A pseudospectral method for optimal control of open quantum systems

Jr-Shin Li,[1,a)] Justin Ruths,[1] and Dionisis Stefanatos[2]
[1]*Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, USA*
[2]*Prefecture of Kefalonia, Argostoli, Kefalonia 28100, Greece*

(Received 31 July 2009; accepted 2 October 2009; published online 28 October 2009)

In this paper, we present a unified computational method based on pseudospectral approximations for the design of optimal pulse sequences in open quantum systems. The proposed method transforms the problem of optimal pulse design, which is formulated as a continuous-time optimal control problem, to a finite-dimensional constrained nonlinear programming problem. This resulting optimization problem can then be solved using existing numerical optimization suites. We apply the Legendre pseudospectral method to a series of optimal control problems on open quantum systems that arise in nuclear magnetic resonance spectroscopy in liquids. These problems have been well studied in previous literature and analytical optimal controls have been found. We find an excellent agreement between the maximum transfer efficiency produced by our computational method and the analytical expressions. Moreover, our method permits us to extend the analysis and address practical concerns, including smoothing discontinuous controls as well as deriving minimum-energy and time-optimal controls. The method is not restricted to the systems studied in this article and is applicable to optimal manipulation of both closed and open quantum systems. © *2009 American Institute of Physics*. [doi:10.1063/1.3253796]

I. INTRODUCTION

The problem of relaxation is ubiquitous in all applications involving coherent control of quantum mechanical phenomena. In these applications, the quantum system of interest interacts with its environment (open quantum system) and relaxes back to some equilibrium state.[1] This relaxation effect leads to degraded signal recovery and, in turn, to the loss of experimental information. Optimal manipulation of open quantum systems in such a way as to produce desired evolutions while minimizing relaxation losses has been a long standing and challenging problem in the area of quantum control.

Various methods employing optimization techniques and principles of optimal control have been developed for the design of pulse sequences that can be used to manipulate quantum systems in an optimal manner. However, a large majority of them are limited to deal with closed quantum systems.[2–8] Recently, relaxation-optimized pulse sequences that maximize the performance of open quantum systems have emerged. In some simple cases, for example, maximizing polarization transfer between a pair of coupled spins in the presence of relaxation, the optimal pulses have been derived analytically using optimal control theory.[9–11] For more general cases, gradient ascent algorithms were proposed to optimize pulse sequences for optimally steering the dynamics of coupled nuclear spins.[12–17] These algorithms, while successful, rely heavily on the computation of an analytic expression for the system evolution propagator and gradients as well as a large number of discretizations over which to evolve the system. This results in expensive computational power and gradient ascent algorithms, in general, inherit slow linear convergence rates.[18]

In this article, we present a unified computational method for optimal pulse sequence design based on pseudospectral approximations. This paper is organized as follows. In Sec. II, we formulate optimal control problems in open quantum systems and introduce pseudospectral methods. In Sec. III, we present several examples to demonstrate the robustness of pseudospectral methods for optimal pulse sequence design. The systems in the examples have been thoroughly studied in our previous work or by others.

II. PSEUDOSPECTRAL METHODS FOR OPEN QUANTUM SYSTEMS

For an open quantum system, the evolution of its density matrix is not unitary. In many applications of interest, the environment can be approximated as an infinite thermostat, the state of which never changes. Under this assumption, the so-called Markovian approximation, it is possible to write the evolution of the density matrix of an open system (master equation) alone in the Lindblad form[19]

$$\dot{\rho} = -i[H(t),\rho] - L(\rho) \quad (\hbar = 1), \tag{1}$$

where $H(t)$ is the system Hamiltonian that generates unitary evolution while the term $L(\rho)$ models relaxation (nonunitary dynamics). The general form of L is

$$L(\cdot) = \sum_{\alpha,\beta} k_{\alpha\beta}[V_\alpha,[V_\beta^\dagger,\cdot]], \tag{2}$$

where $V_{\alpha,\beta}$ are operators that represent various relaxation mechanisms and $k_{\alpha\beta}$ are coefficients that depend on the

a)Electronic mail: jsli@seas.wustl.edu.

physical parameters of the system. The Hamiltonian $H(t)$ has the general form

$$H(t) = H_f + \sum_{i=1}^{m} u_i(t) H_i, \tag{3}$$

where H_f is the free evolution Hamiltonian and $\Sigma_{i=1}^{m} u_i(t) H_i$ is the so-called control Hamiltonian. The latter is used to manipulate the open quantum system by application of electromagnetic pulses $u_i(t)$ of appropriate shape and frequency.

A typical problem of controlling an (finite-dimensional) open quantum system has the following form: starting from some initial state $\rho(0)$ at $t=0$, find optimal pulses $u_i(t)$, $0 \leq t \leq T$, that bring the final density matrix $\rho(T)$ at $t=T$ as "close" as possible to some target operator O. More precisely, find $u_i(t)$ that maximize the final expectation value of O, $\langle O \rangle(T) = \mathrm{trace}\{\rho(T) O\}$. Using the master Eq. (1) and the relation $d\langle O \rangle / dt = \mathrm{trace}\{\dot{\rho} O\}$ that holds for a time-independent operator, we can find a system of ordinary differential equations that describe the time evolution of an open system for the desired transfer $\rho(0) \rightarrow O$,[20]

$$\dot{x} = \left[\mathcal{H}_f + \sum_{i=1}^{m} u_i(t) \mathcal{H}_i \right] x, \tag{4}$$

where $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ is the state vector whose elements are expectation values of the operators participating in the transfer (for example, usually $x_n(t) = \langle O \rangle(t)$), $\mathcal{H}_f, \mathcal{H}_i \in \mathbb{R}^{n \times n}$ are square matrices corresponding to operators H_f, H_i under a fixed basis of the state space (Hilbert space) and $t \in [0,T]$. This gives rise to an optimal control problem that starting from an initial state $x(0)$ (which is related to $\rho(0)$), find the controls $u_i(t)$, $t \in [0,T]$, that maximize $x_n(T) = \langle O \rangle(T)$ subject to the system evolution equations as in Eq. (4). Specific examples are given in Sec. III.

Practical considerations such as power and time constraints guide us to consider a more general cost function (the quantity that we want to maximize or minimize)

$$\min \varphi(T, x(T)) + \int_0^T \mathcal{L}(x(t), u(t)) dt, \tag{5}$$

where $u = (u_1, u_2, \ldots, u_m)^T$ is the control vector, φ is the terminal cost depending on the final state at the terminal time $t=T$, and $\mathcal{L}$ is the running cost depending on the time history of the state and control variables, x and u. For example, if $\varphi = -x_n(T)$ and $\mathcal{L}=0$, then Eq. (5) is equivalent to maximizing $x_n(T)$ as mentioned above, while if $\varphi=0$ and $\mathcal{L} = \Sigma_{i=1}^{m} u_i^2$, then Eq. (5) is equivalent to minimizing the energy of the pulses. In many cases of application, not only the initial state $x(0)$ can be specified but also other end point constraints may be imposed. They can be expressed in a compact form as

$$e(x(0), x(T)) = 0. \tag{6}$$

Additionally, constraints on the state and control variables satisfied along the path of the system may be imposed, such as the amplitude constraints where $|u_i(t)| \leq M$, for all $t \in [0,T]$, where M is the maximum amplitude of the pulses. Such constraints can be expressed as

$$g(x(t), u(t)) \leq 0. \tag{7}$$

This class of optimal control problems of bilinear systems (linear in both state and control) described in Eq. (4) is, in general, analytically intractable. However, they can be efficiently solved by pseudospectral methods.

