Chemistry

Optimal Control of Spin Systems

with Newton-Raphson Methods

David Goodwin, Ilya Kuprov.

University of Southampton

Optimal Control - Gradient Ascent Pulse Engineering (GRAPE)

The evolution of a quantum system can be characterised by a Liouville equation with a vector representation of the density operator, $\frac{\partial}{\partial t}|\hat{\rho}(t)\rangle = -i\hat{\hat{\mathcal{L}}}(t)|\hat{\rho}(t)\rangle$, having the general solution,

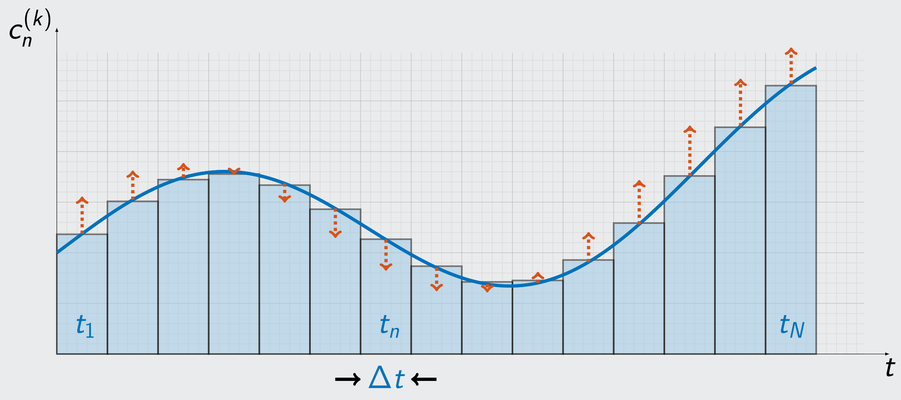
$$|\hat{\rho}(t)\rangle = \exp_{(0)}\left(-i\int_0^t \hat{\hat{\mathcal{L}}}(t)\,\mathrm{d}t\right)|\hat{\rho}(t)\rangle$$

A practical Magnetic Resonance system can be split into two parts; that is beyond experimental control, $\hat{\hat{\mathcal{L}}}_0$, and a set of controllable electromagnetic pulses $\left\{\hat{\hat{\mathcal{L}}}_k\right\}$:

$$\hat{\hat{\mathcal{L}}}(t) = \hat{\hat{\mathcal{L}}}_0 + \sum_k c^{(k)}(t)\hat{\hat{\mathcal{L}}}_k$$

The problem is simplified if the control sequences $\left\{c^{(k)}(t)\right\}$ are assumed to be piecewise constant. For a piecewise constant Hamiltonian, we sequentially multiply each of the, discrete in time, propagators into the initial conditions:

$$\hat{\hat{\mathcal{P}}}_n = \exp\left[-i\left(\hat{\hat{\mathcal{L}}}_0 + \sum_k \underbrace{c^{(k)}(t_n)}_{\| \; c_n^{(k)}}\hat{\hat{\mathcal{L}}}_k\right)\Delta t\right]$$



The expression for fidelity, the overlap between the current state of the system ρ_0 and the target state σ, is

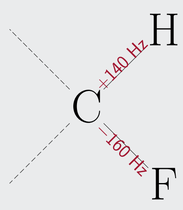
$$J = \mathrm{Re}\,\langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\hat{\hat{\mathcal{P}}}_{N-1}\ldots\hat{\hat{\mathcal{P}}}_2\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle$$

Since $\left\{c_n^{(k)}\right\}$ are vectors of finite dimension, we can use the standard non-linear numerical optimisation to find the maximum of J in their space (it is useful to note that the maximum of J is the same as the minimum of $1-J$).

Simulation System

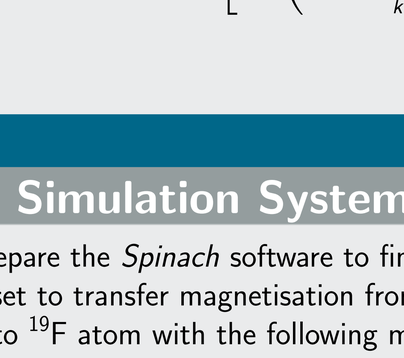
We prepare the *Spinach* software to find a pulse set to transfer magnetisation from ^{1}H atom to ^{19}F atom with the following molecule:

Interaction parameters of a molecular group used in state transfer simulations on a system characterising the fragment of a fluorohydrocarbon molecule (magnetic induction = 9.4 Tesla).

