nodes. Again, the endpoint constraints are met exactly because the LGL grid has nodes at the endpoints. ∎

Lemma 3: Given (x, u), where $x \in H_n^\alpha(\Omega)$, $u \in H_m^\alpha(\Omega)$, and the corresponding interpolation coefficients, $(\bar{x}, \bar{u})$, then the error in the continuous and discrete cost functionals defined in (5) and (18), respectively, due to interpolation is given by

$$|J(x, u) - \bar{J}(\bar{x}, \bar{u})| \le cN^{-\alpha}.$$

Remark 5: Notice that (x, u) and $(\bar{x}, \bar{u})$ are not required to be a feasible solutions to Problem 1 and 2, respectively. This result characterizes the error due to interpolation.

Proof: From (6) and (19) since $\varphi(x(1)) = \varphi(\bar{x}_N)$

$$|J(x, u) - \bar{J}(\bar{x}, \bar{u})| = \left| \int_{-1}^{1} \mathcal{L}(x, u)dt - \sum_{k=0}^{N} \mathcal{L}(\bar{x}_k, \bar{u}_k)w_k \right|.$$

Since $\mathcal{L} \in C^\alpha$ with respect to both the state and control, $x \in H_n^\alpha(\Omega)$ and $u \in H_m^\alpha(\Omega)$, the composite function $\tilde{\mathcal{L}}(t) = \mathcal{L}(x(t), u(t)) \in H^\alpha(\Omega)$. Let $\mathcal{L}_k = \mathcal{L}(\bar{x}_k, \bar{u}_k)$. Substituting these definitions and employing Lemma 1c, we obtain

$$\left| \int_{-1}^{1} \tilde{\mathcal{L}}(t)dt - \sum_{k=0}^{N} \mathcal{L}_k w_k \right| \le cN^{-\alpha} \|\tilde{\mathcal{L}}(t)\|_{(\alpha)}.$$

Since $\tilde{\mathcal{L}} \in H^\alpha(\Omega)$, $\|\tilde{\mathcal{L}}(t)\|_{(\alpha)}$ is finite and the result follows. ∎

V. MAIN RESULT

Theorem 1 (Consistency): Suppose Problem 1 has an optimal solution (x^*, u^*). Given a sequence of optimal solutions to Problem 2, $\{(\bar{x}^\dagger, \bar{u}^\dagger)\}_N$, then the corresponding sequence of interpolating polynomials, $\{(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)\}_N$, has a limit point, $(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$ which is an optimal solution to the original optimal control problem.

Proof: We break the proof into four sections, employing the results from the previous section.

1) By Proposition 1, since (x^*, u^*) is a solution to Problem 1, then for each choice of N, the corresponding interpolation coefficients, $(\bar{x}^*, \bar{u}^*)$, are a feasible solution to Problem 2. By the definition of optimality of $(\bar{x}^\dagger, \bar{u}^\dagger)$

$$\bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \le \bar{J}(\bar{x}^*, \bar{u}^*). \qquad (25)$$

2) By Proposition 2, the limit point of the polynomial interpolation of the discrete optimal solution to Problem 2, $\lim_{N \to \infty}(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = (\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$, is a feasible solution of Problem 1. Therefore, we have, by the definition of the optimality of (x^*, u^*) and the continuity of J

$$J(x^*, u^*) \le \lim_{N \to \infty} J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = J(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger). \qquad (26)$$

3) Using Lemma 3, we can bound the error in the cost between the optimal solution of Problem 1, (x^*, u^*), and the corresponding interpolating coefficients, $(\bar{x}^*, \bar{u}^*)$, as

$$|J(x^*, u^*) - \bar{J}(\bar{x}^*, \bar{u}^*)| \le c_1 N^{-\alpha}. \qquad (27)$$

Similarly, we can bound the error in the cost between the optimal solution of Problem 2, $(\bar{x}^\dagger, \bar{u}^\dagger)$, and the polynomial interpolation of this solution, $(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$, as

$$|J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)| \le c_2 N^{-\alpha}. \qquad (28)$$

Recall that Lemma 3 does not require $(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$ to be a feasible solution of Problem 1. From (27) and (28)

$$\lim_{N \to \infty} \bar{J}(\bar{x}^*, \bar{u}^*) = J(x^*, u^*) \quad (29)$$

$$\lim_{N \to \infty} \left[J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \right] = 0. \quad (30)$$