Spectral methods involve the expansion of functions in terms of orthogonal polynomial basis functions on the domain $[-1,1]$. Using such a basis leads to *spectral accuracy*, namely, the kth coefficient of the expansion decays faster than any inverse power of k,[21] which is analogous to the Fourier series expansion for periodic functions. This property of rapid decay from spectral methods is adapted to solve optimal control problems such as the one described above. It permits the use of relatively low order polynomials to approximate the control and state trajectory functions, $u(t)$ and $x(t)$.

Since the support of the orthogonal polynomial bases is on the interval $[-1,1]$, we first transform the optimal control problem from the time domain $t \in [0,T]$ to $\tau \in [-1,1]$ using the simple affine transformation,

$$\tau(t) = \frac{2t - T}{T}.$$

In a redundant use of notation, we make this transition and reuse the same time variable t. The transformed optimal control problem can now be written as

$$\begin{aligned} &\min \varphi(1, x(1)) + \frac{T}{2} \int_{-1}^{1} \mathcal{L}(x(t), u(t)) dt, \\ &\text{s.t. } \dot{x} = \frac{T}{2} \left[\mathcal{H}_f + \sum_{i=1}^{m} u_i(t) \mathcal{H}_i \right] x, \\ &\quad e(x(-1), x(1)) = 0, \\ &\quad g(x(t), u(t)) \leq 0. \end{aligned} \tag{8}$$

Pseudospectral methods were developed to solve partial differential equations and recently adapted to solve optimal control problems.[22–26] Several of the concepts involved in pseudospectral methods have previously found use in areas of chemical physics, such as implementing Lagrange interpolating polynomials based on Gauss–Lobatto quadrature to enforce boundary conditions in quantum scattering problems.[27] Pseudospectral approximations are a spectral collocation (or interpolation) method in which the differential equation describing the state dynamics is enforced at specific nodes. Spectral collocation is motivated by the Chebyshev equioscillation theorem[28] which states that the best Nth order approximating polynomial $p_N^*(f)$ to a continuous function f on the interval $[-1,1]$ is an interpolating polynomial, as evaluated by the uniform norm,

$$\| f - p_N^*(f) \|_\infty = \min_{p \in \mathbb{P}_N} \| f - p \|_\infty, \tag{9}$$

where $\mathbb{P}_N$ is the space of all polynomials of degree at most N. Since any Nth order interpolating polynomial can be represented in terms of the Lagrange basis functions (or Lagrange

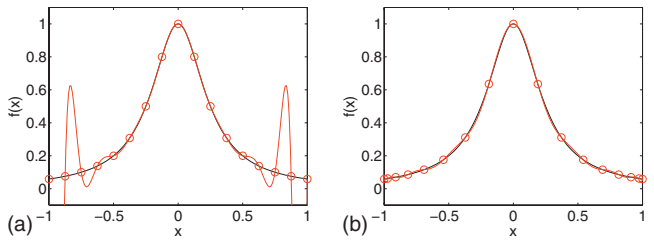

FIG. 1. The $N=16$ order interpolation of the function $f(x)=1/(16x^2+1)$ based on a uniform grid [panel (a)] demonstrates the Runge phenomenon whereas the interpolation based on the LGL grid [panel (b)] does not.

polynomials), we use these functions to express the interpolating approximations of the continuous state and control functions, $x(t)$ and $u(t)$, as in model (8). Given a grid of $N+1$ interpolation nodes within $[-1,1]$, $\Gamma=\{t_0<t_1<\cdots<t_N\}$, the Lagrange polynomials $\{\ell_k\}\in\mathbb{P}_N$, $k\in\{0,1,\ldots,N\}$, are constructed by

$$\ell_k(t)=\prod_{\substack{i=0\\ i\neq k}}^{N}\frac{(t-t_i)}{(t_k-t_i)},$$

which are characterized by the kth polynomial taking unit value at the kth node of the grid and zero value at all other nodes of the grid, i.e., $\ell_k(t_i)=\delta_{ki}$, where δ_{ki} is the Kronecker delta function.[29] Note that Lagrange polynomials form an orthogonal basis with respect to the discrete inner product $\langle p,q\rangle=\Sigma_{k=0}^{N}p(t_k)q(t_k)$.

With these tools, we can now write the Nth order interpolating approximations of the state trajectory and control functions with respect to a given grid Γ of $N+1$ nodes as

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

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

where $\bar{x}_k$ and $\bar{u}_k$ are not only the coefficients of the expansions but also the function values at the kth node due to the definition of the Lagrange polynomials.[22] Because these Lagrange polynomials are constructed based on the choice of these nodes, the approximations made with this basis as in Eqs. (10) and (11) are sensitive to the choice of the nodes. For an arbitrary selection of nodes, as the order of approximation N gets large, Runge phenomenon may occur; that is, there are increasingly larger spurious oscillations near the end points of the $[-1,1]$ domain,[30] as shown in Fig. 1. A selection of Gauss-type nodes with quadratic spacing toward the end points suppresses such oscillation between the interpolation nodes and greatly increases the accuracy of the approximation.[31] It has been shown that for a fixed $N>0$ and a norm given by Eq. (9), Gauss-type nodes are asymptotically close to optimal for interpolating a continuous function over the domain $[-1,1]$.[32]

In order to maintain the advantages of a spectral method while using collocation, we write the Lagrange polynomials in terms of orthogonal polynomials. We choose to focus on the Legendre polynomials which are orthogonal in $L_2[-1,1]$ defined with a weighted inner product,

$$\langle f,g\rangle=\int_{-1}^{1}f(t)g(t)w(t)dt,$$

with a constant weight function $w(t)=1$, $\forall t\in[-1,1]$, where $f,g\in L_2[-1,1]$. Implementing the pseudospectral method with the orthogonal Legendre polynomials determines the grid to be Legendre–Gauss nodes which are the roots of $\dot{L}_N(t)$, the derivative of the Nth order Legendre polynomial. To enforce the method at the end points, we use the Legendre–Gauss–Lobatto (LGL) nodes, which include $t_0=-1$ and $t_N=1$, i.e., $\Gamma^{\mathrm{LGL}}=\{t_j:\dot{L}_N(t)|_{t_j}=0,j=1,\ldots,N-1\}\cup\{-1,1\}$. Then, the Lagrange polynomials $\ell_k(t)$ can be expressed with respect to Γ^{LGL} as[29]

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

where $\{t_k\}\in\Gamma^{\mathrm{LGL}}$, $k=0,1,\ldots,N$.