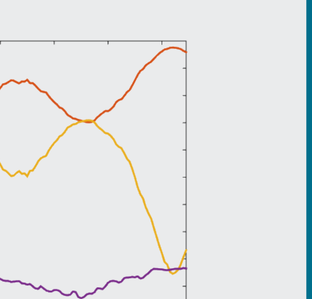


In this case the set of control channels operators are $\left\{\hat{L}_x^{(H)}, \hat{L}_y^{(H)}, \hat{L}_x^{(C)}, \hat{L}_y^{(C)}, \hat{L}_x^{(F)}, \hat{L}_y^{(F)}\right\}$. It is useful to visualise the set of optimal

pulses by the population of their correlation subspaces and population of coherence local at each spin:



The Hessian Matrix is:



- Symmetric,
- Non-singular - one that is invertible,
- Diagonally dominant.

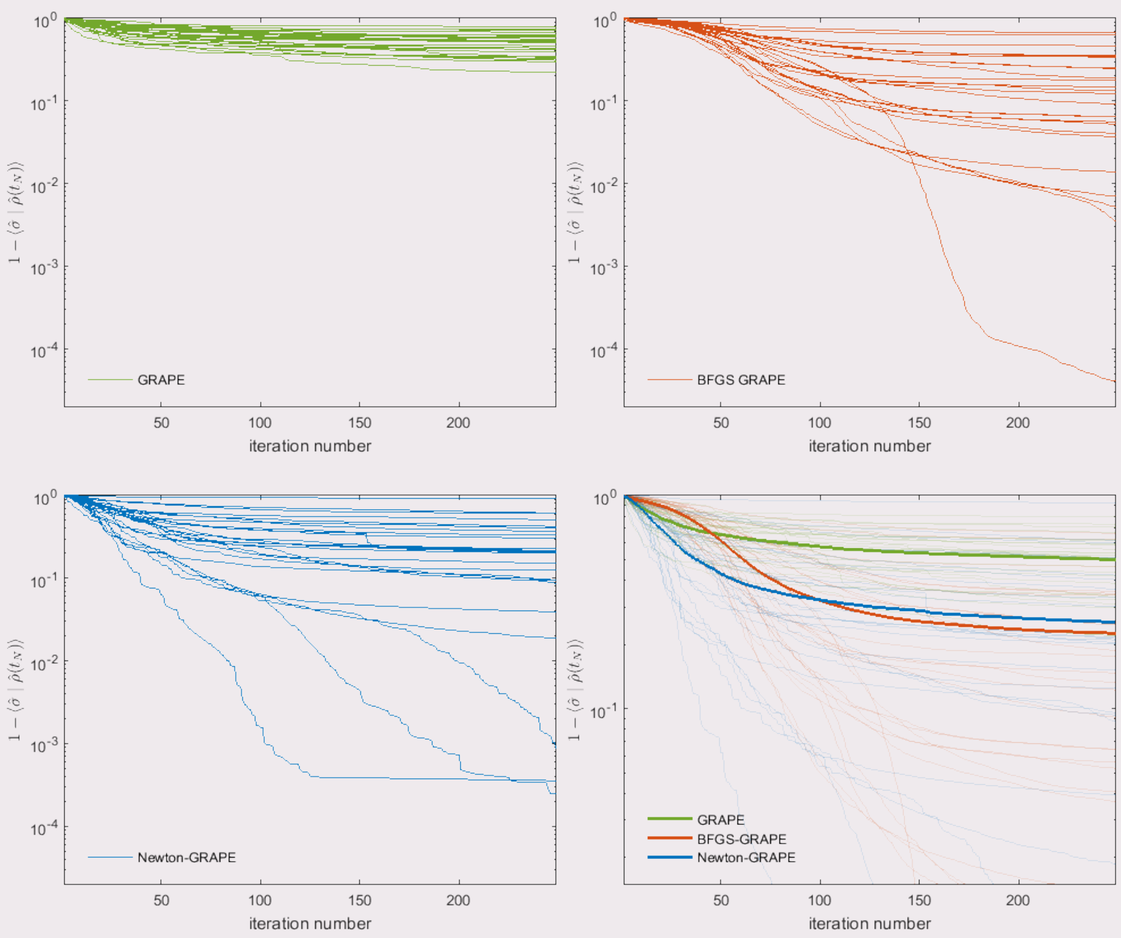
To regularise the Hessian, so it is non-singular and well conditioned, with we take the eigendecomposition of an augmented Hessian matrix:

$$\mathbf{H}_{aug} = \begin{bmatrix}\delta^2\mathbf{H} & \delta\vec{g}\\ \delta\vec{g} & \mathbf{0}\end{bmatrix} = Q\Lambda Q^\dagger$$

where $\vec{g}$ is the gradient, Λ is a diagonal matrix of eigenvalues and Q is a matrix of eigenvectors. We have introduced a constant δ, found iteratively; the region of a radius we trust to give a Hessian that is sufficiently positive definite.

