Multilevel methods for parameter identification problems

Alfio Borzì

Institute for Mathematics and Scientific Computing
Karl-Franzens-University, Graz, Austria

Thirty-second Conference of the Dutch and Flemish Numerical Analysis Communities, Woudschoten Conference, Zeist, The Netherlands, 3 - 5 October, 2007

Research framework

Distributed parameter estimation problems represent an important class of inverse problems with a variety of important applications.

From the point of view of control, bilinear control problems represent a class of nonlinear control strategies with the aim to obtain better system response than possible with linear control.

Inverse Helmholtz problems and quantum control problems

Multigrid methods applied on the full space and on the reduced space

CSMG – FAS with Collective Smoothing

MGOPT

Optimization with PDE constraints

$$\min_{u \in U} J(y,u) \qquad J : Y \times U \to \mathbb{R}$$
$$\text{s.t.} \quad c(y,u) = 0$$

The existence of c_y^{-1} enables a clear distinction between y, the state variable, and $u \in U$, the optimization variable in the admissible set. So we have the mapping $u \mapsto J(y(u), u)$ in the form

$$u \overset{\text{IFT}}{\mapsto} y(u) \mapsto J(y(u), u) =: \hat{J}(u)$$

The solution of this optimization problem is characterized by the following optimality system

$$\begin{aligned} c(y,u) &= 0 \\ c_y(y,u)^* \, p &= -h'(y) \\ \nu \, g'(u) + c_u^* \, p &= 0 \end{aligned}$$

assuming $J(y,u) = h(y) + \nu \, g(u)$, $\nu > 0$. We have $\nabla \hat{J}(u) = \nu \, g'(u) + c_u^* \, p(u)$

Bilinear control problems – Parameter identification

Helmholtz optimal control problem

$$\text{minimize} \quad J(y,u) := \frac{1}{2}\int_\Omega (y-z)^2\, d\Omega + \frac{\beta}{2}\left(\int_\Omega u^2\, d\Omega + \int_\Omega \nabla u \cdot \nabla u \, d\Omega\right)$$

$$\text{subject to} \qquad -\Delta y + uy = f \quad \text{in} \quad \Omega, \qquad y = 0 \quad \text{on} \quad \Gamma.$$

Bose–Einstein condensates control problem

$$\text{minimize} \quad J(\psi,u) := \frac{1}{2}\left(1 - \left|\langle \psi_d | \psi(T)\rangle\right|^2\right) + \frac{\gamma}{2}\int_0^T (\dot{u}(t))^2 \; dt$$

$$\text{subject to} \qquad i\dot{\psi}(x,t) = \left(-\frac{1}{2}\nabla^2 + V(x,u(t)) + g\,|\psi(x,t)|^2\right)\psi(x,t)$$

Helmholtz optimal control problem

$$\text{minimize} \quad J(y,u) := \frac{1}{2}\int_\Omega (y-z)^2 \, d\Omega + \frac{\beta}{2}\left(\int_\Omega u^2 \, d\Omega + \int_\Omega \nabla u \cdot \nabla u \, d\Omega\right)$$

$$\text{subject to} \quad -\Delta y + uy = f \quad \text{in} \quad \Omega, \quad y = 0 \quad \text{on} \quad \Gamma.$$

Lagrangian

$$L(y,u,p) = J(y,u) + [\int_\Omega \nabla p \cdot \nabla y + u\,p\,y \, d\Omega - \int_\Omega pf \, d\Omega]$$

Optimality system

Forward

$$-\Delta y + u\,y = f \quad \text{in} \quad \Omega, \quad y = 0 \quad \text{on} \quad \Gamma$$

Adjoint

$$-\Delta p + u\,p + y = z \quad \text{in} \quad \Omega, \quad y = 0 \quad \text{on} \quad \Gamma$$

Inverse

$$-\Delta u + u + y\,p/\beta = 0 \quad \text{in} \quad \Omega, \quad u = u_b \quad \text{on} \quad \Gamma$$

Bose–Einstein condensates control problem

$$\text{minimize} \quad J(\psi, u) := \frac{1}{2}\left(1 - \left|\langle \psi_d | \psi(T) \rangle\right|^2\right) + \frac{\gamma}{2}\int_0^T \left(\dot{u}(t)\right)^2 \, dt$$

$$\text{subject to} \qquad i\dot{\psi}(x,t) = \left(-\frac{1}{2}\nabla^2 + V(x, u(t)) + g\left|\psi(x,t)\right|^2\right)\psi(x,t)$$

