

Preprint of "GRAPE" article:

N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, S. J. Glaser,

**"Optimal Control of Coupled Spin Dynamics:
Design of NMR Pulse Sequences by Gradient Ascent Algorithms",**

J. Magn. Reson. 172, 296-305 (2005)

Abstract

In this paper, we introduce optimal control-related gradient-based numerical algorithms for the design of pulse sequences in NMR spectroscopy. This methodology is used for designing pulse sequences that maximize the coherence transfer between coupled spins in a given specified time, minimize the relaxation effects in a given coherence transfer step or minimize the time required to produce a given unitary propagator, as desired. The application of these gradient ascent pulse engineering (GRAPE) methods to design pulse sequences that are robust to experimentally important parameter variations, such as chemical shift dispersion or rf variations due to imperfections such as rf-inhomogeneity is also explained.

1 Introduction

In applications of NMR spectroscopy it is desirable to have optimized pulse sequences tailored to specific applications. For example, in multi-dimensional NMR experiments one is often interested in pulse sequences which maximize the coherence transfer between coupled spins in a given specified time, minimize the relaxation effects in a given coherence transfer step or minimize the time required to produce a given unitary propagator. From an engineering perspective all these problems are challenges in optimal control [1, 2] where one is interested in tailoring the excitation to a dynamical system to maximize some performance criterion. In this paper we present gradient ascent algorithms for optimizing pulse sequences (control laws) for steering the dynamics of coupled nuclear spins. Similar methods and their variants have been applied in Laser spectroscopy [3, 4, 5, 7]. In NMR, this approach has been used to design band-selective pulses [8, 9, 10], robust broadband excitation and inversion pulses [11, 12, 13]. However, previous studies in NMR were limited to uncoupled spin systems whose dynamics is governed by the Bloch equations. It is important to note that the optimal control principles are standard text book material in applied optimal control [1, 2]. The focus of this paper is the application of these methods for some important problems in NMR. Previously, gradient-based optimizations of NMR pulse sequences for coupled spin systems have almost exclusively relied on gradients computed by the difference method. One important exception are analytical derivatives introduced by Levante et al. [14] for pulse sequence optimizations, where the performance can be expressed in terms of the eigenvalues and eigenfunctions of the total propagator.

The paper is organized as follows. In section 2, we present the basic theoretical ideas and numerical optimization algorithms directly applicable to the problem of pulse design. To illustrate the method, we present three simple but non-trivial applications to coupled spin systems both in the presence and in the absence of relaxation. In section 3.1, we look at the problem of finding maximum coherence transfer achievable in a given time and the design of pulse sequences that achieve this transfer. In section 3.2, the algorithm is used to find relaxation optimized pulse sequences that perform desired coherence transfer operations with minimum losses. In section 3.3, we design pulse sequences that produce a desired unitary propagator in a network of coupled spins in minimal time. In all examples, we compare the results obtained by the numerical optimization algorithm with optimal solutions obtained by analytical arguments based on geometric optimal control theory. In the conclusion section, we discuss the convergence properties of the proposed algorithm and possible extensions.

2 Theory

2.1 Transfer between hermitian operators in the absence of relaxation

To fix ideas, we first consider the problem of pulse design for polarization or coherence transfer in the absence of relaxation. The state of the spin system is characterized by the density operator $\rho(t)$, and its equation of motion is the Liouville-von Neuman equation [15]

$$\dot{\rho}(t) = -i \left[\left(\mathcal{H}_o + \sum_{k=1}^m u_k(t) \mathcal{H}_k \right), \rho(t) \right], \quad (1)$$

where \mathcal{H}_o is the free evolution Hamiltonian, \mathcal{H}_k are the radio-frequency (rf) Hamiltonians corresponding to the available control fields and $u(t) = (u_1(t), u_2(t), \dots, u_m(t))$ represents the vector of amplitudes that can be changed and which is referred to as control vector. The problem is to find the optimal amplitudes $u_k(t)$ of the rf fields that steer a given initial density operator $\rho(0) = \rho_o$ in a specified time T to a density operator $\rho(T)$ with maximum overlap to some desired target operator C . For hermitian operators ρ_o and C , this overlap may be measured by the standard inner product

$$\langle C | \rho(T) \rangle = \text{tr} \{ C^\dagger \rho(T) \}. \quad (2)$$

(For the more general case of non-hermitian operators, see section 2.2). Hence, the performance index Φ_o of the transfer process can be defined as

$$\Phi_o = \langle C | \rho(T) \rangle. \quad (3)$$

In the following, we will assume for simplicity that the chosen transfer time T is discretized in N equal steps of duration $\Delta t = T/N$ and during each step, the control amplitudes u_k are constant, i.e. during the j^{th} step the amplitude $u_k(t)$ of the k^{th} control Hamiltonian is given by $u_k(j)$ (c.f. Fig. 1). The time-evolution of the spin system during a time step j is given by the propagator

$$U_j = \exp \left\{ -i \Delta t \left(\mathcal{H}_o + \sum_{k=1}^m u_k(j) \mathcal{H}_k \right) \right\}. \quad (4)$$

The final density operator at time $t = T$ is

$$\rho(T) = U_N \dots U_1 \rho_o U_1^\dagger \dots U_N^\dagger \quad (5)$$

and the performance function Φ_o (Eq. 3) to be maximized can be expressed as

$$\Phi_o = \langle C | U_N \dots U_1 \rho_o U_1^\dagger \dots U_N^\dagger \rangle. \quad (6)$$

Using the definition of the inner product (c.f. Eq. 2) and the fact that the trace of a product is invariant under cyclic permutations of the factors, this can be rewritten as

$$\Phi_o = \langle \underbrace{U_{j+1}^\dagger \dots U_N^\dagger C U_N \dots U_{j+1}}_{\lambda_j} \mid \underbrace{U_j \dots U_1 \rho_o U_1^\dagger \dots U_j^\dagger}_{\rho_j} \rangle, \quad (7)$$

where ρ_j is the density operator $\rho(t)$ at time $t = j\Delta t$ and λ_j is the backward propagated target operator C at the same time $t = j\Delta t$. Let us see how the performance Φ_o changes when we perturb the control amplitude $u_k(j)$ at time step j to $u_k(j) + \delta u_k(j)$. From Eq. (4) the change in U_j to first order in $\delta u_k(j)$ is given by

$$\delta U_j = -i\Delta t \delta u_k(j) \bar{\mathcal{H}}_k U_j \quad (8)$$

with

$$\bar{\mathcal{H}}_k \Delta t = \int_0^{\Delta t} U_j(\tau) H_k U_j(-\tau) d\tau \quad (9)$$

and

$$U_j(\tau) = \exp\{-i \tau (\mathcal{H}_o + \sum_{k=1}^m u_k(j) \mathcal{H}_k)\}. \quad (10)$$

This follows from the standard formula

$$\frac{d}{dx} e^{A+xB} \Big|_{x=0} = e^A \int_0^1 e^{A\tau} B e^{-A\tau} d\tau. \quad (11)$$

For small Δt (when $\Delta t \ll \|\mathcal{H}_o + \sum_{k=1}^m u_k(j) \mathcal{H}_k\|^{-1}$), $\bar{H}_k \approx H_k$ and using Eqs. (7) and (8) we find to first order in Δt

$$\frac{\delta \Phi_o}{\delta u_k(j)} = - \langle \lambda_j \mid i \Delta t [\mathcal{H}_k, \rho_j] \rangle. \quad (12)$$

Observe we increase the performance function Φ_o if we choose

$$u_k(j) \rightarrow u_k(j) + \epsilon \frac{\delta \Phi_o}{\delta u_k(j)}, \quad (13)$$

where ϵ is a small step size. This forms the basis of the following algorithm, which we denote GRAPE (gradient ascent pulse engineering) in order to distinguish it from conventional gradient approaches used in NMR based on difference methods.

Basic GRAPE algorithm

- 1) Guess initial controls $u_k(j)$.
- 2) Starting from ρ_o , calculate $\rho_j = U_j \dots U_1 \rho_o U_1^\dagger \dots U_j^\dagger$ for all $j \leq N$.
- 3) Starting from $\lambda_N = C$, calculate $\lambda_j = U_{j+1}^\dagger \dots U_N^\dagger C U_N \dots U_{j+1}$ for all $j \leq N$.
- 4) Evaluate $\delta \Phi_o / \delta u_k(j)$ and update the $m \times N$ control amplitudes $u_k(j)$ according to Eq. (13).

