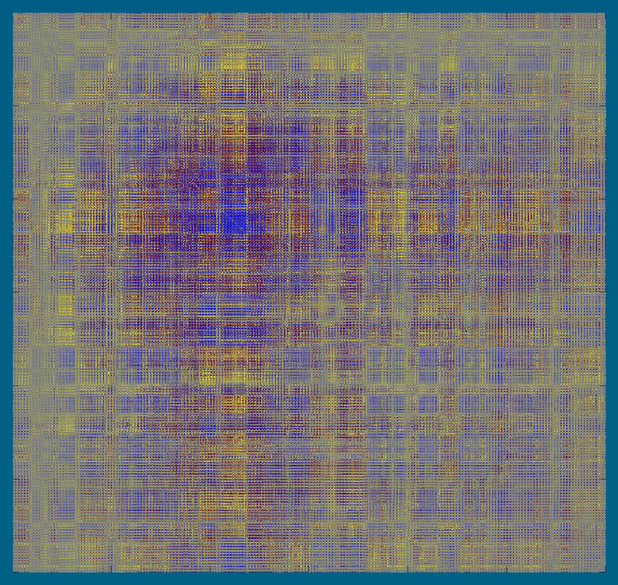
Chemistry

Hessian Regularisation

Krotov-BFGS & GRAPE-Newton-Raphson Methods

David Goodwin, Sean Wright, Ilya Kuprov.

University of Southampton



Introduction

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

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

The Hamiltonian of a practical NMR quantum system can be split into two parts:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_d + \sum_k c_n^{(k)}\hat{\mathcal{H}}_k$$

where $\hat{\mathcal{H}}_d$ is the *drift* or *free Hamiltonian*, and $\hat{\mathcal{H}}_k$ are the parts of the Hamiltonian that can be changed externally, called the *control Hamiltonian*.

We assume that $c_n^{(k)}$ are piecewise-constant, so "slicing" the problem into discrete time intervals Δt. The density matrix can be calculated at a time T using the propagators $\hat{\mathcal{U}}_n$:

$$\rho(T) = \hat{\mathcal{U}}_N\cdots\hat{\mathcal{U}}_1\rho_0\hat{\mathcal{U}}_1^\dagger\cdots\hat{\mathcal{U}}_N^\dagger,\quad \hat{\mathcal{U}}_n = \exp\left(-i\Delta t\left(\hat{\mathcal{H}}_d + \sum_k c_n^{(k)}\hat{\mathcal{H}}_k\right)\right)$$

The state of a magnetic resonance system can be controlled using a set of electromagnetic pulses. Design of these pulses may prove difficult for control of complicated systems; numerical optimisation methods can be used to find a maximum "overlap" between the desired state and the state produced by the set of pulses, with the pulse schedule being the parameter of the objective function in the optimisation problem. Here, the metric of "overlap" is indicated by the fidelity.

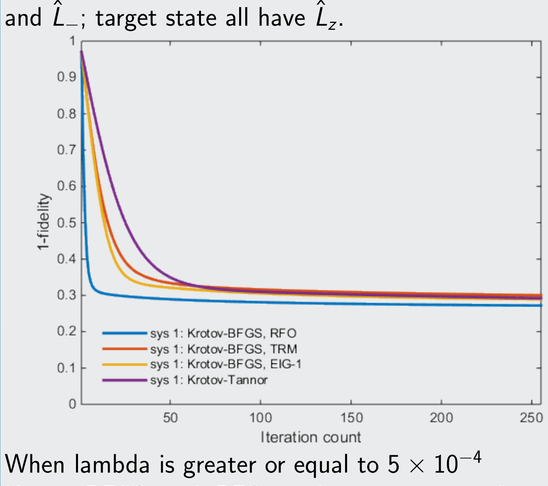
This work will attempt to reproduce the krotov-BFGS optimisation method (Tannor 1992, Eitan 2011) and to develop a scalable Newton-Raphson optimisation method for the GRAPE method (de Fouquieres 2011) in the software toolbox *Spinach* (Hogben 2011).

The work presented here finds that both of these methods need regularisation of the Hessian matrix (the matrix of second order derivatives) for optimal control of most systems.