From the interpolation as in Eq. (10), we have

$$\frac{d}{dt}I_N\,x(t)=\sum_{k=0}^{N}\bar{x}_k\dot{\ell}_k(t).$$

Using special recursive identities for the derivative of Legendre polynomials,[26] we have at the LGL nodes $t_j\in\Gamma^{\mathrm{LGL}}$, $j=0,1,\ldots,N$,

$$\frac{d}{dt}I_N\,x(t_j)=\sum_{k=0}^{N}\bar{x}_k\dot{\ell}_k(t_j)=\sum_{k=0}^{N}D_{jk}\bar{x}_k, \tag{12}$$

where D_{jk} are jkth elements of the constant $(N+1)\times(N+1)$ differentiation matrix D defined by[33]

$$D_{jk}=\begin{cases}\dfrac{L_N(t_j)}{L_N(t_k)}\dfrac{1}{t_j-t_k}, & j\neq k\\ -\dfrac{N(N+1)}{4}, & j=k=0\\ \dfrac{N(N+1)}{4}, & j=k=N\\ 0, & \text{otherwise}.\end{cases} \tag{13}$$

In addition, the integral cost functional in the optimal control problem (8) can be approximated by the Gaussian quadrature. In particular, LGL quadrature is used to enforce end point conditions and defined as

$$\int_{-1}^{1}f(t)dt=\sum_{i=1}^{N}f(t_i)w_i,\quad w_i=\int_{-1}^{1}\ell_i(t)dt, \tag{14}$$

which is exact for $f\in\mathbb{P}_{2N-1}$ when $\{t_i\}\in\Gamma^{\mathrm{LGL}}$.[21] Therefore, the choice of LGL nodes not only achieves close to optimal interpolation error by preventing increasingly spurious oscillations as N gets large but also ensures the accuracy of the numerical integration.

Compiling Eqs. (10), (12), and (14) we can convert the optimal control problem as in Eq. (8) into the following finite-dimensional constrained minimization problem by discretizing the states and controls with an interpolation

scheme, representing the differential equation through recursive definition of spectral derivatives, and expressing integral terms with the Gaussian quadrature,

$$\min \varphi(T,\bar{x}_N) + \frac{T}{2}\sum_{i=0}^{N} \mathcal{L}(\bar{x}_i,\bar{u}_i)w_i,$$

$$\text{s.t.} \sum_{k=0}^{N} D_{jk}\bar{x}_k = \frac{T}{2}\left[\mathcal{H}_f + \sum_{i=1}^{m} \bar{u}_{ij}\mathcal{H}_i\right]\bar{x}_j,$$

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

$$g(\bar{x}_j,\bar{u}_j) \le 0, \quad \forall\, j \in \{0,1,\ldots,N\},$$

where $\bar{u}_{ij}$, $\bar{x}i=1,\ldots,m$, are components of the vector $\bar{u}_j$ denoting the value of the control function u_i at the jth LGL node t_j, namely, $\bar{u}_j=(\bar{u}_{1j},\ldots,\bar{u}_{mj})^T=(u_1(t_j),\ldots,u_m(t_j))^T$. Solvers for this type of constrained minimization problem are readily available and straightforward to implement.

III. EXAMPLES FROM NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY IN LIQUIDS

In this section we show the robustness and efficiency of the pseudospectral method by applying it to a series of optimal control problems on open quantum systems that arise in NMR spectroscopy of proteins in liquids. These control problems were selected because analytical expressions for their optimal solutions have been derived in literature,[9,10,17] making them well suited for testing the performance of the pseudospectral method on open quantum systems.

A. Pair of coupled heteronuclear spins

The first open quantum system from liquid state NMR that we consider is an isolated pair of heteronuclear spins 1/2 (spins that belong to different nuclear species), denoted as I_1 (for example ^{1}H) and I_2 (for example, ^{13}C or ^{15}N), with a scalar coupling J.[20] In a doubly rotating frame, which rotates with each spin at its resonance (Larmor) frequency, the free evolution Hamiltonian for this system is $H_f=2JI_{1z}I_{2z}$, where $I_{1z}=\sigma_{1z}/2$, $I_{2z}=\sigma_{2z}/2$ and σ_{1z},σ_{2z} are the Pauli spin matrices for spins I_1 and I_2, respectively. Note that this Hamiltonian is valid in the so-called weak coupling limit, where the resonance frequencies of the spins satisfy $|\omega_1-\omega_2| \gg J$ and thus the Heisenberg coupling $(\boldsymbol{I}_1\cdot\boldsymbol{I}_2)$, which is the characteristic indirect coupling between spins in isotropic liquids, can be approximated by the scalar coupling $(I_{1z}I_{2z})$.[20]

The most important relaxation mechanisms in NMR spectroscopy in liquid solutions are due to dipole-dipole (DD) interaction and chemical shift anisotropy (CSA), as well as their interference effects, i.e., DD-CSA cross correlation.[20] We initially consider the spin system without cross-correlated relaxation.

1. Spin pair without cross-correlated relaxation

Here we consider the open quantum system with only DD and CSA relaxation ignoring the cross-correlated relaxation. This case approximates, for example, the situation for deuterated and ^{15}N-labeled proteins in H_2O at moderately high magnetic fields (e.g., 10 T), where the ^{1}H–^{15}N spin pairs are isolated and CSA relaxation is small. Furthermore, we focus on slowly tumbling molecules in the so-called spin diffusion limit.[34] In this case, longitudinal relaxation rates $(1/T_1)$ are negligible compared to transverse relaxation rates $(1/T_2)$.[34]

For this coupled two-spin system, the free evolution of the density matrix ρ in the doubly rotating frame is given by the following master equation:[9]

$$\dot{\rho} = -iJ[2I_{1z}I_{2z},\rho] - k_{\mathrm{DD}}[2I_{1z}I_{2z},[2I_{1z}I_{2z},\rho]] - k^1_{\mathrm{CSA}}[I_{1z},[I_{1z},\rho]] - k^2_{\mathrm{CSA}}[I_{2z},[I_{2z},\rho]], \tag{15}$$

where J is the scalar coupling constant, k_{DD} is the DD relaxation rate, and $k^1_{\mathrm{CSA}},k^2_{\mathrm{CSA}}$ are CSA relaxation rates for spins I_1,I_2, respectively. These relaxation rates depend on various physical parameters of the system, such as the gyromagnetic ratios of the spins, the internuclear distance, and the correlation time of the rotational tumbling.[20]