4) We are now ready to assemble the various pieces of this proof. Combining (29) and (25) we have

$$\lim_{N \to \infty} \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \leq \lim_{N \to \infty} \bar{J}(\bar{x}^*, \bar{u}^*) = J(x^*, u^*).$$

Adding the result from (26)

$$\lim_{N \to \infty} \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \leq J(x^*, u^*) \leq \lim_{N \to \infty} J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger). \quad (31)$$

Since the difference between the left and right sides, as given by (30), decreases to zero as $N \to \infty$, the quantities $\bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)$ and $J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$ converge to $J(x^*, u^*)$. In particular,

$$0 \leq \lim_{N \to \infty} \left[J(x^*, u^*) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \right]$$
$$\leq \lim_{N \to \infty} \left[J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \right] = 0.$$

Thus, the optimal discrete cost $\bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)$ of Problem 2 and the continuous cost $J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$ of the corresponding interpolation polynomials converge to the optimal cost $J(x^*, u^*)$ of Problem 1. Moreover, $(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$ is a feasible solution to Problem 1 and achieves the optimal cost. Therefore, $(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$ is an optimal solution to Problem 1. ∎

Remark 6 (Ensemble Extension): The nature in which the ensemble extension enters into the multidimensional pseudospectral method makes it straightforward to extend this convergence proof to the ensemble case. Section II-D showed the simplicity of the derivative term in multidimensional sampling with (16). Similarly, in the ensemble case, the constraints corresponding to the dynamics (19) operate entirely in parallel for different parameter values. The additional integration in the cost function over the parameter domain, as in (4) adds another layer of quadrature approximation that can be shown to converge with arguments similar to those presented above.

VI. CONCLUSION

In this work, we have presented a cohesive perspective and methodology for optimal control of inhomogeneous ensembles, as particularly motivated by compelling problems in quantum control and extendable to both parameterized systems in, for example, neuroscience [22] and uncertain systems throughout science and engineering. Such systems are mathematically characterized by considering a parameterized family of differential equations indexed by a parameter vector that shows variation. Applying this rigorous framework prompts us to solve the corresponding optimal control problems with computational methods of particular form. The notion of polynomial approximation entering into the controllability analysis of the Bloch equations indicates that a modified pseudospectral method is a prime candidate. The method has natural extensions which we develop to model ensemble variation. This direct collocation method transforms the continuous-time optimal control problem into an algebraic nonlinear programming problem, which we show to be effective in a variety of applications. We supplied additional and more general arguments for the convergence of this method, in particular relaxing several assumptions and discussing the convergence characteristic of the multidimensional pseudospectral method for optimal ensemble control.

APPENDIX A
DIMENSIONLESS BLOCH EQUATIONS

The Bloch equations without relaxation, $\dot{\mathcal{M}} = \mathcal{M} \times \gamma \mathcal{B}_{\text{eff}}$, utilizes the classical description of interacting electromagnetic forces, where $\mathcal{M}$ is the spin magnetization vector, γ is the gyromagnetic ratio, the effective externally applied field is $\mathcal{B}_{\text{eff}} = (B_1 \cos(\omega_0 t + \phi), B_1 \sin(\omega_0 t + \phi), B_0)'$, $B_1(t)$, and B_0 are the amplitudes of the applied fields in the transverse plane and z direction, respectively, and $\phi(t)$ is the phase angle [1]. Conventionally, the fields are given as frequencies $\gamma \mathcal{B}_{\text{eff}} = (\omega_{1x}, \omega_{1y}, \omega_0)$ and measured in units of Hertz. Using the generators of rotation

$$\Omega_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \Omega_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$\Omega_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

the Bloch equations are be given by

$$\frac{d}{dt} \mathcal{M}(t) = \left[\omega_0 \Omega_z + \omega_{1y}(t) \Omega_y + \omega_{1x}(t) \Omega_x \right] \mathcal{M}(t). \quad (32)$$