5) With this as the new controls, go to step 2).

The algorithm is terminated if the change in the performance index Φ_o is smaller than a chosen threshold value.

In principle, the choice of starting $u_k(j)$ can be completely random. However, an educated guess might lead to faster convergence. Clearly, since the algorithm is based on a gradient ascent procedure, there is no guarantee that it will converge to a global minimum. However at each step the algorithm moves in the direction of increasing performance (c.f. Fig. 1), so we can be assured that it converges to control amplitudes that are extremal points of the desired performance function. To expedite the process of this convergence, we can adopt standard conjugate gradient methods [2].

The important advantages of the optimal control related approach are best highlighted by comparing the GRAPE algorithm to conventionally used numerical difference methods to calculate the gradient $\delta\Phi_o/\delta u_k(j)$ by computing Φ_o for the given pulse sequence $u_k(j)$ as well as for small variations of all $m \times N$ control amplitudes. For example, for $N = 500$ and $m = 4$, the conventional approach would require to calculate 2001 full time evolutions of the density operator from $t = 0$ to T . In contrast, the GRAPE approach to calculate the same gradient $\delta\Phi_o/\delta u_k(j)$ only requires two full time evolutions (one to propagate ρ_o from $t = 0$ to T and one to back-propagate λ_N from $t = T$ to 0), i.e. it is orders of magnitude faster. This makes it possible to efficiently optimize NMR pulse sequences in much larger parameter spaces. As conventional approaches were typically limited to a few dozens of control variables, a typical strategy was to restrict the optimization to certain pulse families, such as composite pulses with a limited number of flip and phase angles [16, 17], Gaussian pulse cascades [18], spline functions [19] or Fourier expansions [20]. In contrast, the GRAPE algorithm allows for much higher flexibility as the number of pulse parameters to be optimized can be orders of magnitude larger compared to conventional approaches.

2.2 Transfer between non-hermitian states in the absence of relaxation

For non-hermitian operators ρ_o and C (e.g. $\rho_o = S^- = S_x - iS_y$ and $C = I^- = I_x - iI_y$, c.f. section 3.1), Φ_o as defined in Eq. (3) cannot be used directly as a performance index for the optimization, because in general it is not real valued. Depending on the application [21], suitable performance functions for non-hermitian operators are the real part of Φ_o or the absolute value of Φ_o :

$$\Phi_1 = \text{Re}(\Phi_o) = \text{Re}\langle C|\rho(T)\rangle \quad (14)$$

$$= \text{Re}\langle (C^x + i C^y)|U_N \dots U_1 (\rho_o^x + i \rho_o^y) U_1^\dagger \dots U_N^\dagger\rangle \quad (15)$$

or

$$\Phi_2 = |\Phi_o|^2 = |\langle C|\rho(T)\rangle|^2 = \langle C|\rho(T)\rangle\langle\rho(T)|C\rangle, \quad (16)$$

where C^x and $i C^y$ are the hermitian and skew-hermitian parts of the target operator C and ρ_o^x and $i \rho_o^y$ are the hermitian and skew-hermitian parts of ρ_o .

For the performance function Φ_1 we find the gradient to first order in Δt

$$\frac{\delta\Phi_1}{\delta u_k(j)} = -\langle \lambda_j^x | i\Delta t[\mathcal{H}_k, \rho_j^x] \rangle - \langle \lambda_j^y | i\Delta t[\mathcal{H}_k, \rho_j^y] \rangle, \quad (17)$$

where ρ_j^x and ρ_j^y are the hermitian and skew-hermitian parts of $\rho_j = \rho_j^x + i\rho_j^y$ and similarly λ_j^x and λ_j^y are the hermitian and skew-hermitian parts of $\lambda_j = \lambda_j^x + i\lambda_j^y$.

For the performance function Φ_2 the gradient to first order in Δt is given by

$$\begin{aligned} \frac{\delta\Phi_2}{\delta u_k(j)} &= -\langle \lambda_j | i\Delta t[\mathcal{H}_k, \rho_j] \rangle \langle \rho_N | C \rangle - \langle C | \rho_N \rangle \langle i\Delta t[\mathcal{H}_k, \rho_j] | \lambda_j \rangle \\ &= -2\text{Re}\{\langle \lambda_j | i\Delta t[\mathcal{H}_k, \rho_j] \rangle \langle \rho_N | C \rangle\}. \end{aligned} \quad (18)$$

Using the gradient $\delta\Phi_1/\delta u_k(j)$ or $\delta\Phi_2/\delta u_k(j)$ instead of $\delta\Phi_o/\delta u_k(j)$ in step 4, the basic GRAPE algorithm described in section 2.1 can also be applied to optimize the transfer between non-hermitian operators.

2.3 Relaxation-optimized coherence transfer

In Liouville space [15], the equation of motion for the density operator in the presence of relaxation can be written as

$$\dot{\rho} = \hat{\mathcal{L}}\rho, \quad (19)$$

where $\hat{\mathcal{L}} = -i\hat{\mathcal{H}} + \hat{\Gamma}$ is the Liouville superoperator, $\hat{\mathcal{H}}$ is the Hamilton superoperator and $\hat{\Gamma}$ is the relaxation superoperator (including thermal correction [22] if appropriate). For simplicity, here we consider the transfer between hermitian operators ρ_o and C , but the results can be easily generalized to non-hermitian operators (c.f. section 2.2). According to Eq. (3), a suitable performance function is

$$\Phi_o = \langle C|\rho(T)\rangle, \quad (20)$$

where now the final density operator $\rho(T)$ is given by

$$\rho(T) = \hat{L}_N \dots \hat{L}_1 \rho_o \quad (21)$$

with

$$\hat{L}_j = \exp\{ \hat{\mathcal{L}} \Delta t \}. \quad (22)$$

Hence, the performance function can be expressed as

$$\begin{aligned} \Phi_o &= \langle C | \hat{L}_N \dots \hat{L}_1 \rho_o \rangle \\ &= \langle \underbrace{\hat{L}_{j+1}^\dagger \dots \hat{L}_N^\dagger C}_{\lambda_j} | \underbrace{\hat{L}_j \dots \hat{L}_1 \rho_o}_{\rho_j} \rangle \end{aligned} \quad (23)$$

and as in Eq. (12) to first order in Δt

$$\frac{\delta \Phi_o}{\delta u_k(j)} = -\langle \lambda_j | i\Delta t \hat{\mathcal{H}}(\rho_j) \rangle = -\langle \lambda_j | i\Delta t [\mathcal{H}_k, \rho_j] \rangle \quad (24)$$

where in the presence of relaxation, λ_j and ρ_j are defined in Eq. (23).

2.4 Synthesis of unitary transformations

Now we consider the problem to create in a given time T a desired unitary propagator. The equation of motion for the propagator of a closed quantum system is

$$\dot{U} = -i(\mathcal{H}_o + \sum_{k=1}^m u_k(t) \mathcal{H}_k) U. \quad (25)$$

At $t = 0$, the initial propagator is $U(0) = \mathbf{1}$.