One problem of interest in NMR experiments is to find the pulses (controls), $\omega_x(t)$ and $\omega_y(t)$ applied in the x and y directions, respectively, for optimal polarization transfer $I_{1z}\rightarrow I_{2z}$ from one spin to the other. This transfer is suitably done in two steps: $I_{1z}\rightarrow 2I_{1z}I_{2z}\rightarrow I_{2z}$. Since these two steps are symmetric, the optimal controls for the second step are symmetric to those of the first one. Thus, we only need to concentrate on the first step and the objective is to maximize the transfer $I_{1z}\rightarrow 2I_{1z}I_{2z}$. In other words, starting from the initial state $\rho(0)=I_{1z}$, we tend to maximize the final expectation value of the target operator $O=2I_{1z}I_{2z}$, i.e., $\langle 2I_{1z}I_{2z}\rangle(T)=\mathrm{trace}\{\rho(T)2I_{1z}I_{2z}\}$, where T is the final time of the experiment and $\rho(T)$ is controlled by $\omega_x(t)$ and $\omega_y(t)$. Using the master Eq. (15), we can find differential equations that describe the time evolution of the expectations of the operators participating in the desired transfer as presented in Sec. II. The corresponding equations in matrix form are

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\\ \dot{x}_4\end{bmatrix} = \begin{bmatrix}0 & -u_1 & 0 & 0\\ u_1 & -\xi & -1 & 0\\ 0 & 1 & -\xi & -u_2\\ 0 & 0 & u_2 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}, \tag{16}$$

where $x_1=\langle I_{1z}\rangle$, $x_2=\langle I_{1x}\rangle$, $x_3=\langle 2I_{1y}I_{2z}\rangle$, $x_4=\langle 2I_{1z}I_{2z}\rangle$, and $\xi=(k_{\mathrm{DD}}+k^1_{\mathrm{CSA}})/J$, and the controls $u_1(t)=\omega_y(t)/J$ and $u_2(t)=\omega_x(t)/J$ are the normalized (with respect to J) transverse components of the applied magnetic field. Note that the above system (16) has the bilinear form as shown in Eq. (4).

Consequently, we now arrive at an optimal control problem for the transfer $I_{1z}\rightarrow 2I_{1z}I_{2z}$ that is to find $u_1(t)$ and $u_2(t)$, $0\le t\le T$, such that starting from $x(0)=(1,0,0,0)^T$, $x_4(T)$ is maximized subject to the evolution Eq. (16). This problem has been solved analytically and the resulting analytical pulse was denoted as ROPE.[9] It is shown there that the maximum achievable value of x_4, i.e., the efficiency η_1 of the transfer is given as a function of parameter ξ by

$$\eta_1 = \sqrt{\xi^2+1} - \xi. \tag{17}$$

Using the pseudospectral method presented in this paper, we calculated numerically optimal controls $u_1(t),u_2(t)$ for

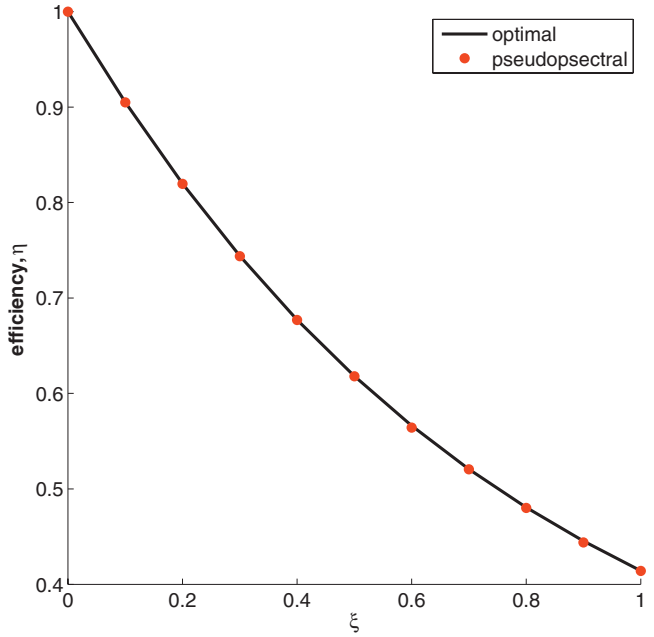

FIG. 2. The efficiency of the transfer $x_1 \rightarrow x_4$ in system (16) achieved by the pseudospectral method, as a function of the relaxation parameter ξ in the range [0,1]. The theoretically calculated maximum efficiency given by Eq. (17) is also shown.

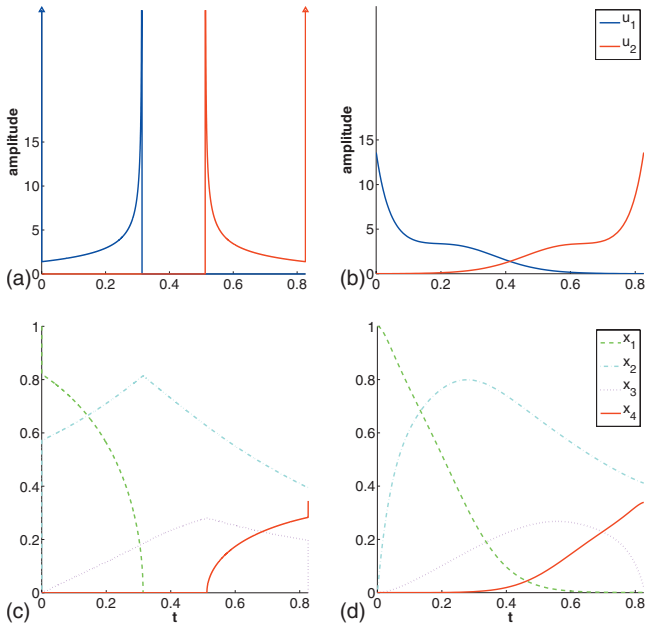

FIG. 3. Pseudospectral controls [panel (b)] and state trajectories [panel (d)] are compared to analytic ROPE (Ref. 9) controls (panel a) and trajectories [panel (c)] for $\xi=1$. Each of the hard pulses at $t=0$ and $t=T$ [panel (a)] corresponds to a 35° rotation and transfer the state from $x(0^-)=[1,0,0,0]^T$ to $x(0^+)=[\cos 35°, \sin 35°, 0, 0]^T$ and from $x(T^-)=[0, x_2(T), \eta \sin 35°, \eta \cos 35°]^T$ to $x(T^+)=[0, x_2(T), 0, \eta]^T$, respectively, in near instantaneous time.

various values of ξ in the range $[0,1]$ to maximize the corresponding achievable values of $x_4(T)$. The method was implemented in MATLAB using the third party KNITRO nonlinear programming solver from Ziena optimization. The problem was approximated using 25 ($N=24$) nodes and with the terminal time free to vary, but with a maximum time of $T=10$. A unified and general method should not use any prior knowledge so the solver was given an arbitrary initial guess for the controls. In this case, we take $u_1(t)=u_2(t)=1$ and $T=1$. The termination tolerance on the cost function of the solver was set at 1×10^{-8}. In Fig. 2 we plot the value $x_4(T)$ achieved by the pseudospectral method for $\xi \in [0,1]$. For comparison, we also plot the maximum efficiency given in Eq. (17). The excellent agreement shows the efficiency of the method to approximate optimal solutions.