If we consider variation in the applied electromagnetic fields B_0 and B_1, we can express (32) in matrix form as in (33) at the bottom of the page, where $\omega \in [-\beta, \beta]$ and $\epsilon \in [1 - \delta, 1 + \delta]$, $0 \leq \delta \leq 1$. For calculation and computation, it is useful to transform the Bloch equations into the so-called rotating frame and normalize the system by a nominal pulse amplitude A to yield a dimensionless equation. Solutions based on the dimensionless equation can then be scaled for a specific choice of nominal amplitude. Consider a transformation $M = \exp(-\omega_0 \Omega_z t) \mathcal{M}$. In addition, we scale time with $\tau = At$.

$$\frac{d}{dt} \begin{bmatrix} \mathcal{M}_x(t,\omega,\epsilon) \\ \mathcal{M}_y(t,\omega,\epsilon) \\ \mathcal{M}_z(t,\omega,\epsilon) \end{bmatrix} = \gamma \begin{bmatrix} 0 & -(\omega_0 + \omega) & \epsilon B_1 \sin(\omega_0 t + \phi) \\ \omega_0 + \omega & 0 & -\epsilon B_1 \cos(\omega_0 t + \phi) \\ -\epsilon B_1 \sin(\omega_0 t + \phi) & \epsilon B_1 \cos(\omega_0 t + \phi) & 0 \end{bmatrix} \begin{bmatrix} \mathcal{M}_x(t,\omega,\epsilon) \\ \mathcal{M}_y(t,\omega,\epsilon) \\ \mathcal{M}_z(t,\omega,\epsilon) \end{bmatrix} \quad (33)$$

It is straightforward to show that the new state equation is given by

$$\frac{d}{d\tau}M(\tau,\omega,\epsilon) = \left[\omega\Omega_z + \epsilon u(\tau)\Omega_y + \epsilon v(\tau)\Omega_x \right] M(\tau,\omega,\epsilon)$$

where $\tau \in [0, AT \times 2\pi]$, $\omega \in [-B, B]$, $B = \beta/A$, and

$$u(\tau) = \frac{\gamma B_1\left(\frac{\tau}{A}\right)}{A}\cos\left(\phi\left(\frac{\tau}{A}\right)\right)$$

$$v(\tau) = \frac{\gamma B_1\left(\frac{\tau}{A}\right)}{A}\sin\left(\phi\left(\frac{\tau}{A}\right)\right)$$

(all dimensionless). Note the 2π factor in the time scaling is introduced to convert from units of Hertz to radians/second. Designing the time-varying controls $u(\tau)$ and $v(\tau)$ is equivalent to the original design of amplitude $B_1(t)$ and phase $\phi(t)$.

APPENDIX B
PRACTICAL IMPLEMENTATION

The purpose of the pseudospectral framework discussed in this paper is to discretize a continuous optimal ensemble control problem into a discrete nonlinear programming problem. The methods discussed in Sections II-C and II-D and applications in Section III leave a significant amount of flexibility in terms of how to implement this approximation methodology.

The approximation techniques involved in the pseudospectral method are straightforward and can be accomplished with any scientific-capable programming language or interface. Our prototyping in this area began in MATLAB; in fact, there are a few MATLAB based implementations of the entire pseudospectral method, for which only system-specific details must be written by the user. In favor of flexibility, this work was done entirely with custom software.

Ultimately, the Legendre pseudospectral method evidences a straightforward pattern, which can be directly written into a nonlinear programming problem. AMPL is a very general modeling language to pose optimization problems. Our implementation specifically employs AMPL with a commercial software KNITRO to solve the nonlinear programming problems, wrapped by several python scripts that generate, for example, the D matrix and LGL nodes of appropriate dimension. Moreover, the condensed notation of the AMPL language provides a succinct way to state optimal ensemble control problems.

APPENDIX C
RELLICH'S THEOREM

*Theorem 2 (Rellich's Theorem [42], p. 272):* Suppose $\{f_k\}$ is a sequence in H^α such that
1) $\sup_k \|f_k\|_{(\alpha)} < \infty$;
2) the f_k's are all supported in a fixed compact set V.
Then there is a subsequence $\{f_{k_j}\}$ which converges in H^β for all $\beta < \alpha$.

REFERENCES

[1] R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions.* Oxford, U.K.: Clarendon, 1987.

[2] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information.* Cambridge, U.K.: Cambridge Univ. Press, 2011.

[3] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, III, M. Rance, and N. J. Skelton, *Protein NMR Spectroscopy.* New York: Academic, 2007.