First we consider the problem to approach a desired propagator U_F by applying a pulse sequence $u_j(t)$ such that at the final time

$$\|U_F - U(T)\|^2 = \|U_F\|^2 - 2\text{Re}\langle U_F | U(T) \rangle + \|U(T)\|^2 \quad (26)$$

is minimized, which is equivalent to maximizing $\text{Re}\langle U_F | U(T) \rangle$. Hence we can define the performance function to be optimized by the pulse sequence as

$$\begin{aligned} \Phi_3 &= \text{Re}\langle U_F | U(T) \rangle \\ &= \text{Re}\langle U_F | U_N \dots U_1 \rangle \\ &= \text{Re}\langle \underbrace{U_{j+1}^\dagger \dots U_N^\dagger U_F}_{P_j} | \underbrace{U_j \dots U_1}_{X_j} \rangle. \end{aligned} \quad (27)$$

and the corresponding gradient $\delta\Phi_3/\delta u_k(j)$ to first order in Δt is given by

$$\frac{\delta\Phi_3}{\delta u_k(j)} = -\text{Re}\langle P_j | i\Delta t \mathcal{H}_k X_j \rangle. \quad (28)$$

While the performance index Φ_3 may be of theoretical interest, for practical applications, it is sufficient to approach the target propagator U_F only up to an arbitrary phase factor $\exp\{i\varphi\}$ and

$$\|U_F - e^{i\varphi} U(T)\|^2 = \|U_F\|^2 - 2\text{Re}\langle U_F | e^{i\varphi} U(T) \rangle + \|U(T)\|^2 \quad (29)$$

is to be minimized for choice of φ , which is equivalent to maximizing the performance function

$$\begin{aligned} \Phi_4 &= |\langle U_F | U(T) \rangle|^2 \\ &= \langle U_F | U_N \dots U_1 \rangle \langle U_1 \dots U_N | U_F \rangle \\ &= \langle P_j | X_j \rangle \langle X_j | P_j \rangle \end{aligned} \quad (30)$$

with the operators X_j and P_j as defined in Eq. (27). The corresponding gradient $\delta\Phi_4/\delta u_k(j)$ to first order in Δt is given by

$$\begin{aligned} \frac{\delta\Phi_4}{\delta u_k(j)} &= -\langle P_j | X_j \rangle \langle i\Delta t \mathcal{H}_k X_j | P_j \rangle - \langle P_j | i\Delta t \mathcal{H}_k X_j \rangle \langle X_j | P_j \rangle \\ &= -2\text{Re}\{\langle P_j | i\Delta t \mathcal{H}_k X_j \rangle \langle X_j | P_j \rangle\}. \end{aligned} \quad (31)$$

2.5 Reduction of rf power and limited rf amplitudes

In the given formulation of the optimization problem, it is also straight-forward to add to any of the previously defined performance functions Φ_i , a penalty

$$\Phi_{rf} = \alpha \sum_{j=1}^N \sum_{k=1}^m \{u_k(j)\}^2 \Delta t \quad (32)$$

for the total rf power applied during the pulse sequence to minimize sample heating, where α is a weight of the penalty imposed for excessive rf-power. Hence, the gradient simply contains an additional term

$$\frac{\delta\Phi_{rf}}{\delta u_k(j)} = -2 \alpha u_k(j) \Delta t. \quad (33)$$

If the maximum rf amplitude is limited, this can be taken into account in the algorithm described in section 2.1 by resetting the amplitude to the maximum amplitude if it is exceeded after step 4 (see e.g. [12]).

2.6 Robustness

For practical applications, it is often desirable to achieve the optimum performance for a range of parameters ω , such as a given range of chemical shifts and/or a given range of rf amplitudes to take into account the effects of rf inhomogeneity or rf miscalibration. If the range of parameters is sampled at discrete values ω_p , the total performance Φ_{tot} can be measured by summing over the performance of systems parameterized by ω_p :

$$\Phi_{tot} = \sum_p \Phi(\omega_p). \quad (34)$$

For example, for the case of hermitian transfer, with $\Phi(\omega_p) = \Phi_o(\omega_p)$ (c.f. section 2.1)

$$\begin{aligned} \Phi_{tot} &= \sum_p \langle C | U_N(\omega_p) \dots U_1(\omega_p) \rho_o U_1^\dagger(\omega_p) \dots U_N^\dagger(\omega_p) \rangle \\ &= \sum_p \langle \lambda_j(\omega_p) | \rho_j(\omega_p) \rangle \end{aligned} \quad (35)$$

and

$$\frac{\delta \Phi_{tot}}{\delta u_k(j)} = - \sum_p \langle \lambda_j(\omega_p) | i\Delta t [\mathcal{H}_k, \rho_j(\omega_p)] \rangle. \quad (36)$$

3 Examples

3.1 Time-optimal coherence-order selective in-phase transfer

As a practical example, we consider coherence-order selective in-phase transfer ($I^- \rightarrow S^-$) [23] in a heteronuclear two-spin system in the absence of relaxation. Here we are interested in the following question: What is the minimum time to achieve a specified amount of coherence transfer, or conversely, what is the maximum possible coherence transfer amplitude in any given time T in the absence of relaxation, i.e. under unitary evolution? This is a simple, but non-trivial example, which has only recently been solved analytically based on principles of geometric control [24, 25]. Hence, this constitutes an ideal test case for the presented GRAPE algorithm because numerically optimized transfer amplitudes can be directly compared to the theoretical benchmark provided by the analytical result. We assume that both spins S and I are on-resonance in the doubly rotating frame. The free evolution Hamiltonian of the spin system is

$$\mathcal{H}_o = 2\pi J I_z S_z, \quad (37)$$

where J is the heteronuclear coupling constant. The initial density operator term of interest is $\rho_o = I^- = I_x - iI_y$ and the target operator is $C = S^- = S_x - iS_y$. In most practical applications of coherence-order-selective coherence transfer [21, 23, 26], the goal is to maximize $|\langle I^- | \rho(T) \rangle|$. Hence, the appropriate performance function is Φ_2 (c.f. Eq. (16)). The normalized absolute value of the transfer amplitude for a given mixing period T is defined as [21]

$$\eta(T) = \frac{|\langle S^- | \rho(T) \rangle|}{\|I^-\| \|S^-\|} = \frac{1}{2} |\langle S^- | \rho(T) \rangle|. \quad (38)$$

In our numerical optimizations based on the GRAPE algorithm, the heteronuclear coupling J was chosen to be 1 Hz. 30 pulses were optimized with total durations T in the range between 0 and 1.5 s (c.f. Fig. 2), each pulse was digitized in steps $\Delta t = 0.002$ s. For each time step Δt , the x and y rf amplitudes irradiated at spins I and S were optimized: $u_1(j) = \nu_x^I(j)$, $u_2 = \nu_y^I(j)$, $u_3 = \nu_x^S(j)$, $u_4 = \nu_y^S(j)$ and $\mathcal{H}_1 = 2\pi I_x$, $\mathcal{H}_2 = 2\pi I_y$, $\mathcal{H}_3 = 2\pi S_x$, $\mathcal{H}_4 = 2\pi S_y$ (c.f. Eq. 1). For example for $T = 1.5$ s, this resulted in a total number of 6000 optimization parameters. For each value of T , the gradient flow algorithm was started with initial sequences $u_k(j)$ which were created by assigning a random value to every tenth point and using a cubic spline fit to fill in the amplitudes $u_k(j)$ of the intermediate time points. This resulted in random but relatively smooth initial pulse amplitudes. In these optimizations, the maximum rf amplitude was not limited and we also did not include a penalty for increased rf power (c.f. section 2.5). In Fig. 2, the numerically optimized transfer efficiencies $\eta(T)$ (circles) are superimposed with the analytical curve (solid line), representing time-optimal pulses [24, 25]. For all chosen total durations T , the maximum transfer efficiency $\eta(T)$ found by the gradient algorithm converged to the analytically derived optimum values. The minimum time to reach full transfer ($\eta = 1$) is $\tau^* = 3/(2J)$ [24]. For $T < T^*$, the optimal transfer amplitude is up to 12.5% larger compared to the transfer amplitude of heteronuclear isotropic mixing [27, 28, 29]. For example, Fig. 3 shows an optimized pulse sequence found by the GRAPE algorithm for $T = 0.5$ s. In order to simplify the comparison of the pulses applied to spins I and S , the arbitrary relative phase of the pulse sequence applied to spin S was shifted by 200° . The figure shows that up to this relative phase shift of the S pulse, the sequences are almost identical, as expected in order to create the required effective Hamiltonian [24]. Note that there is an infinite number of possible pulse sequences, which create the optimal average Hamiltonian and hence for each value of T , many optimal solutions exist and the pulse sequences found by the GRAPE algorithm depend strongly on the initial random sequence.