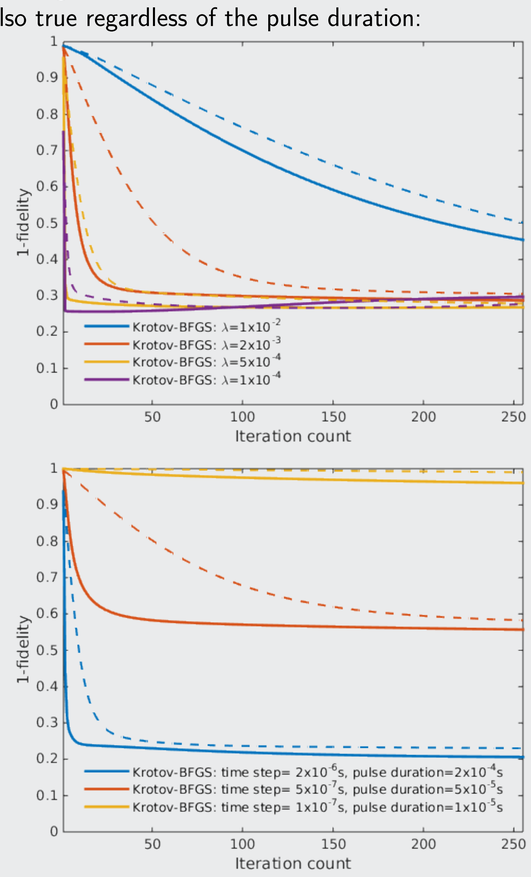
System 1

H-H-H-H-H-H-H-H-H-H-H

11 Hydrogen atoms in an initial state of a mixture of $\hat{L}_+$ and $\hat{L}_-$; target state all have $\hat{L}_z$.



When lambda is greater or equal to 5×10^{-4} Krotov-BFGS, with RFO regularisation, outperformed the standard Krotov-Tannor algorithm; reaching convergence in a smaller number of iterations. This is also true regardless of the pulse duration:

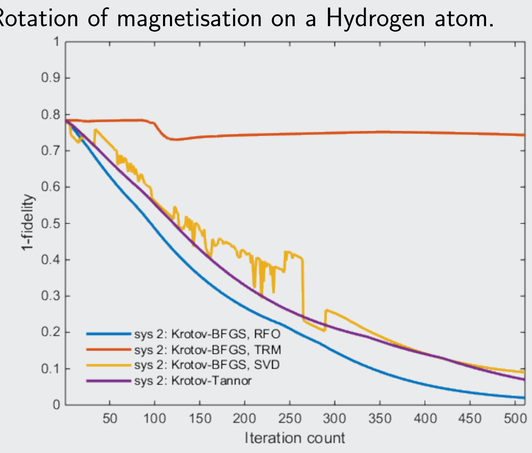


The $\lambda = 1^{-4}$ plot indicates that the algorithm failed to converge and should be investigated further in order to improve the stability of the algorithm.

System 2

H

Rotation of magnetisation on a Hydrogen atom.



Gradient Calculation

Optimisation problems using quasi-Newton methods (e.g. BFGS) require a gradient calculation. This is reduced to:

- Propagate forwards from the source (current state ρ_0)
- Propagate backwards from the target state σ
- Compute the expectation of the derivative:

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

Expectation of Second Order Derivatives

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:

$$\left\langle \sigma\middle|\hat{\mathcal{U}}_N\hat{\mathcal{U}}_{N-1}\cdots\frac{\partial^2}{\partial {c_{n=t}^{(k)}}^2}\hat{\mathcal{U}}_{n=t}\cdots\hat{\mathcal{U}}_2\mathcal{U}_1\middle|\rho_0\right\rangle$$

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

$$\left\langle \sigma\middle|\hat{\mathcal{U}}_N\hat{\mathcal{U}}_{N-1}\cdots\frac{\partial}{\partial c_{n=t_2}^{(k)}}\hat{\mathcal{U}}_{n=t_2}\cdots\frac{\partial}{\partial c_{n=t_1}^{(k)}}\hat{\mathcal{U}}_{n=t_1}\cdots\hat{\mathcal{U}}_2\mathcal{U}_1\middle|\rho_0\right\rangle$$

- Computation to scale with $O(n)$ by storing propagators from gradient calculation.
- Problem now reduces to finding n second-order derivatives on the block diagonal of the Hessian.

Hessian Element Calculations: Augmented Exponentials

- based on Sophie Schirmer's method (unpublished) for calculating gradients based on an augmented exponential (Van Loan 1978).

$$\exp\begin{pmatrix} -i\hat{\hat{L}}\Delta t & -i\hat{\hat{H}}_{n_1}^{(k_1)}\Delta t & \mathbf{0} \\ \mathbf{0} & -i\hat{\hat{L}}\Delta t & -i\hat{\hat{H}}_{n_2}^{(k_2)}\Delta t \\ \mathbf{0} & \mathbf{0} & -i\hat{\hat{L}}\Delta t \end{pmatrix} = \begin{pmatrix} e^{-i\hat{\hat{L}}\Delta t} & \frac{\partial}{\partial c_{n_1}^{(k_1)}}e^{-i\hat{\hat{L}}\Delta t} & \frac{1}{2}\frac{\partial^2}{\partial c_{n_1}^{(k_1)}\partial c_{n_2}^{(k_2)}}e^{-i\hat{\hat{L}}\Delta t} \\ \mathbf{0} & e^{-i\hat{\hat{L}}\Delta t} & \frac{\partial}{\partial c_{n_2}^{(k_2)}}e^{-i\hat{\hat{L}}\Delta t} \\ \mathbf{0} & \mathbf{0} & e^{-i\hat{\hat{L}}\Delta t} \end{pmatrix}$$