$$\lambda_{min} = \max\left[0, -\min(\Lambda)\right]$$
$$\mathbf{H}_{reg} = Q(\Lambda + \lambda_{min}\hat{I})Q^\dagger$$

Simulation Results



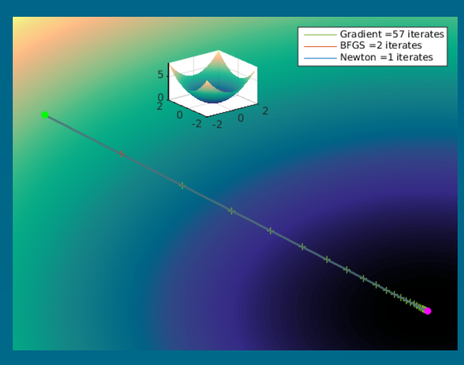
Numerical optimization is concerned with finding an extrema of an objective function. We consider two variables in finding the minimiser of a function:

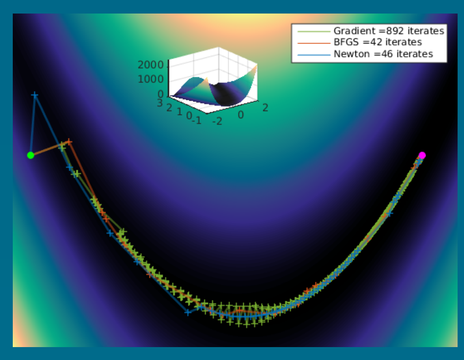
- Step length, α - the distance moved away from the point in the current iterate.
- Search direction, p_k - the direction to move away from the current point in the iterate.

For the line search method, we can find the minimizer considering the above two variables:

$$\min_{\alpha>0} f\left(x_k + \alpha p_k\right)$$

where we begin at some initial point, x_0, and generate a set of iterates using $x_{k+1} = x_k + \alpha_k p_k$, until a solution has been found to some defined level of accuracy. A well defined, convex objective function is best optimized with a Newton-Raphson method.





The Newton-Raphson Method expands the Taylor series to second order;

$$f\left(x_k + p\right) \approx f_x + p^T\nabla f_k + \frac{1}{2}p^T\nabla^2 f_k p$$

where the right hand side of the above formula is minimised by the search direction:

$$p_k = -\mathbf{H}_k^{-1}\nabla f_k$$

where $\mathbf{H}_k = \nabla^2 f_k$ is the Hessian matrix. Quasi-Newton methods use an approximation to the Hessian (using a formula such as the BFGS update). Gradient descent method approximates the Hessian matrix to be the unit matrix

Efficient Propagator Derivative Calculations

The numerical optimisation method simulated in the results section above require a gradient calculation. This is reduced to:

$$J = \langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\hat{\hat{\mathcal{P}}}_{N-1}\hat{\hat{\mathcal{P}}}_{N-2}\hat{\hat{\mathcal{P}}}_{N-3}\underbrace{\hat{\hat{\mathcal{P}}}_{N-4}\ldots\hat{\hat{\mathcal{P}}}_3\hat{\hat{\mathcal{P}}}_2\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle}$$

(I) propagate forwards from source

$$\frac{\partial}{\partial c_{N-3}^{(k)}}\hat{\hat{\mathcal{P}}}_{N-3}$$

(III) compute expectation of the derivative

(II) propagate backwards from target

$$J = \overbrace{\langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\hat{\hat{\mathcal{P}}}_{N-1}\hat{\hat{\mathcal{P}}}_{N-2}}\hat{\hat{\mathcal{P}}}_{N-3}\hat{\hat{\mathcal{P}}}_{N-4}\ldots\hat{\hat{\mathcal{P}}}_3\hat{\hat{\mathcal{P}}}_2\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle$$

The total cost of the gradient of J is therefore one forward simulation, one backward simulation and (n steps) $\times$ (k controls) derivatives of matrix exponentials with respect to scalar parameters. The expectation of first order derivatives is

$$\left\langle\frac{\partial J}{\partial c_{n=t}^{(k)}}\right\rangle = \langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\hat{\hat{\mathcal{P}}}_{N-1}\cdots\frac{\partial}{\partial c_{n=t}^{(k)}}\hat{\hat{\mathcal{P}}}_{n=t}\cdots\hat{\hat{\mathcal{P}}}_2\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle$$