3.2 Relaxation-optimized pulse elements (ROPE)

As a second example, we consider the problem to achieve optimal coherence transfer in the presence of relaxation. For an isolated two-spin system in the spin-diffusion limit, it has recently been demonstrated that currently used pulse sequence elements such as INEPT [30] are far from optimal. For example, if dipolar relaxation between an isolated pair of spins is the dominant relaxation mechanism, the in-phase to anti-phase transfer (I_x to $2I_zS_x$) via analytically derived relaxation optimized pulse sequence elements (ROPE) [31, 32] is up to a factor of $e/2 = 1.36$ more efficient than the traditional INEPT transfer. Here, we demonstrate the application of the GRAPE algorithm to the numerical optimization of ROPE-type sequences and compare the results to the analytical solutions.

We consider a system, consisting of two coupled heteronuclear spins $1/2$, denoted I and S , with a coupling constant of $J = 194$ Hz. In the spin diffusion limit only the transverse relaxation rate k is nonzero, assuming pure dipole-dipole relaxation (without CSA/DD cross-correlation effects) [31]. We consider the case where the transverse relaxation rate k as defined in [31] is equal to the coupling constant, i.e. $k/J=1$ (Here, a thermal correction of the relaxation superoperator need not be included if the transfer element is used as a mixing step [22]). For the transfer $I_z \rightarrow 2I_zS_z$, the initial density operator is $\rho(0) = I_z$ and the desired target operator: $C = 2I_zS_z$. Both spins are assumed to be on-resonance in a doubly rotating frame. Pulse shapes consisting of $N = 75$ discrete time steps were optimized for various pulse durations T , using the gradient $\delta\Phi_o/\delta u_k(j)$ given in Eq. (24). As in the previous example, a random initial sequence was created for each value of T by assigning a random rf amplitudes to every tenth time point and using a cubic spline fit to interpolate the amplitudes $u_k(j)$ of the intermediate time points.

Fig. 4 shows the transfer efficiency of the numerically optimized sequences (black circles). For comparison, the figure also shows the analytical curve representing the theoretical limit [31] of the transfer efficiency as a function of T . E.g., for $T = 2.11$ ms, (i.e. $T/J^{-1} = 0.408$), the numerically optimized pulse sequence is shown in Fig. 5 A. This pulse shape is very close to the analytically derived ROPE pulse [31]. It is interesting to note that the numerically optimized pulse closely approaches the theoretical limit with a finite maximum rf amplitude. In the center of the pulse, x and y rf amplitudes are slightly overlapping, whereas this is not the case in the analytical solution, which has a short delay in the center of the sequence. However, the characteristic ROPE transfer mechanism (c.f. Figs. 4 and 6 in Ref. [31]) is evident in Fig. 5 B, which shows the trajectories of

the non-vanishing terms of the density operator under the action of the pulse shape shown in Fig. 5 A. In contrast to INEPT, a large portion of I_x is immediately transformed to I_z , which is protected from relaxation in the present model. Consequently, I_z is brought in an optimal trajectory back to the transverse plain in the first phase of the transfer. In the last phase, $2I_yS_z$ is lifted in an optimal way to $2I_zS_z$, which is again protected against relaxation [31]. Although for simplicity, CSA relaxation was not considered in this example, it is straight-forward to include CSA relaxation as well as the effects of cross-correlation in the relaxation matrix and to numerically optimize corresponding pulses (data not shown) [33]. Furthermore, the algorithm is not limited to two coupled spins and more complicated relaxation networks can be taken into account.

3.3 Time-optimal implementation of unitary transformations

This example illustrates the use of the GRAPE algorithm in the development of pulse sequences that implement a desired unitary propagator in minimum time. We consider a chain of three heteronuclear spins with coupling constants $J_{12} = J_{23} = J$, $J_{13} = 0$. In a multiple-rotating frame, in which the three heteronuclear spins are on resonance, the free evolution Hamiltonian \mathcal{H}_o is

$$\mathcal{H}_o = 2\pi J I_{1z} I_{2z} + 2\pi J I_{2z} I_{3z}. \quad (39)$$

Many applications in NMR spectroscopy [34, 35] and NMR quantum computing [36, 37, 38] require unitary transformations of the form

$$\mathcal{U}_{zzz}(\alpha) = \exp\{-i\alpha 4I_{1z}I_{2z}I_{3z}\}. \quad (40)$$

We recently derived analytically the minimum time $T^*(\alpha)$ to create $\mathcal{U}_{zzz}(\alpha)$. The corresponding pulse sequences [39, 40] are considerably shorter than conventional implementations of these unitary propagators [34, 35, 41, 42]. For $0 \leq \alpha \leq \pi/2$, the minimum time T^* is given by [39, 40]

$$T^*(\alpha) = \frac{\sqrt{\alpha(2\pi - \alpha)}}{\pi J} \quad (41)$$

and $T^*(n\pi \pm \alpha) = T^*(\alpha)$, where n is an arbitrary integer.

Here, we used the gradient $\delta\Phi_3/\delta u_k(j)$ defined in Eq. (28), where the initial unitary propagator is the identity matrix ($U(0) = \mathbf{1}$) and the target operator is $U(T) = \mathcal{U}_{zzz}(\alpha)$ for six values of α between 0 and $\pi/2$ (c.f. Fig. 6). The heteronuclear couplings J were chosen to be 1 Hz and each pulse was digitized in steps $\Delta t = 0.0025$ s. For each time step Δt , the x and y rf amplitudes irradiated at spins I_1 and I_2 and I_3 were optimized. For each value of α , random initial pulse sequences

were numerically optimized for various pulse durations T , in order to determine the minimum T , for which the numerical algorithm finds a performance index of $\Phi_3/\text{Tr}\{\mathbf{1}\} = 1$. In this series of optimizations, T was incremented in steps of 0.05 s for each value of α . The shortest durations T , for which a numerical value of 1.0 was found for $\Phi_3/\text{Tr}\{\mathbf{1}\}$ are indicated by circles in Fig. 6. These durations represent upper numerical limits for the minimum time T^* . Fig. 7 shows the numerically optimized pulse sequence for $\alpha = \pi/8$ and $T = 0.5$ s. It is qualitatively similar to the analytically derived sequence, which consists only of rf pulses irradiated at spin I_2 [39]. The curve representing the analytical solution of T^* (c.f. Eq. 41) is also shown in Fig. 6 for comparison. The asterisks represent the longest durations T with $\Phi_3/\text{Tr}\{\mathbf{1}\} < 1.0$. For $\alpha = \pi/40, \pi/16, \pi/8, \pi/4, 3\pi/8$, and $\pi/2$, the numerical values of $\Phi_3/\text{Tr}\{\mathbf{1}\}$ at the times T indicated by asterisks in Fig. 6 were 0.99995, 0.99998, 0.9987, 0.9997, 0.9985, and 0.9994, respectively.

4 Conclusions

In this paper we have presented a streamlined derivation of analytical gradients for the design of pulse sequences in NMR spectroscopy. We applied these optimal control related algorithms to the design of pulse shapes for problems involving transfer of coherence between coupled spins and synthesis of unitary propagators in a network of coupled spins. Although the theory and numerical principles are textbook material in the area of optimal control, its application to the control of coupled spin dynamics is new and promising. It should be noted that the proposed gradient ascent algorithms are not guaranteed to converge to a globally optimal pulse shape. All that can be said is the proposed algorithms will converge to a stationary point of the performance function. To speed up convergence, the algorithm can be further modified by using adaptive step sizes for updating the control amplitudes as well as by using conjugate gradients instead of ordinary ones. Yet, all these issues are technicalities of implementation that have not been addressed here, where we highlighted the basic ideas. In future extensions, we plan to test these type of variations in order to speed up algorithms. Note that variations of similar ideas have appeared in other fields of coherent control [3, 6, 43], where iterative modifications of controls yielded improved pulse shapes. All these methods only guarantee convergence to some critical point that does not have to be the global optimum. A standard modification to the gradient ascent adds some noise when updating the control amplitudes in order to avoid getting trapped in local minima.