[4] R. van Handel, J. K. Stockton, and H. Mabuchi, "Feedback control of quantum state reduction," *IEEE Trans. Autom. Control*, vol. 50, no. 6, pp. 768–780, Jun. 2005.

[5] C. D'Helon and M. R. James, "Stability, gain, and robustness in quantum feedback networks," *Phys. Rev. A*, vol. 73, p. 053803, 2006.

[6] J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, and G. Wagner, "Optimal pulse design in quantum control: A unified computational method," *Proc. Nat. Acad. Sci.*, vol. 108, no. 5, pp. 1879–1884, 2011.

[7] M. H. Levitt, "Composite pulses," *Prog. Nucl. Magn. Reson. Spectrosc.*, vol. 18, pp. 61–122, 1986.

[8] E. L. Hahn, "Spin echoes," *Phys. Rev.*, vol. 80, pp. 580–594, 1950.

[9] A. J. Shaka and R. Freeman, "Composite pulses with dual compensation," *J. Magn. Reson.*, vol. 55, pp. 487–493, 1983.

[10] J. Pauly, R. Le Roux, D. Nishimura, and A. Macovski, "Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm," *IEEE Trans. Med. Imaging*, vol. 10, pp. 53–65, 1991.

[11] S. Conolly, D. Nishimura, and A. Macovski, "Optimal control to the magnetic resonance selective excitation problem," *IEEE Trans. Med. Imaging*, vol. MI-5, no. 2, pp. 106–115, Jun. 1986.

[12] T. E. Skinner, T. Reiss, B. Luy, N. Khaneja, and S. J. Glaser, "Application of optimal control theory to the design of broadband excitation pulses for high resolution NMR," *J. Magn. Reson.*, vol. 163, pp. 8–15, 2003.

[13] C. T. Kehlet *et al.*, "Improving solid-state NMR dipolar recoupling by optimal control," *J. Am. Chem. Soc.*, vol. 126, pp. 10202–10203, 2004.

[14] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, "Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms," *J. Magn. Reson.*, vol. 172, pp. 296–305, 2005.

[15] I. I. Maximov, Z. Tosner, and N. C. Nielsen, "Optimal control design of NMR and dynamic nuclear polarization experiments using monotonically convergent algorithms," *J. Chem. Phys.*, vol. 128, p. 184505, 2008.

[16] J.-S. Li and N. Khaneja, "Ensemble control of bloch equations," *IEEE Trans. Autom. Contr.*, vol. 54, no. 3, pp. 528–536, Mar. 2009.

[17] J.-S. Li, "Ensemble control of finite-dimensional time-varying linear systems," *IEEE Trans. Autom. Contr.*, vol. 56, no. 2, pp. 345–357, Feb. 2011.

[18] J.-S. Li and N. Nhaneja, "Control of inhomogeneous quantum ensembles," *Phys. Rev. A*, vol. 73, p. 030302, 2006.

[19] J. Ruths and J.-S. Li, "A multidimensional pseudospectral method for optimal control of quantum ensembles," *J. Chem. Phys.*, vol. 134, p. 044128, 2011.

[20] G. Lindblad, "On the generators of quantum dynamical semigroups," *Commun. Math. Phys.*, vol. 48, pp. 119–130, 1976.

[21] J.-S. Li, J. Ruths, and D. Stefanatos, "A pseudospectral method for optimal control of open quantum systems," *J. Chem. Phys.*, vol. 131, p. 164110, 2009.

[22] J.-S. Li, "Control of a network of spiking neurons," in *Proc. 8th IFAC Symp. Nonlinear Control Syst.*, Sep. 2010.

[23] K. Beauchard, J.-M. Coron, and P. Rouchon, "Controllability issues for continuous-spectrum systems and ensemble controllability of Bloch equations," *Commun. Math. Phys.*, vol. 296, pp. 525–557, 2010.

[24] K. Malanowski and H. Maurer, "Sensitivity analysis for state constrained optimal control problems," *Discrete Contin. Dyn. Syst.*, vol. 4, no. 2, pp. 241–272, 1998.

[25] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods.* Berlin, Germany: Springer, 2006.