Krotov-BFGS

GRAPE optimises the cost functional (fidelity) with respect to the control pulse sequence. The Krotov approach (Tannor 1992) is to find an iterative proceedure that guarantees improvement to the cost functional at every step. This is achieved by solving Euler-Lagrange equations:

$$J = \langle\psi(T)|\sigma\rangle\langle\sigma|\psi(T)\rangle - \mathrm{Re}\int_0^T\left\langle\chi\middle|\frac{d}{dt} + i\hat{\mathcal{H}}_0 + i\sum_k c^{(k)}(t)\hat{\mathcal{H}}_k\middle|\psi\right\rangle - \lambda\sum_k dt\left\|c^{(k)}(t)\right\|^2$$

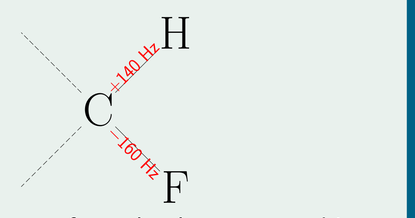
$$\frac{d}{dt}|\psi_n\rangle = -i\left(\hat{\mathcal{H}}_0 + \sum_k c_n^{(k)}(t)\hat{\mathcal{H}}_k\right)|\psi_n\rangle \qquad \frac{d}{dt}|\chi_n\rangle = -i\left(\hat{\mathcal{H}}_0 + \sum_k \tilde{c}_n^{(k)}(t)\hat{\mathcal{H}}_k\right)|\chi_n\rangle$$

$$c_n^{(k)}(t) = c_{n-1}^{(k)}(t) + \frac{1}{\lambda}\mathrm{Im}\left\langle\chi_{n-1}\middle|\hat{\mathcal{H}}_k\middle|\psi_n\right\rangle \qquad \tilde{c}_n^{(k)}(t) = \tilde{c}_{n-1}^{(k)}(t) + \frac{1}{\lambda}\mathrm{Im}\left\langle\chi_n\middle|\hat{\mathcal{H}}_k\middle|\psi_n\right\rangle$$

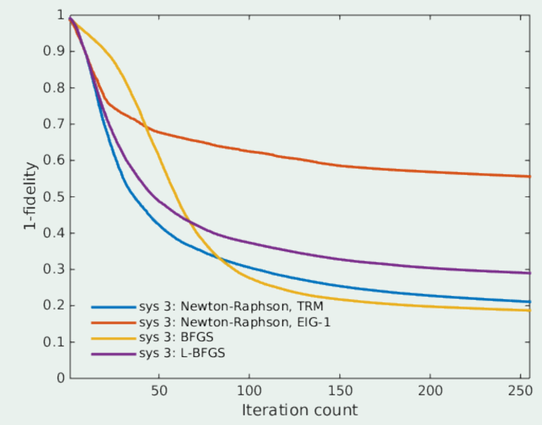
where $|\psi_n(0)\rangle \equiv |\text{initial state}\rangle$ and $|\chi(T)\rangle \equiv |\text{target state}\rangle\langle\text{target state}|\psi(T)\rangle$.

- *de Fouquieres, Schirmer, Glaser, Kuprov; J. Mag. Res., 212, 412-417, (2011).*
- *Hogben, Krzystyniak, Charnock, Hore, Kuprov; J. Mag. Res., 208, 179-194, (2011).*
- *Van Loan; IEEE Trans., 23(3), 395–404, (1978).*
- *Tannor, Kazakov, Orlov; Time-Dependent Quantum Molecular Dynamics, Plenum, 347-360, (1992).*
- *Eitan, Mundt, Tannor; Phys. Rev. A, 83, 053426 (2011).*

System 3



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



The Hessian Matrix

The Hessian Matrix is:

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

We must regularise the Hessian matrix so it is non-singular and well conditioned.

EIG-1

Take the eigendecomposition of the Hessian matrix:

$$\mathbf{H} = Q\Lambda Q^\dagger$$

where Λ is a diagonal matrix of eigenvalues and Q is a matrix whose columns are the corresponding eigenvectors.

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

TRM

Similar to the eigendecomposition method above, except we introduce a constant δ; the region of a radius we trust to give a Hessian that is sufficiently positive definite.

$$\mathbf{H}_{reg} = Q(\Lambda + \delta\lambda_{min}\hat{I})Q^\dagger$$

RFO

This is the same as the TRM method, except we make the eigendecomposition of an augmented Hessian:

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

where $\vec{g}$ is the gradient (vector is first order derivatives).

SVD

Similar to the eigendecomposition method, except we regularise according to the singular values of the Hessian matrix.