Optimality system

$$i\dot{\psi} = \left(-\frac{1}{2}\nabla^2 + V_u + g|\psi|^2\right)\psi$$

$$i\dot{p} = \left(-\frac{1}{2}\nabla^2 + V_u + 2g|\psi|^2\right)p + g\,\psi^2\,p^*$$

$$\gamma\ddot{u} = -\Re e\langle\psi|\frac{\partial V_u}{\partial u}|p\rangle\,,$$

with the initial and terminal conditions

$$\psi(0) = \psi_0 \text{ and } ip(T) = -\langle\psi_d|\psi(T)\rangle\,\psi_d$$

$$u(0) = 0 \quad u(T) = 1$$

Multigrid strategies

CSMG – apply classical multigrid techniques to the PDE optimality system. Full space approach.

Design of appropriate smoothing schemes

MGOPT – the multigrid cycling structure defines an outer iteration and classical optimization schemes represent the inner iteration loop to minimize $\hat{J}$. Reduced space approach.

Choice of coarse spaces, ...

CSMG – Collective Smoothing Multi-Grid

- **Multigrid FAS-(m_1, m_2) method for solving $A_k(u_k) = f_k$.**
  1. If $k = 1$ solve $A_k(u_k) = f_k$ directly (e.g., repeated application of S_k).
  2. Pre-smoothing steps on the fine grid: $u_k^{(l)} = S_k(u_k^{(l-1)}, f_k)$, $l = 1, \ldots, m_1$;
  3. Computation of the residual: $r_k = f_k - A_k(u_k^{(m_1)})$;
  4. Restriction of the residual: $r_{k-1} = I_k^{k-1} r_k$;
  5. Set $u_{k-1} = \dot{I}_k^{k-1} u_k^{(m_1)}$;
  6. Set $f_{k-1} = r_{k-1} + A_{k-1}(u_{k-1})$
  7. Call m times FAS-(m_1, m_2) to solve $A_{k-1}(u_{k-1}) = f_{k-1}$;
  8. Coarse-grid correction: $u_k^{(m_1+1)} = u_k^{(m_1)} + I_{k-1}^k (u_{k-1} - \dot{I}_k^{k-1} u_k^{(m_1)})$;
  9. Post-smoothing steps on the fine grid: $u_k^{(l)} = S_k(u_k^{(l-1)}, f_k)$, $l = m_1 + 2, \ldots, m_1 + m_2 + 1$;

Collective smoothing iteration – Helmholtz problem

Define Δ_h in the compact form

$$\Delta_h v_h|_{ij} = \frac{1}{h^2} \Big(\sum_{s,t \in \omega_{ij},\, s,t \neq i,j} c_{st}\, v_{st} - c_{ij} v_{ij} \Big).$$

Set $A_{ij} = \sum_{s,t \in \omega_{ij},\, s,t \neq i,j} c^y_{st}\, y_{st} + h^2 \tilde{f}_{ij}$ similarly $p \rightarrow B_{ij}$ and $u \rightarrow C_{ij}$. The following system for the three scalar variables y_{ij}, p_{ij}, and u_{ij} is obtained

$$\begin{aligned} -A_{ij} + c^y_{ij}\, y_{ij} + h^2\, u_{ij}\, y_{ij} &= 0 \\ -B_{ij} + c^p_{ij}\, p_{ij} + h^2\, u_{ij}\, p_{ij} + h^2\, y_{ij} &= 0 \\ -C_{ij} + c^u_{ij}\, u_{ij} + h^2\, u_{ij} + h^2\, y_{ij}\, p_{ij}/\beta &= 0 \end{aligned}$$

It admits multiple solutions represented by the zeros of a quartic polynomial equation in u_{ij}. Two of these solutions are complex conjugate. Numerical instabilities of standard iterations occur because of the presence of the two real solutions.

Quartic polynomial equation

Construct the quartic polynomial equation and solve it exactly (off-line).
We have $y_{ij} = y_{ij}(u_{ij})$ and $p_{ij} = p_{ij}(u_{ij})$ as functions of u_{ij} as follows

$$y_{ij}(u_{ij}) = A_{ij}/(c^y_{ij} + h^2\, u_{ij})$$

and

$$p_{ij}(u_{ij}) = \frac{(c^y_{ij}\, B_{ij} - A_{ij} h^2 + B_{ij} h^2 u_{ij})}{(c^p_{ij} + h^2 u_{ij})(c^y_{ij} + h^2 u_{ij})}$$

We assume that $c^p_{ij} + h^2 u_{ij} \neq 0$ and $c^y_{ij} + h^2 u_{ij} \neq 0$ at any ij.