In the work introduced here, we have not only improved upon pulse sequences, but the GRAPE al-

gorithms have lead us to novel coherence transfer pathways. Further investigation has even triggered analytical solutions to optimal pulse shapes as well as optimal pulse sequences. We have also used special instantiations of the GRAPE algorithm for designing broadband excitation pulses in uncoupled spin systems [11], which are examples of robust control for a range of spin system parameters, such as chemical shift and rf amplitude. With the given gradients $\delta\Phi_i/\delta u_k(j)$, it is also straight forward to suppress undesired coherence transfers while simultaneously optimizing desired transfers. For example, this can be achieved by defining the overall quality factor as a (weighted) sum of e.g. Φ_2 for the desired transfer and $-\Phi_2$ for the undesired transfer. A practical problem is the choice of the number of time steps for the discretization of pulse shapes for a given control problem. This is directly related to the number of pulse parameters to be optimized. The discretization should be chosen to ensure that the condition given for Eq. (12) is approximately satisfied. In the presented examples, the chosen number of pulse parameters was sufficiently large to achieve the previously known theoretical performance limits, but we have not explored in detail the minimal number of pulse sequence parameters necessary to achieve the theoretical bounds. In practice, this may be done by increasing the number of pulse sequence parameters until convergence of the performance index is reached. In a recent paper [44], first applications of the GRAPE algorithm to polarization transfer in solid state NMR have been presented. This forms a further example demonstrating the optimization of robust pulse sequences for a large range of parameters, e.g. due to the powder average of dipolar couplings and the possibility to include a time-varying free-evolution Hamiltonian as in magic angle sample spinning. The algorithm introduced here is expected to be a very useful tool for developing superior pulse sequences in multiple spin systems.

Acknowledgments

N.K. would like to acknowledge DARPA QUIST grant 496020-01-1-0556, NSF 0218411 and NSF 0133673. S.G. thanks the Deutsche Forschungsgemeinschaft for grants Gl 203/3-1 and Gl 203/4-2 and the Fonds der Chemischen Industrie.

References

- [1] V. F. Krotov, Global methods in optimal control, Marcel Decker, New York (1996).
- [2] A. Bryson, Jr., Y.-C. Ho, Applied optimal control, Hemisphere, Washington, D.C. (1975).

- [3] D. J. Tannor, S. A. Rice, Control of selectivity of chemical reaction via control of wave packet evolution, *J. Chem. Phys.* **83**, 5013-5018 (1985).
- [4] S. Shi, H. Rabitz, Quantum mechanical optimal control of physical observables in microsystems, *J. Chem. Phys.* **92** 364- 376 (1990).
- [5] W. Jakubetz, E. Kades, J. Manz, State-selective excitation of molecules by means of optimized ultrashort infrared laser pulses, *J. Phys. Chem.* **97**, 12609-12619 (1993).
- [6] Y. Ohtsuki, G. Turinici, H. Rabitz, Generalized monotonically convergent algorithms for solving quantum optimal control problems, *J. Chem. Phys.* **120**, 5509-5517 (2004).
- [7] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, K. Kompa, Whither the Future of Controlling Quantum Phenomena?, *Science* **288**, 824-828 (2000).
- [8] S. Conolly, D. Nishimura, A. Macovski, Optimal control solutions to the magnetic resonance selective excitation problem, *IEEE Trans. Med. Imag.* **MI-5**, 106115 (1986).
- [9] J. Mao, T. H. Mareci, K. N. Scott, E. R. Andrew, Selective inversion radiofrequency pulses by optimal control, *J. Magn. Reson.* **70** ,310318 (1986).
- [10] D. Rosenfeld, Y. Zur, Design of adiabatic selective pulses using optimal control theory, *Magn. Reson. Med.* **36**, 401409 (1996).
- [11] T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja, S. J. Glaser, Application of optimal control theory to the design of broadband excitation pulses for high resolution NMR, *J. Magn. Reson.* **163**, 8-15 (2003).
- [12] T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja, S. J. Glaser, Reducing the duration of broadband excitation pulses using optimal control with limited rf amplitude, *J. Magn. Reson.* **167**, 68-74 (2004).
- [13] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, B. Luy, Exploring the limits of broadband excitation and inversion pulses, *J. Magn. Reson.* **170**, 236-243 (2004).
- [14] T. O. Levante, T. Bremi, R. R. Ernst, Pulse-sequence optimizations with analytical derivatives. Application to deuterium decoupling in oriented phases, *J. Magn. Reson. A* **121**, 167-177 (1996).
- [15] R. R. Ernst, G. Bodenhausen, A. Wokaun, Principles of nuclear magnetic resonance in one and two dimensions, Clarendon Press, Oxford (1987).

- [16] M. H. Levitt, Composite pulses, *Prog. NMR Spectrosc.* **18**, 61-122 (1986).
- [17] M. H. Levitt, Composite pulses, in Encyclopedia of Nuclear Magnetic Resonance, vol. 2, p.1396, (Eds.-in-Chief D.M.Grant and R.K.Harris), John Wiley & Sons, Chichester-NewYork-Brisbane-Singapore, 1996.
- [18] L. Emsley, G. Bodenhausen, Gaussian pulse cascades: new analytical functions for rectangular selective inversion and in-phase excitation in NMR, *Chem. Phys. Lett.* **165**, 469-476 (1990).
- [19] B. Ewing, S. J. Glaser and G. P. Drobny, Development and optimization of shaped NMR pulses for the study of coupled spin systems” *Chem. Phys.* **147**, 121-129 (1990).
- [20] D. B. Zax, G. Goelman, S. Vega, Amplitude-modulated composite pulses, *J. Magn. Reson.* **80**, 375-382 (1988).
- [21] S. J. Glaser, T. Schulte-Herbrüggen, M. Sieveking, O. Schedletzky, N. C. Nielsen, O. W. Sørensen, C. Griesinger, Unitary control in quantum ensembles, Maximizing signal intensity in coherent spectroscopy, *Science* **208**, 421-424 (1998).
- [22] M. H. Levitt, L. Di Bari, The homogeneous master equation and the manipulation of relaxation networks, *Bull. Magn. Reson.* **16**, 94-114 (1994).
- [23] M. Sattler, P. Schmidt, J. Schleucher, O. Schedletzky, S. J. Glaser, C. Griesinger, Novel pulse sequences with sensitivity enhancement for in-phase coherence transfer employing pulsed field gradients, *J. Magn. Reson. B* **108**, 235-242 (1995).
- [24] N. Khaneja, R. W. Brockett, S. J. Glaser, Time-optimal control in spin systems, *Phys. Rev. A* **63**, 03208 (2001).
- [25] T. O. Reiss, N. Khaneja, S. J. Glaser, Time-optimal coherence-order-selective transfer of in-phase coherence in heteronuclear IS spin systems, *J. Magn. Reson.*, **154**, 192-195 (2002).
- [26] A. G. Palmer, J. Cavanagh, , P. E. Wright, M. Rance, Sensitivity improvement in proton-detected heteronuclear correlation spectroscopy, *J. Magn. Reson.* **93**, 151-170 (1991).
- [27] P. Caravatti, L. Braunschweiler, and R. R. Ernst, Heteronuclear correlation spectroscopy in rotating solids, *Chem. Phys. Lett.* **100**, 305-310 (1983).
- [28] D. P. Weitekamp, J. R. Garbow, and A. Pines, Determination of dipole coupling constants using heteronuclear multiple quantum NMR, *J. Chem. Phys.* **77**, 2870-2883 (1982).