Efficient calculation of the expectation of first order derivatives can be made utilising the work of C.Van Loan; using an augmented exponential in the following form

$$\exp\begin{pmatrix}-i\hat{\hat{\mathcal{L}}}\Delta t & -i\hat{\hat{\mathcal{L}}}_n^{(k)}\Delta t\\ \mathbf{0} & -i\hat{\hat{\mathcal{L}}}\Delta t\end{pmatrix} = \begin{pmatrix}e^{-i\hat{\hat{\mathcal{L}}}\Delta t} & \frac{\partial}{\partial c_n^{(k)}}e^{-i\hat{\hat{\mathcal{L}}}\Delta t}\\ \mathbf{0} & e^{-i\hat{\hat{\mathcal{L}}}\Delta t}\end{pmatrix}$$

extracting the derivative from the upper right block. In practice, the exponential is calculated using a two-point finite difference stencil with Krylov propagation. The Newton-Raphson method is a second order

method, additionally requiring the explicit calculation of the Hessian matrix. This requires the expectation of the second order derivatives:

$$\left\langle\frac{\partial^2 J}{\partial c_n^2}\right\rangle = \langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\cdots\hat{\hat{\mathcal{P}}}_{n+1}\frac{\partial^2\hat{\hat{\mathcal{P}}}_n}{\partial c_n^2}\hat{\hat{\mathcal{P}}}_{n-1}\cdots\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle$$

$$\left\langle\frac{\partial^2 J}{\partial c_n\partial c_{n+1}}\right\rangle = \langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\cdots\hat{\hat{\mathcal{P}}}_{n+2}\frac{\partial\hat{\hat{\mathcal{P}}}_{n+1}}{\partial c_{n+1}}\frac{\partial\hat{\hat{\mathcal{P}}}_n}{\partial c_n}\hat{\hat{\mathcal{P}}}_{n-1}\cdots\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle$$

$$\left\langle\frac{\partial^2 J}{\partial c_m\partial c_n}\right\rangle = \langle\sigma|\,\hat{\hat{\mathcal{P}}}_N\cdots\hat{\hat{\mathcal{P}}}_{n+1}\frac{\partial\hat{\hat{\mathcal{P}}}_n}{\partial c_n}\hat{\hat{\mathcal{P}}}_{n-1}\cdots\hat{\hat{\mathcal{P}}}_{m+1}\frac{\partial\hat{\hat{\mathcal{P}}}_m}{\partial c_m}\hat{\hat{\mathcal{P}}}_{m-1}\cdots\hat{\hat{\mathcal{P}}}_1\,|\rho_0\rangle$$

- Computation to scale with $O(n\times k)$ by storing propagators from gradient calculation.
- Problem now reduces to finding $n\times k$ second-order derivatives on the block diagonal of the Hessian with a 3×3 augmented exponential:

$$\exp\begin{pmatrix}-i\hat{\hat{\mathcal{L}}}\Delta t & -i\hat{\hat{\mathcal{L}}}_n^{(k)}\Delta t & \mathbf{0}\\ \mathbf{0} & -i\hat{\hat{\mathcal{L}}}\Delta t & -i\hat{\hat{\mathcal{L}}}_m^{(k)}\Delta t\\ \mathbf{0} & \mathbf{0} & -i\hat{\hat{\mathcal{L}}}\Delta t\end{pmatrix} = \begin{pmatrix}e^{-i\hat{\hat{\mathcal{L}}}\Delta t} & \frac{\partial}{\partial c_n^{(k)}}e^{-i\hat{\hat{\mathcal{L}}}\Delta t} & \frac{1}{2}\frac{\partial^2}{\partial c_n^{(k)}\partial c_m^{(k)}}e^{-i\hat{\hat{\mathcal{L}}}\Delta t}\\ \mathbf{0} & e^{-i\hat{\hat{\mathcal{L}}}\Delta t} & \frac{\partial}{\partial c_m^{(k)}}e^{-i\hat{\hat{\mathcal{L}}}\Delta t}\\ \mathbf{0} & \mathbf{0} & e^{-i\hat{\hat{\mathcal{L}}}\Delta t}\end{pmatrix}$$

Bibliography

- *Kuprov; J. Mag. Res. 233, 107–112, (2013).*
- *Floether, de Fouquieres, Schirmer,; New J. Phys. 14, 073023, (2012).*
- *de Fouquieres, Schirmer, Glaser, Kuprov; J. Mag. Res. 212, 412-417, (2011).*
- *Hogben, Krzystyniak, Charnock, Hore, Kuprov; J. Mag. Res. 208, 179-194, (2011).*
- *Khaneja, Reiss, Kehlet, Schulte-Herbruggen, Glaser; J. Mag. Res. 172, 296-305, (2005).*
- *Nocedal, Wright; Numerical optimization, (1999).*
- *Najfeld, Havel; Adv. App. Math. 16, 321-375, (1995).*
- *Van Loan; IEEE Trans. 23(3), 395–404, (1978).*