Another clear advantage of the pseudospectral method well illustrated by this problem is that the calculated control pulses are smooth functions. Figure 3(a) shows the discontinuities of the analytically derived optimal pulses.[9] In particular, notice the high-amplitude spikes at the beginning and end of each component of the analytical pulse. Such discontinuities can be challenging, if not impossible, to implement in practice and high amplitudes can be hazardous for the experiment sample, equipment, and human subjects as in magnetic resonance imaging (MRI). The pulse amplitude derived by the pseudospectral method, shown in Fig. 3(b), is easily implementable and maintains low values despite achieving transfer efficiencies within 1×10^{-3} of the theoretical optimal values. The pseudospectral pulse shown in Fig. 3(b) is attained from an optimization that minimizes energy while maintaining a desired transfer efficiency (in practice, this can be set to a desired efficiency of 1 and the solver will find the least infeasible solution). Therefore, not only is the pseudospectral pulse without discontinuities but it also accomplishes the transfer with 45% less energy than the ROPE pulse.

2. Spin pair with cross-correlated relaxation

If DD-CSA cross-correlated relaxation cannot be neglected, the master equation as in Eq. (15) is then modified to incorporate it as[10]

$$\dot{\rho} = -iJ[2I_{1z}I_{2z},\rho] - k_{\mathrm{DD}}[2I_{1z}I_{2z},[2I_{1z}I_{2z},\rho]] - k^1_{\mathrm{CSA}}[I_{1z},[I_{1z},\rho]] - k^2_{\mathrm{CSA}}[I_{2z},[I_{2z},\rho]] - k^1_{\mathrm{DD/CSA}}[2I_{1z}I_{2z},[I_{1z},\rho]] - k^2_{\mathrm{DD/CSA}}[2I_{1z}I_{2z},[I_{2z},\rho]],$$

where $k_{\mathrm{DD}}, k^1_{\mathrm{CSA}}, k^2_{\mathrm{CSA}}$ are autorelaxation rates due to DD relaxation, CSA relaxation of spin I_1, CSA relaxation of spin I_2, and $k^1_{\mathrm{DD/CSA}}, k^2_{\mathrm{DD/CSA}}$ are cross-correlation rates of spins I_1 and I_2 due to interference effects between DD and CSA relaxation mechanisms.

Using this master equation, we can find the following equations for the ensemble averages:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 & -u_1 & u_2 & 0 & 0 & 0 \\ u_1 & -\xi_a & 0 & -1 & -\xi_c & 0 \\ -u_2 & 0 & -\xi_a & -\xi_c & 1 & 0 \\ 0 & 1 & -\xi_c & -\xi_a & 0 & -u_2 \\ 0 & -\xi_c & -1 & 0 & -\xi_a & u_1 \\ 0 & 0 & 0 & u_2 & -u_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}, \tag{18}$$

where $x_1=\langle I_{1z}\rangle$, $x_2=\langle I_{1x}\rangle$, $x_3=\langle I_{1y}\rangle$, $x_4=\langle 2I_{1y}I_{2z}\rangle$, $x_5=\langle 2I_{1x}I_{2z}\rangle$, $x_6=\langle 2I_{1z}I_{2z}\rangle$, $\xi_a=(k_{\mathrm{DD}}+k^1_{\mathrm{CSA}})/J$, $\xi_c=k^1_{\mathrm{DD/CSA}}/J$,

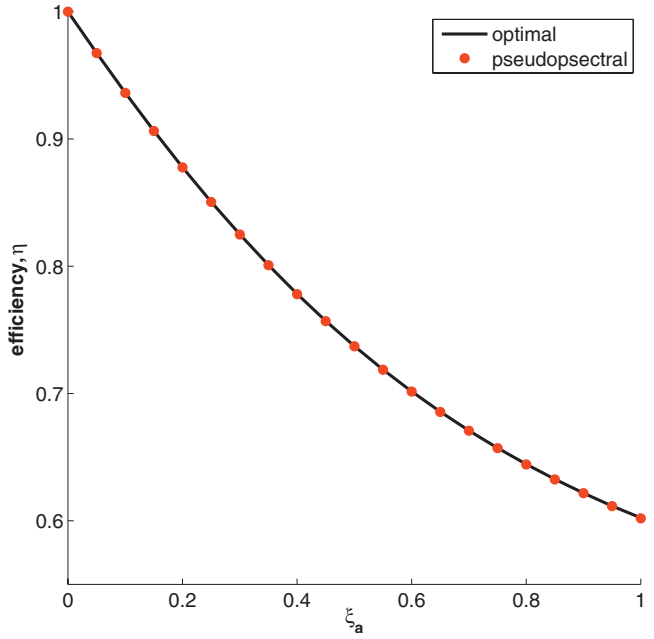

FIG. 4. The efficiency of the transfer $x_1 \rightarrow x_6$ in system (18) achieved by the pseudospectral method, as a function of the relaxation parameter ξ_a in the range [0,1], with $\xi_c=0.75\xi_a$. The theoretically calculated maximum efficiency given by Eq. (19) is also shown.

and $u_1(t), u_2(t)$ are the available controls as before. Starting from $x(0)=(1,0,0,0,0,0)^T$, we want to design $u_1(t)$ and $u_2(t)$ that maximize $x_6(T)$ subject to Eq. (18).

This problem has also been solved analytically and the analytical pulse was denoted as CROP.[10] It is shown there that the maximum achievable value of x_6, i.e., the efficiency η_2 of the transfer, is given by the same formula as before,

$$\eta_2 = \sqrt{\xi^2 + 1} - \xi, \tag{19}$$

but now

$$\xi = \sqrt{\frac{\xi_a^2 - \xi_c^2}{1 + \xi_c^2}}. \tag{20}$$

Using the pseudospectral method introduced in this paper, we calculated numerically optimal controls $u_1(t), u_2(t)$ for various values of ξ_a over [0,1] and with $\xi_c=0.75\xi_a$, to maximize the corresponding achievable values of $x_6(T)$. Using the same MATLAB program and KNITRO solver, the optimal control problem was approximated by 25 ($N=24$) nodes with a free terminal time (maximum $T=5$). A similar constant initial guess was used and the cost function tolerance was set to 1×10^{-5}. In Fig. 4 we plot the values of $x_6(T)$ achieved by the pseudospectral method and the maximum efficiency given in Eq. (19). Again, an excellent agreement is observed. The CROP and pseudospectral control pulse components are plotted in Fig. 5. The pseudospectral pulse in Fig. 5(b) is optimized as a time-optimal pulse that achieves a transfer efficiency within 1×10^{-3} of the analytical optimal value $\eta=0.6022$ with a duration $T=2.2389$, 44% shorter than the CROP pulse given the same bound in control amplitude, further highlighting the flexibility of this numerical method.