The quartic polynomial equation in u_{ij} results from the optimality condition

$$-C_{ij} + c^u_{ij}\, u_{ij} + h^2\, u_{ij} + h^2\, y_{ij}(u_{ij})\, p_{ij}(u_{ij})/\beta = 0$$

Smoothing step

The quartic polynomial equation in u_{ij} admits four solutions. Two of them are complex and two are real solutions. What is the right one ?

The solution $\tilde{u}_{ij}$ that locally minimizes the ij-component of J

$$J_{ij}(\tilde{u}_{ij}) = \frac{1}{2}h^2\left(y_{ij}(\tilde{u}_{ij}) - z_{ij}\right)^2 + \frac{\beta}{2}\left(h^2\left(\tilde{u}_{ij}\right)^2 + \frac{1}{2}(u_{i\pm1j} - \tilde{u}_{ij})^2 + \frac{1}{2}(u_{ij\pm1} - \tilde{u}_{ij})^2\right)$$

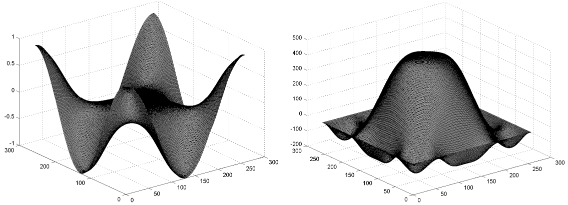
in the spirit of subspace correction schemes

Given $u_{ij} = \tilde{u}_{ij}$ one obtains the updates

$$y_{ij}(u_{ij}) \qquad \text{and} \qquad p_{ij}(u_{ij})$$

A robust and efficient smoothing iteration results.

Numerical results



Solution of the optimal control problem with $y = 0$ on Γ; $\beta = 10^{-7}$. The state (left) and the control (right); 256×256 mesh.

Convergence and tracking properties; Dirichlet b.c.

β	ρ	$\|y - z\|_{L^2(\Omega)}$
10^{-5}	0.09	$3.5\,10^{-1}$
10^{-7}	0.09	$2.9\,10^{-1}$
10^{-9}	0.3	$2.5\,10^{-1}$
10^{-11}	0.8^*	$1.8\,10^{-1}$

We take $z = \cos(2\pi x_1)\cos(2\pi x_2)$.

MGOPT framework

The MGOPT solution to the optimization problem $\min_u \hat{J}(u)$ requires to define a hierarchy of minimization problems

$$\min_{u_k} \hat{J}_k(u_k) \qquad k = 1, 2, \ldots, L$$

where $u_k \in V_k$ and $\hat{J}_k(\cdot)$ is the reduced cost functional.

Among spaces V_k, restriction operators $I_k^{k-1} : V_k \rightarrow V_{k-1}$ and prolongation operators $I_{k-1}^k : V_{k-1} \rightarrow V_k$ are defined.
Require that $(I_k^{k-1}u, v)_{k-1} = (u, I_{k-1}^k v)_k$ for all $u \in V_k$ and $v \in V_{k-1}$.

We also choose an optimization scheme as 'smoother'

$$u_k^{(l)} = O_k\left(u_k^{(l-1)}\right)$$

That provides sufficient reduction

$$\hat{J}_k(O_k(u_k^{(l)})) < \hat{J}_k(u_k^{(l)}) - \eta \|\nabla \hat{J}_k(u_k^{(l)})\|^2$$

for some $\eta \in (0, 1)$.

MGOPT method

Step 1. If $k = 1$ solve $\min_{u_k}(\hat{J}_k(u_k) - f_k u_k)$ directly, i.e. $\nabla \hat{J}_k(u_k) = f_k$.

Step 2. Pre-optimization. $u_k^{(l)} = O_k\,(u_k^{(l-1)})$, $l = 1, \ldots, m_1$. $(f_L = 0.)$

Step 3. Computation of the fine-to-coarse gradient correction

$$\phi_{k-1} = \nabla \hat{J}_{k-1}(I_k^{k-1} u_k^{(m_1)}) - I_k^{k-1} \nabla \hat{J}_k(u_k^{(m_1)}), \qquad f_{k-1} = I_k^{k-1} f_k + \phi_{k-1}.$$

Step 4. Call m times MGOPT to solve $J_{k-1}(\tilde{u}_{k-1}) = \min_{u_{k-1}} J_{k-1}(u_{k-1})$ where

$$J_{k-1}(u_{k-1}) = \hat{J}_{k-1}(u_{k-1}) - f_{k-1} u_{k-1}.$$

Step 5. Coarse-to-fine minimization step with line-search (α) given by

$$u_k^{(m_1+1)} = u_k^{(m_1)} + \alpha\, I_{k-1}^{k}(\tilde{u}_{k-1} - I_k^{k-1} u_k^{(m_1)}).$$

Step 6. Post-optimization. $u_k^{(l)} = O_k\,(u_k^{(l-1)})$, $l = m_1 + 2, \ldots, m_1 + m_2 + 1$.