- [29] S. J. Glaser and J. J. Quant, Homonuclear and heteronuclear Hartmann-Hahn transfer in isotropic liquids, in "Advances in magnetic and optical resonance" (W. S. Warren, Ed.), Vol. 19, pp. 59 - 252, Academic Press, San Diego (1996).
- [30] G. A. Morris, R. Freeman, Enhancement of nuclear magnetic resonance signals by polarization transfer, *J. Am. Chem. Soc.* **101**, 760-762 (1979).
- [31] N. Khaneja, T. Reiss, B. Luy, S. J. Glaser, Optimal control of spin dynamics in the presence of relaxation, *J. Magn. Reson.* **162**, 311-319 (2003).
- [32] D. Stefanatos, N. Khaneja, S. J. Glaser, Optimal control of coupled spins in presence of longitudinal relaxation, *Phys. Rev. A* **69**, 022319 (2004).
- [33] N. Khaneja, B. Luy, S. J. Glaser, Boundary of quantum evolution under decoherence, *Proc. Natl. Acad. Sci. USA* **100**, 13162-13166 (2003).
- [34] O. W. Sørensen, Polarization transfer experiments in high-resolution NMR spectroscopy, *Prog. NMR Spectrosc.* **21**, 503-569 (1989).
- [35] A. Meissner, O. W. Sørensen, I -spin n -quantum coherences in $I_n S$ spin systems employed for E.COSY-type measurement of heteronuclear long-range coupling constants in NMR, *Chem. Phys. Lett.* **276**, 97-102 (1997).
- [36] D. G. Cory, A. F. Fahmy, T. F. Havel, Ensemble quantum computing by NMR spectroscopy, *Proc. Natl. Acad. Sci* **94**, 1634-1639 (1997).
- [37] N. Gershenfeld, I. L. Chuang, Bulk spin-resonance quantum computation, *Science* **275**, 350-356 (1997).
- [38] C. H. Bennet, D. P. Di Vincenzo, Quantum information and computation, *Nature* **404**, 247-255 (2000).
- [39] N. Khaneja, S. J. Glaser, R. W. Brockett, Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer, *Phys. Rev. A*, **65**, 032301 (2002).
- [40] T. O. Reiss, N. Khaneja, S. J. Glaser, Broadband geodesic pulses for three spin systems: Time-optimal realization of effective trilinear coupling terms and indirect SWAP Gates, *J. Magn. Reson.* **165**, 95-101 (2003).

- [41] C. H. Tseng, S. Somaroo, Y. Sharf, E. Knill, R. Laflamme, T. F. Havel, D. G. Cory, Quantum simulation of a three-body interaction Hamiltonian on an NMR quantum computer, *Phys. Rev. A*, **61**, 012302 (2000).
- [42] J. Kim, J.-S. Lee, S. Lee, Implementing unitary operators in quantum computation, *Phys. Rev. A*, **61**, 032312 (2000).
- [43] C. M. Tesch, L. Kurtz, R. de Vivie-Riedle, Applying optimal control theory for elements of quantum computation in molecular systems, *Chem. Phys. Lett.* **343**, 633-641 (2001).
- [44] C. Kehlet, A. C. Sivertsen, M. Bjerring, T. O. Reiss, N. Khaneja, S. J. Glaser, N. C. Nielsen, Improving solid-state NMR dipolar recoupling by optimal control, *J. Am. Chem. Soc.* **126**, 10202-10203 (2004)

Figure Captions

Figure 1: Schematic representation of a control amplitude $u_k(t)$, consisting of N steps of duration $\Delta t = T/N$. During each step j , the control amplitude $u_k(j)$ is constant. The vertical arrows represent gradients $\frac{\delta\Phi_o}{\delta u_k(j)}$, indicating how each amplitude $u_k(j)$ should be modified in the next iteration in order to improve the performance function Φ_o .

Figure 2: For the coherence order-selective coherence transfer $S^- \rightarrow I^-$ in a system consisting of two heteronuclear spins $1/2$, the numerically optimized transfer efficiencies $\eta(T)$ (circles) and the analytically derive time-optimal transfer efficiency (solid line) [24, 25] are shown.

Figure 3: Example of a numerically optimized pulse shape for coherence order-selective coherence transfer $I^- \rightarrow S^-$ found by the GRAPE algorithm for a given total transfer time of $T = 1/(2J)$ (c.f. Fig. 2). Panel A shows the x amplitude (solid curve) and y amplitude (dashed curve) of the rf field irradiated at spin I and panel B shows the x amplitude (solid curve) and y amplitude (dashed curve) of the rf field irradiated at spin S .

Figure 4: The efficiency η of the transfer from I_z to $2I_zS_z$ in the presence of dipole-dipole relaxation in the spin diffusion limit as a function of the sequence duration T (details, see text). The circles show the efficiencies of five numerically optimized sequences of different durations T and the curve represents the theoretical limit [31].

Figure 5: Example of a numerically optimized pulse sequence (A) of duration $T = 0.408J^{-1}$ for the transfer of I_z to $2I_zS_z$ in the presence of dipole-dipole relaxation (c.f. Fig. 4). ν_x and ν_y correspond to the x (solid curve) and y (dashed curve) rf amplitude irradiated at spin I . Panel (B) shows the corresponding trajectories of the non-vanishing density operator terms during the relaxation-optimized pulse sequence.

Figure 6: The solid curve shows the analytical solution of the minimum time $T^*(\alpha)$ [39] for the creation of a propagator $\mathcal{U}_{zzz}(\alpha) = \exp\{-i\alpha 4I_{1z}I_{2z}I_{3z}\}$ in a spin system consisting of three heteronuclear spins with couplings $J_{12} = J_{23} = J$ and $J_{13} = 0$. For six values of α , pulse sequences with various durations T were optimized using the gradient $\delta\Phi_3/\delta u_k(j)$ (c.f. Eq. 28) in steps $\Delta T=0.05 J^{-1}$. For each α , the largest value of T , for which the optimized numerical value $\Phi_3/\text{Tr}\{\mathbf{1}\}$ was found to be smaller than 1.0 is indicated by an asterisk. The shortest value of T , for which the optimized numerical value $\Phi_3/\text{Tr}\{\mathbf{1}\}$ was found to be 1.0 is indicated by a circle.

Figure 7: Example of a numerically optimized pulse sequence for the creation of the propagator $\mathcal{U}_{zzz}(\alpha)$ corresponding to the data point represented by a circle at $\alpha = \pi/8$ in Fig. 6. The x and y amplitudes of the rf pulse irradiated at spin I_2 are shown. In comparison, the numerically optimized rf amplitudes irradiated at spins I_1 and I_3 are less than 0.5 % at each point during the sequence (data not shown) and can be neglected.