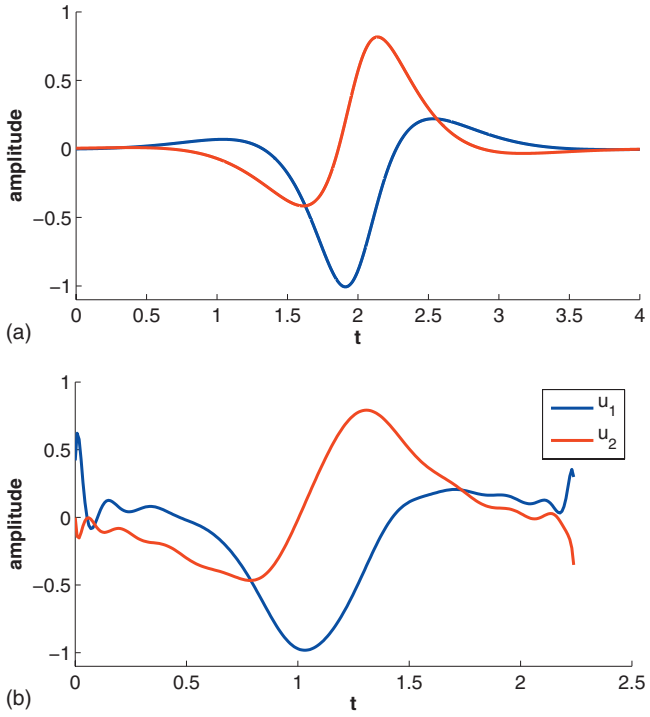

FIG. 5. The CROP pulse [panel (a)] is compared to the time-optimal pseudospectral control pulse [panel (b)] for $\xi_a=1$ and $\xi_c=0.75$. The pseudospectral control pulse achieves a transfer efficiency within 1×10^{-3} of the analytical optimal value $\eta=0.6022$ with a duration $T=2.2389$, 44% shorter than the CROP pulse given the same bound in control amplitude.

B. Three spin chain

The next open quantum system that we consider is a three spin chain (spins I_1, I_2, I_3) with equal scalar couplings between nearest neighbors. In a suitably chosen (multiple) rotating frame, which rotates with each spin at its resonance (Larmor) frequency, the Hamiltonian H_f that governs the free evolution of the system is $H_f=\sqrt{2}J(I_{1z}I_{2z}+I_{2z}I_{3z})$. The common coupling constant is written in the form $\sqrt{2}J$ for normalization reasons. As in the first example described in Sec. III A 1, we neglect cross-correlated relaxation and focus on slowly tumbling molecules in the spin diffusion limit. The corresponding master equation is

$$\begin{aligned}\dot{\rho} = &- i\sqrt{2}J[I_{1z}I_{2z} + I_{2z}I_{3z}, \rho] - k_{\mathrm{DD}}[2I_{1z}I_{2z}, [2I_{1z}I_{2z}, \rho]] \\ &- k_{\mathrm{DD}}[2I_{2z}I_{3z}, [2I_{2z}I_{3z}, \rho]] - k_{\mathrm{DD}}[2I_{3z}I_{1z}, [2I_{3z}I_{1z}, \rho]] \\ &- k^1_{\mathrm{CSA}}[I_{1z}, [I_{1z}, \rho]] - k^2_{\mathrm{CSA}}[I_{2z}, [I_{2z}, \rho]] \\ &- k^3_{\mathrm{CSA}}[I_{3z}, [I_{3z}, \rho]].\end{aligned} \tag{21}$$

For this system we examine the polarization transfer from spin I_1 to spin I_3, $I_{1z} \rightarrow I_{3z}$. This transfer is suitably done in three steps, $I_{1z} \rightarrow 2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z} \rightarrow I_{3z}$. The first and last steps are similar to the previously examined spin transfer, thus we concentrate on the intermediate step $2I_{1z}I_{2z} \rightarrow 2I_{2z}I_{3z}$, which corresponds to $\rho(0)=2I_{1z}I_{2z}$ and $O=2I_{2z}I_{3z}$. Similarly, from the master Eq. (21), we can derive the associated differential equations which describe the time evolution of the expectation values of operators participating in this transfer,

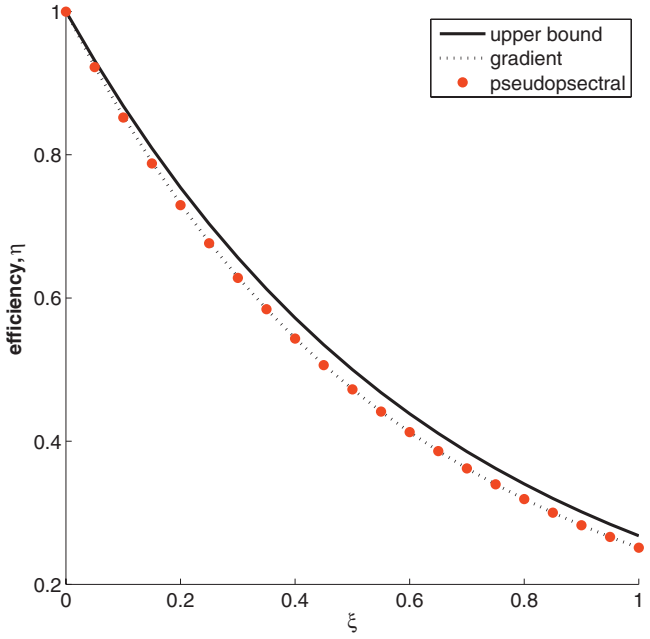

FIG. 6. The efficiency for transfer $x_1 \rightarrow x_5$ in system (22) achieved by the pseudospectral method, as a function of the relaxation parameter ξ in the range [0,1]. The corresponding numerical results achieved by a state-of-the-art gradient ascent algorithm as well as the theoretically calculated upper bound for the maximum efficiency, given by Eq. (23), are also shown.