[26] G. Elnagar, M. A. Kazemi, and M. Razzaghi, "The pseudospectral Legendre method for discretizing optimal control problems," *IEEE Trans. Autom. Control*, vol. 40, no. 10, pp. 1793–1793, Oct. 1995.

[27] I. Ross and F. Fahroo, "Legendre pseudospectral approximations of optimal control problems," in *New Trends in Nonlinear Dynamics and Control, edited by*, W. Kang, Ed. *et al.* Berlin, Germany: Springer, 2003, pp. 327–342.

[28] F. Fahroo and I. Ross, "Costate estimation by a Legendre pseudospectral method," *J. Guid. Control Dynam.*, vol. 24, pp. 270–277, 2001.

[29] G. Szego, *Orthogonal Polynomials*. New York: American Mathematical Society, 1959.

[30] J. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. New York: Dover, 2000.

[31] D. Gottlieb, Y. Hussaini, and S. Orszag, "Theory and applications of spectral methods," in *Spectral Methods for Partial Differential Equations*, R. Voigt, D. Gottlieb, and M. Y. Hussaini, Eds. Philadelphia: SIAM, 1984, pp. 1–117.

[32] F. Bloch, "Nuclear induction," *Phys. Rev.*, vol. 70, pp. 460–474, 1946.

[33] S. Conolly, D. Nishimura, and A. Macovski, "Optimal control solutions to the magnetic resonance selective excitation problem," *IEEE Trans. Med. Imag.*, vol. MI-5, no. 2, pp. 106–115, Jun. 1986.

[34] R. Ruskov and A. Korotkov, "Spectrum of qubit oscillations from generalized Bloch equations," *Phys. Rev. B*, vol. 67, p. 075303, 2003.

[35] G. A. Morris and R. Freeman, "Enhancement of nuclear magnetic resonance signals by polarization transfer," *J. Amer. Chem. Soc.*, vol. 101, pp. 760–762, 1979.

[36] G. S. Payne and M. O. Leach, "Surface-coil polarization transfer for monitoring tissue metabolism *in vivo*," *Magn. Reson. Med.*, vol. 43, pp. 510–516, 2000.

[37] M. Möttöen, R. Sousa, J. Zhang, and K. B. Whaley, "High fidelity one-qubit operations under random telegraph noise," *Phys. Rev. A*, vol. 73, p. 022332, 2006.

[38] Q. Gong, W. Kang, and M. Ross, "A pseudospectral method for the optimal control of constrained feedback linearizable systems," *IEEE Trans. Autom. Control*, vol. 51, no. 7, pp. 1115–1129, Jul. 2006.

[39] W. Kang, "The rate of convergence for a pseudospectral optimal control method," in *Proc. 47th IEEE Conf. Decision Control*, Cancun, Mexico, Dec. 2008, pp. 521–527.

[40] Q. Gong, M. Ross, W. Kang, and F. Fahroo, "Connections between the covector mapping theorem and convergence of pseudospectral methods for optimal control," *Comput. Optim. Appl*, vol. 41, pp. 307–335, 2008.

[41] A. Isidori, *Nonlinear Control Systems*. London, U.K.: Springer, 1995.

[42] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*. New York: Wiley, 1984.
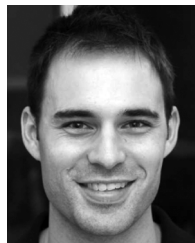
Justin Ruths received the B.S. degree from Rice University, Houston, TX, the M.S. degree from Columbia University, New York, and the Ph.D. degree in systems science and mathematics from Washington University, St. Louis, MO.

He is currently an Assistant Professor at the Singapore University of Technology and Design in the Engineering Systems and Design Pillar. His research interests are in the areas of computational optimal control theory and large-scale complex systems.

Jr-Shin Li (M'06) received the B.S. and M.S. degrees from National Taiwan University, Taipei, and the Ph.D. degree in applied mathematics from Harvard University, Cambridge, MA, in 2006.

He is currently an Assistant Professor in electrical and systems engineering with a joint appointment in the Division of Biology and Biomedical Sciences at Washington University, St. Louis. His research interests are in the areas of control theory and optimization. His current work is on the control of large-scale complex systems with applications ranging from quantum mechanics and neuroscience to bioinformatics.

Dr. Li was a recipient of the NSF Career Award in 2007 as well as the AFOSR Young Investigator Award in 2009.