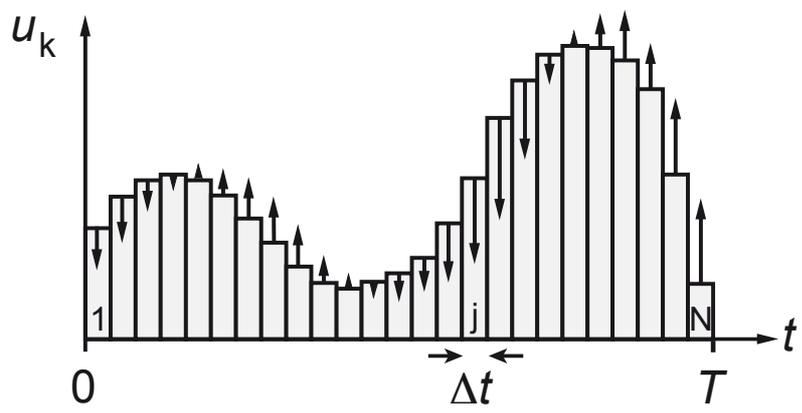


Figure 1

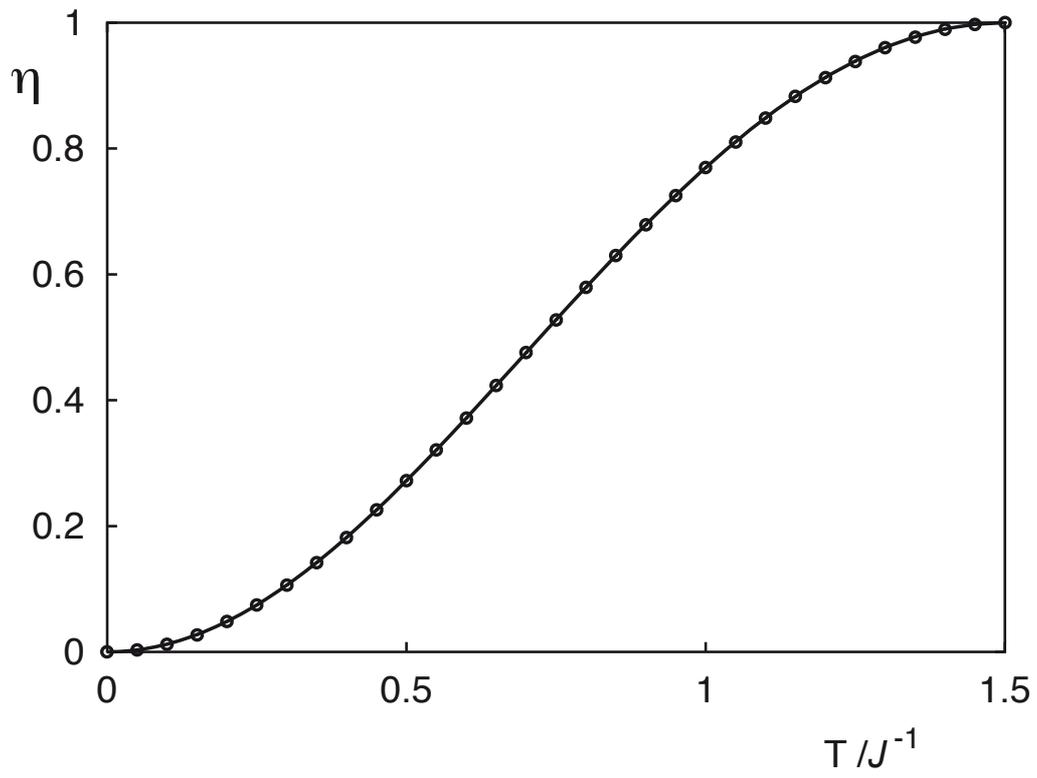


Figure 2

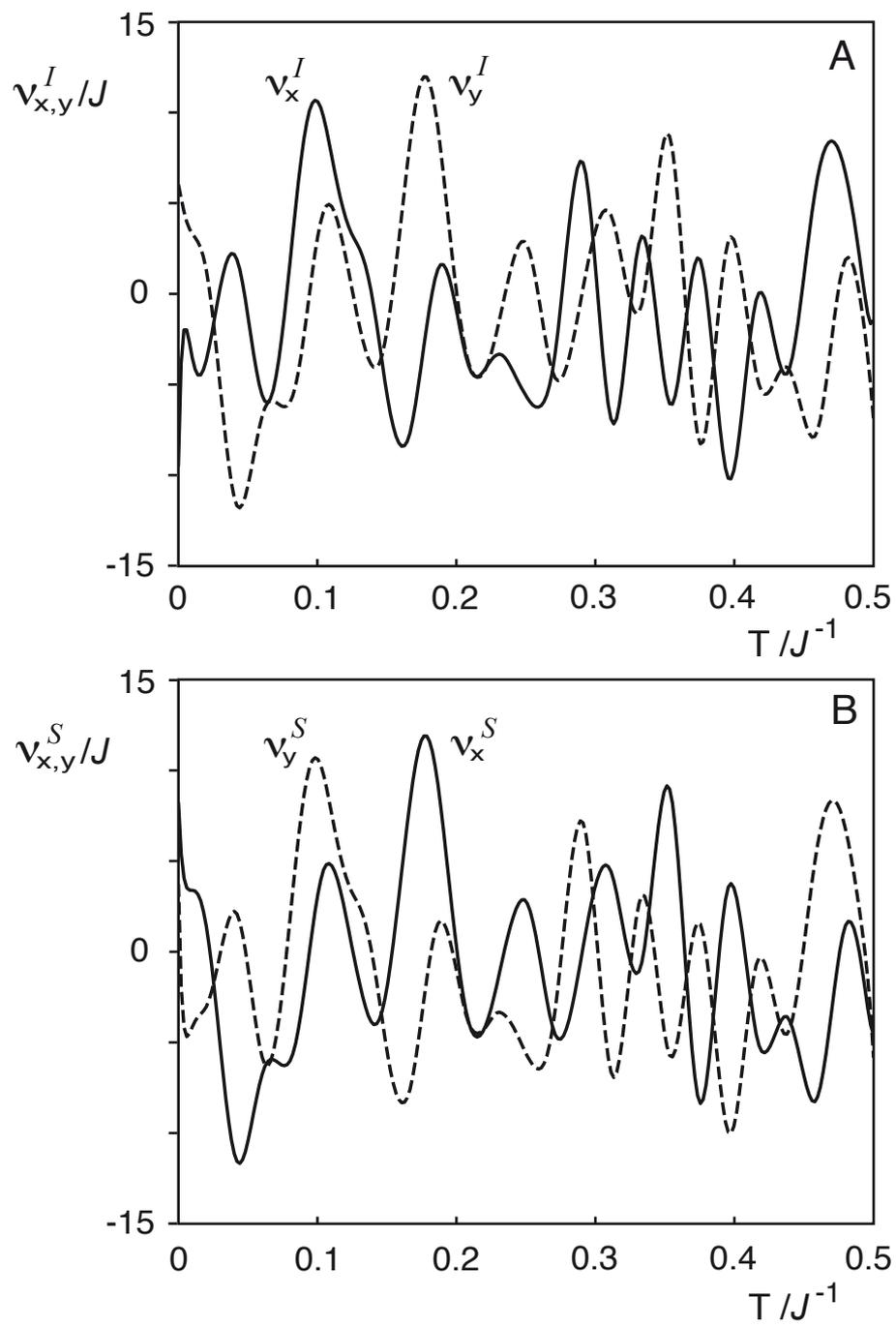


Figure 3

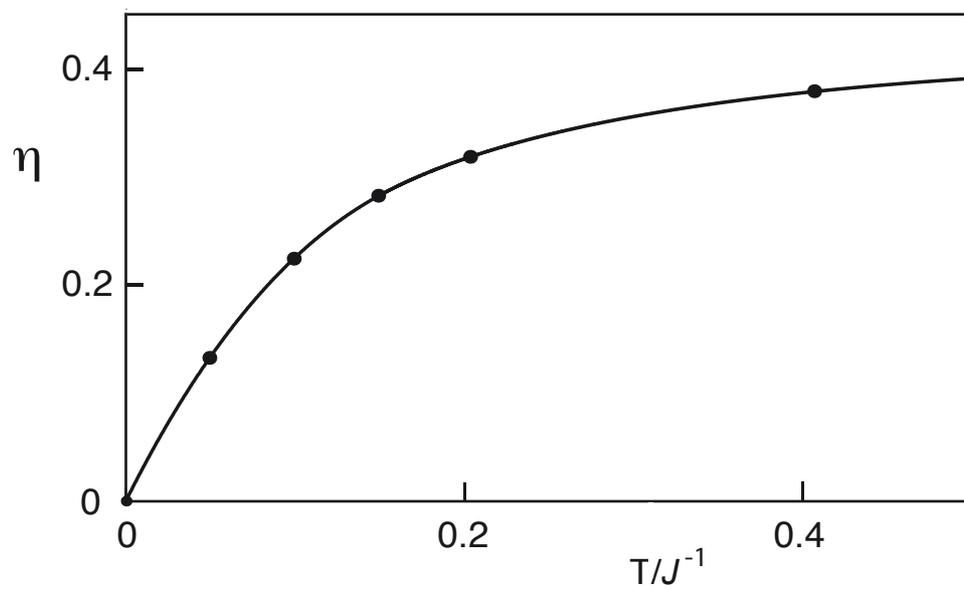