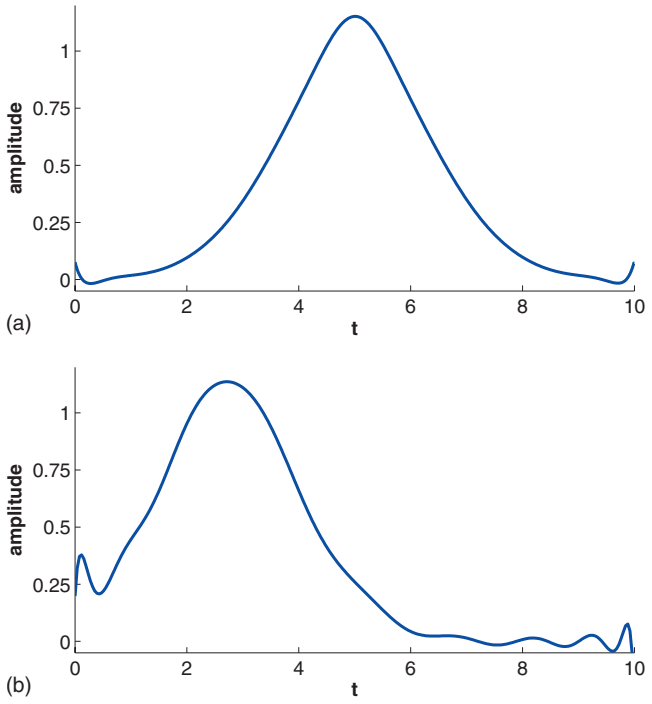

FIG. 7. The gradient derived control pulse [panel (a)] is compared to the pseudospectral control pulse [panel (b)] for $\xi=1$.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} 0 & -u & 0 & 0 & 0 \\ u & -\xi & -1 & 0 & 0 \\ 0 & 1 & -\xi & -1 & 0 \\ 0 & 0 & 1 & -\xi & -u \\ 0 & 0 & 0 & u & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \tag{22}$$

where $x_1=\langle 2I_{1z}I_{2z}\rangle$, $x_2=\langle 2I_{1z}I_{2x}\rangle$, $x_3=\langle\sqrt{2}(2I_{1z}I_{2y}I_{3z}+I_{2y}/2)\rangle$, $x_4=-\langle 2I_{2x}I_{3z}\rangle$, $x_5=\langle 2I_{2z}I_{3z}\rangle$, and $\xi=(2k_{\mathrm{DD}}+k_{\mathrm{CSA}}^2)/J$.[17] Transverse relaxation rate is normalized with respect to the scalar coupling J and the normalized relaxation parameter is ξ. The control function $u(t)=\omega_y(t)/J$ is the y-component of the applied magnetic field. The x-component creates an equivalent path from x_1 to x_5 and needs not be considered.[17]

The corresponding optimal control problem is to find, starting from $x(0)=(1,0,0,0,0)^T$, $u(t)$ that maximizes $x_5(T)$ subject to Eq. (22). It has been shown in Ref. 17 that a strict upper bound, η_3, of the maximum achievable value of x_5 is characterized by ξ,

$$\eta_3 = \frac{(\sqrt{\xi^2+2}-\xi)^2}{2}. \tag{23}$$

It is important to note that this upper bound was derived from an augmented version of the system in Eq. (22), which can be studied analytically. This bound was used in Ref. 17 to evaluate the performance of a gradient ascent algorithm for calculating optimal $u(t)$ for system (22). Therefore, the true optimal efficiencies of Eq. (22) are unknown, however, the optimal efficiencies of the augmented system serve as a strict upper bound.

In this paper, the optimal control $u(t)$ was calculated by the pseudospectral method for various values of ξ in the range [0,1] to maximize the corresponding values $x_5(T)$, where the optimal control problem was approximated by 25 ($N=24$) nodes with a free terminal time (maximum $T=10$). A similar constant initial guess was used and the cost function tolerance was set to 1×10^{-6}. The values of $x_5(T)$ achieved by the pseudospectral method are displayed in Fig. 6. The corresponding numerical results by the state-of-the-art gradient ascent algorithm used in Ref. 17 and the analytical efficiency upper bound given in Eq. (23) are also shown in the same figure. Observe the excellent agreement between the efficiencies of the two numerical methods, lower of course than the analytical upper bound. In Fig. 7 we show the gradient control pulse with the pseudospectral control pulse for the case $\xi=1$. It is worth observing that the gradient method, which optimizes over the control variables, uses a discretization of 1000 points on $[0,T]$, leading to $1000\times(1$ control$)=1000$ decision variables. The pseudospectral method uses a discretization with 25 points and optimizes over both the states and controls, i.e., $25\times(5$ states $+1$ control$)=150$ decision variables. This factor on the order of five difference in the number of decision variables of the discretized system is one of the key advantages of the pseudospectral method. In the former two systems where there were two control variables, the gradient method would double the number of decision variables, making a factor on the order of ten difference.

C. Numerical convergence analysis

Here we present convergence results for the pseudospectral optimizations presented in this paper. The convergence property is related to the conditions under which a sequence of discretized optimization solutions, provided existing, converges to the original optimal control solution as the number of nodes (discretizations) increases. Convergence rates of

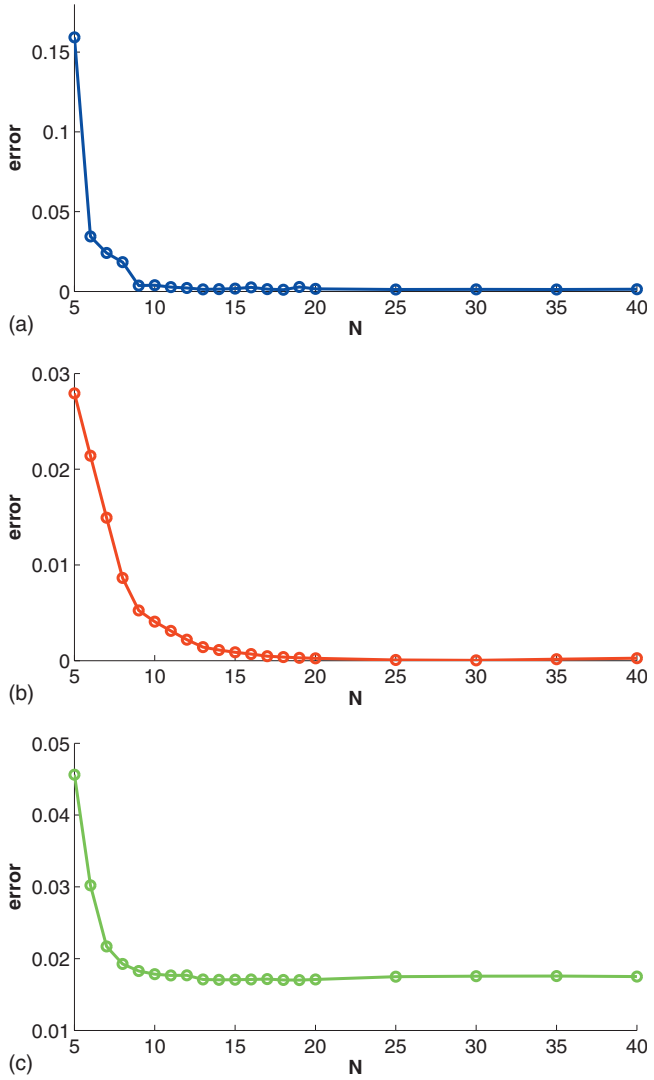

FIG. 8. Numerical results are shown for the convergence of the pseudospectral transfer efficiency to the optimal transfer efficiency in the cases of spin pair, $\xi=1$ [panel (a)]; spin pair with cross-correlated relaxation, $\xi_a=1$, $\xi_c=0.75$ [panel (b)]; three spin chain, $\xi=1$ [panel (c)]. The error in the spin pair examples is the difference between the pseudospectral transfer efficiency and the optimal efficiency, whereas in the three spin chain case the error is the difference between the pseudospectral transfer efficiency and the upper bound, which we do not expect to converge to zero.