Figure 4

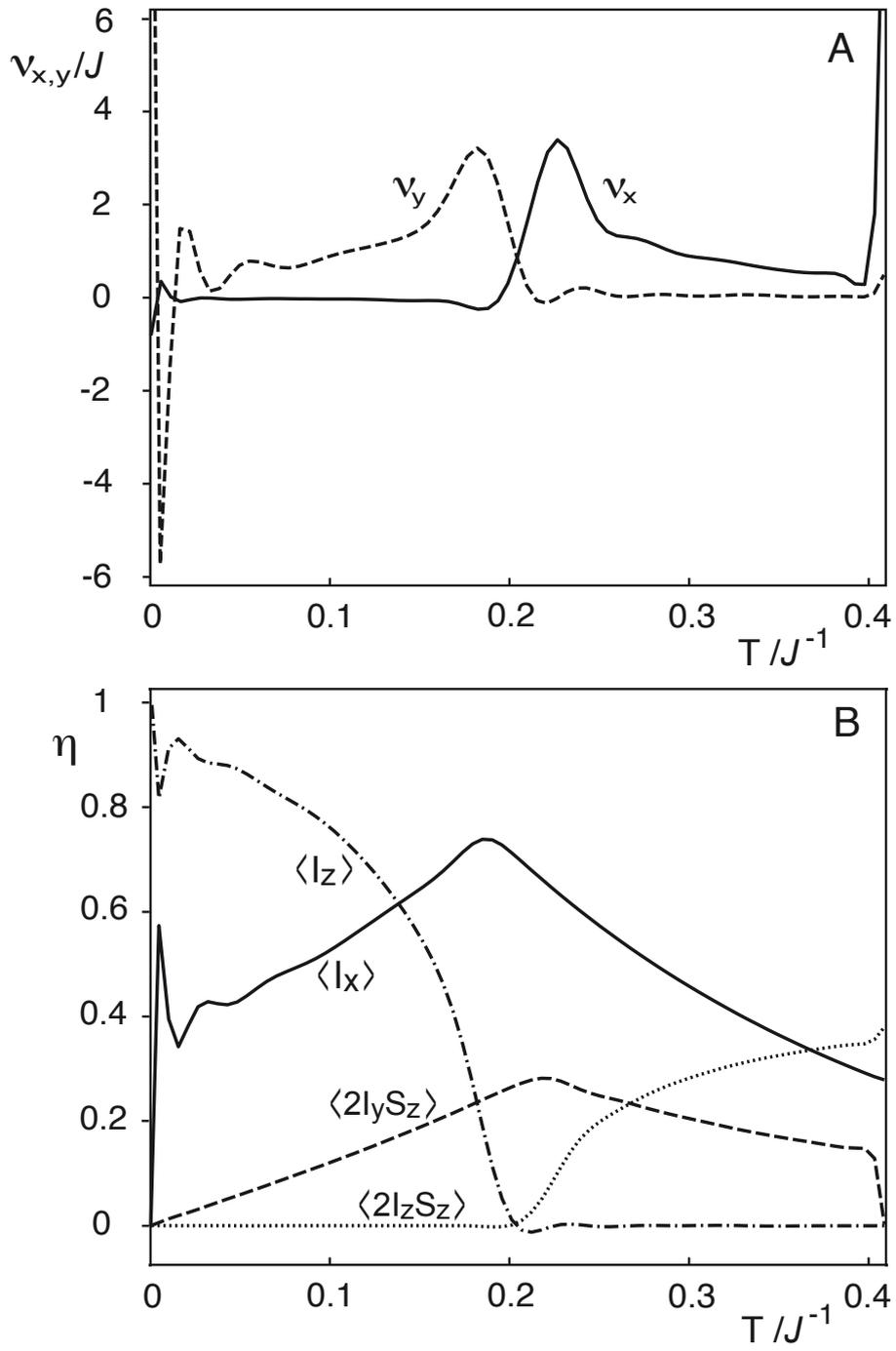


Figure 5

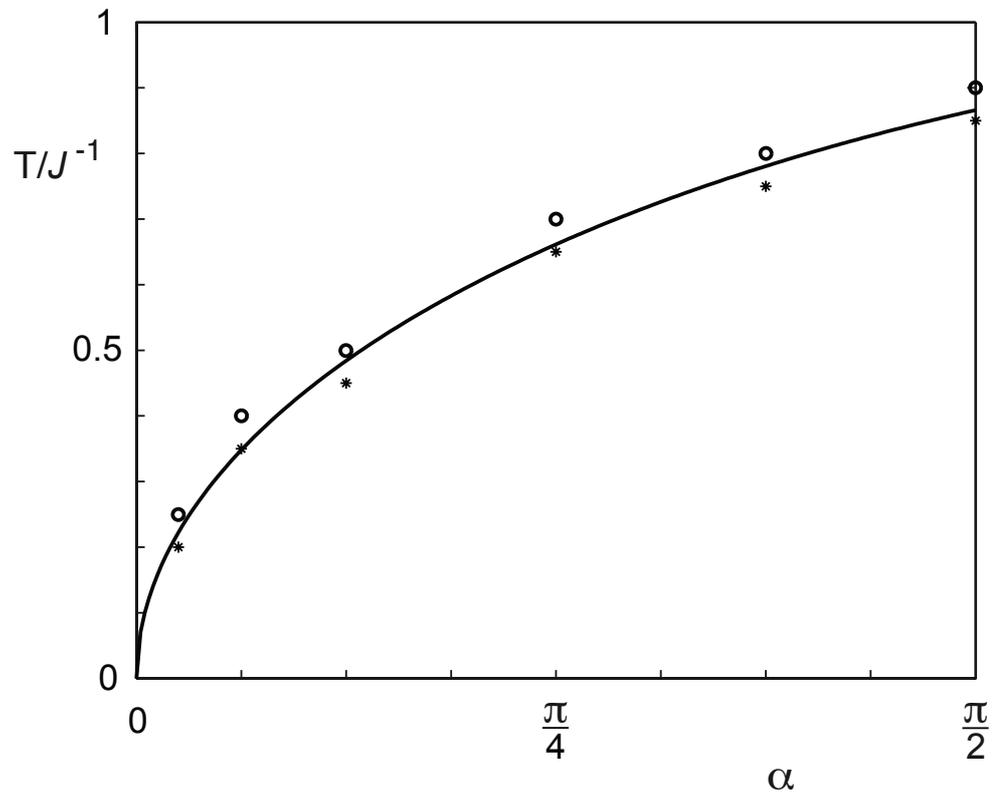


Figure 6

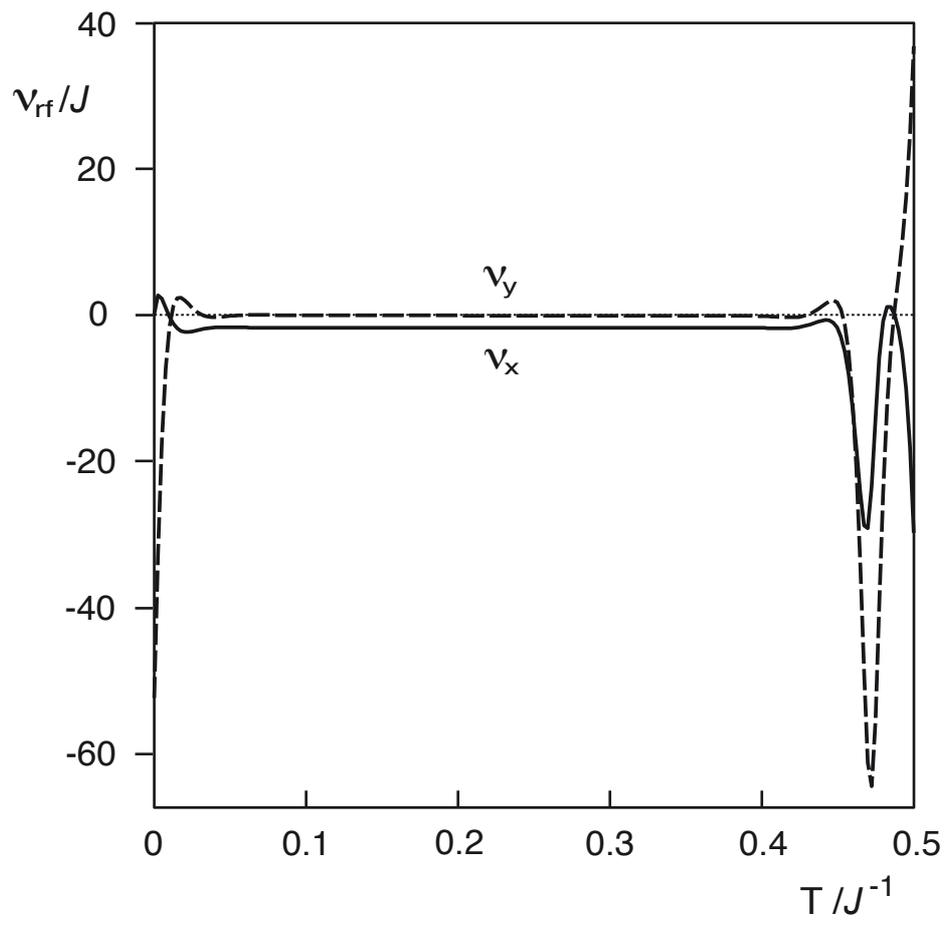


Figure 7