pseudospectral methods as applied to optimal control problems have so far been derived in only the case, in which the nonlinear dynamics are feedback linearizable.[24] However, the bilinear dynamics for modeling quantum control systems as discussed in this article, in general, do not fall into this category. Since analytic convergence results are challenging to identify for general systems, we analyzed the convergence numerically to guarantee that solutions do converge as the number of nodes, N, increases. Figure 8 presents this analysis for each system and confirms a robust convergence. The rapid convergence which is characteristic of the pseudospectral method can also be seen clearly as well as justification that the node selection, $N=24$, is sufficient to accurately solve these illustrative examples.

IV. CONCLUSION

In this paper, we presented a method based on pseudospectral approximations to effectively discretize and solve optimal control problems associated with pulse sequence design for open quantum mechanical systems. Examples from NMR spectroscopy in liquid solutions evidenced the flexibility and efficiency of the proposed methods. In these examples, pseudospectral methods generated pulses that achieve performance similar to that of analytical methods, and it should be noted that these "approximated" optimal pulses found by pseudospectral methods are always smooth. A strength of pseudospectral methods is that they provide a robust technique which is easily extendible and implementable. In addition, it has been shown empirically that these methods have exponential convergence rates,[35] while state-of-the-art algorithms such as gradient methods typically evidence linear convergence. Pseudospectral methods provide a universal tool for solving pulse design problems for dissipative quantum systems coming from different physical contexts (NMR, MRI, Quantum Optics, etc.). Some immediate extensions of the methods presented here include considering the optimal pulse design problem of steering a family of open or closed quantum systems with different dynamics. A concrete example is to consider a family of coupled spin systems where each one is characterized by a different relaxation rate, e.g., ξ can take values from a positive interval $[\xi_1,\xi_2]$. In this case, the optimal control problem would seek to accomplish the maximum transfer, η, over all systems, namely, to maximize $\int_{\xi_1}^{\xi_2}\eta(\xi)d\xi$ subject to the system evolution as in Eq. (16), (18), and (22). Experimental verification of the performance of the pseudospectral pulses is also of keen interest and currently being pursued.

ACKNOWLEDGMENTS

This work was supported by the NSF under Grant No. 0747877.

[1] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2007).
[2] K. Kobzar, T. Skinner, N. Khaneja, S. Glaser, and B. Luy, J. Magn. Reson. **170**, 236 (2004).
[3] K. Kobzar, T. Skinner, N. Khaneja, S. Glaser, and B. Luy, J. Magn. Reson. **194**, 58 (2008).
[4] J. Pauly, P. Le Roux, D. Nishimura, and A. Macovski, IEEE Trans. Med. Imaging **10**, 53 (1991).
[5] S. Conolly, D. Nishimura, and A. Macovski, IEEE Trans. Med. Imaging **MI-5**, 106 (1986).
[6] J. Mao, T. H. Mareci, K. N. Scott, and E. R. Andrew, J. Magn. Reson. **70**, 310 (1986).
[7] N. Khaneja, R. Brockett, and S. J. Glaser, Phys. Rev. A **63**, 032308 (2001).
[8] B. Pryor and N. Khaneja, J. Chem. Phys. **125**, 194111 (2006).
[9] N. Khaneja, T. Reiss, B. Luy, and S. J. Glasser, J. Magn. Reson. **162**, 311 (2003).
[10] N. Khaneja, B. Luy, and S. J. Glaser, Proc. Natl. Acad. Sci. U.S.A. **100**, 13162 (2003).
[11] D. Stefanatos, N. Khaneja, and S. J. Glaser, Phys. Rev. A **69**, 022319 (2004).
[12] N. Khaneja, T. Reiss, C. Kehlet T. S.-Herbruggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).
[13] Y. Ohtsuki, G. Turinici, and H. Rabitz, J. Chem. Phys. **120**, 5509 (2004).
[14] T. E. Skinner, T. Reiss, B. Luy, N. Khaneja, and S. J. Glaser, J. Magn. Reson. **163**, 8 (2003).
[15] N. Khaneja, J.-S. Li, C. Kehlet, B. Luy, and S. J. Glaser, Proc. Natl. Acad. Sci. U.S.A. **101**, 14742 (2004).
[16] D. P. Frueh, T. Ito, J.-S. Li, G. Wagner, S. J. Glaser, and N. Khaneja, J. Biomol. NMR **32**, 23 (2005).
[17] D. Stefanatos, S. J. Glaser, and N. Khaneja, Phys. Rev. A **72**, 062320

(2005).
[18] D. P. Bertsekas, *Nonlinear Programming* (Athena Scientific, Belmont, MA, 1999).
[19] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[20] M. Goldman, *Quantum Description of High-Resolution NMR in Liquids* (Clarendon, Oxford, 1988).
[21] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods* (Springer, Berlin, 2006).
[22] G. Elnagar, M. A. Kazemi, and M. Razzaghi, IEEE Trans. Autom. Control **40**, 1793 (1995).
[23] I. Ross and F. Fahroo, in *New Trends in Nonlinear Dynamics and Control*, edited by W. Kang, M. Xiao, and C. Borges (Springer, Berlin, 2003), p. 327.
[24] Q. Gong, W. Kang, , and I. Ross, IEEE Trans. Autom. Control **51**, 1115 (2006).
[25] F. Fahroo and I. Ross, J. Guid. Control Dynam. **24**, 270 (2001).
[26] P. Williams, ANZIAM J. **47**, C101 (2006).
[27] D. Manolopoulos, Chem. Phys. Lett. **152**, 23 (1988).
[28] P. J. Davis, *Interpolation and Approximation* (Blaisdell, New York, 1963).
[29] G. Szego, *Orthogonal Polynomials* (American Mathematical Society, New York, 1959).
[30] B. Fornberg, *A Practical Guide to Pseudospectral Methods* (Cambridge University Press, New York, 1998).
[31] J. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. (Dover, New York, 2000).
[32] S. Smith, Ann. Math. Informaticae **33**, 109 (2006).
[33] D. Gottlieb, Y. Hussaini, and S. Orszag, in *Spectral Methods for Partial Differential Equations*, edited by R. Voigt, D. Gottlieb, and M. Y. Hussaini (SIAM, Philadelphia, 1984), p. 1.
[34] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon, Oxford, 1987).
[35] L. N. Trefethen, *Spectral Methods in MATLAB* (SIAM, Philadelphia, 2000).