Classical optimization scheme: The NCG method

For pre- and post-optimization we can use the nonlinear conjugate gradient scheme to minimize the locally convex $\hat{J}(u)$. Denote with $g(u) = \nabla \hat{J}(u)$.

Note: Evaluation of $\hat{J}(u_k)$ and of $g_k = \nabla \hat{J}_k(u_k)$ requires a forward (state) and a backwards (adjoint) solve.

We have $d_{k+1} = -g_{k+1} + \beta_k \, d_k$ where (Dai & Yuan SIOPT '99)

$$\beta_k = \frac{\|g_{k+1}\|^2}{(d_k, y_k)}.$$

We require that the steplength α_k for $u_{k+1} = u_k + \alpha_k \, d_k$ satisfies

$$\hat{J}(u_k) - \hat{J}(u_k + \alpha_k d_k) \geq -\delta \alpha_k \, (g_k, d_k)$$

$$(g(u_k + \alpha_k d_k), d_k) > \sigma g(_k, d_k)$$

where $0 < \delta < \sigma < 1/2$.

NCG Scheme

Step 1. Given $k = 1$, u_1, $d_1 = -g_1$, if $\|g_1\| < tol$ then stop.

Step 2. Compute $\tau_k > 0$ satisfying the standard Wolfe conditions.

Step 3. Let $u_{k+1} = u_k + \tau_k \, d_k$.

Step 4. Compute $g_{k+1} = \nabla \hat{J}(u_{k+1})$.
If $\|g_{k+1}\| < tol_{abs}$ or $\|g_{k+1}\| < tol_{rel} \, \|g_1\|$ or $k = k_{\max}$ then stop.

Step 5. Compute $\beta_k = \frac{\|g_{k+1}\|^2}{(d_k, y_k)}$.

Step 6. Let $d_{k+1} = -g_{k+1} + \beta_k \, d_k$.

Step 7. Set $k = k + 1$, goto Step 2.

Cascadic acceleration

The cascadic approach results from combining nested iteration techniques with a (one-grid) iterative scheme.
$k = k_0, \ldots, k_f$ index of grid hierarchy.
u_{k_0} given starting approximation on the coarsest grid.
I_k^{k+1} interpolation operator from k to $k+1$.
$NCG_k(u_k)$ the basic iteration; $*$ denotes the resulting solution.

Cascadic NCG (CNCG) method

Step 1. Given $k = k_0$, $u_{k_0}^*$.

Step 2. Compute $u_k = NCG_k(u_k^*)$.

Step 3. If $k = k_f$ then stop.

Step 4. Else if $k < k_f$ then interpolate $u_{k+1}^* = I_k^{k+1} u_k$.

Step 5. Set $k = k+1$, goto Step 2.

Transport of Bose Einstein condensates in magnetic microtraps

We consider transport of Bose-Einstein condensates in magnetic microtraps, controllable by external parameters such as wire currents or radio-frequency fields.

The mean-field dynamics of the condensate is described by the Gross-Pitaevskii equation (GPE)

$$i\dot{\psi}(x,t) = \left(-\frac{1}{2}\nabla^2 + V(x,u(t)) + g\,|\psi(x,t)|^2\right)\psi(x,t)$$

$V(x,u(t))$ is a three-dimensional potential produced by a magnetic microtrap. $u(t)$ is a control parameter that describes the variation of the confining potential.

Purpose of control

Through $u(t)$ it is possible to manipulate the Bose-Einstein condensate, e.g., to split and reunite it by varying the potential from a single to a double well.

Suppose that initially the system is in the ground state ψ_0 for the potential $V(x, u = 0)$.

We seek for optimal time evolution of $u(t)$ that allows to channel the system from the initial state ψ_0 at time zero to a desired state ψ_d to be the ground state for the potential $V(x, u = 1)$ at time T.

Optimal control formulation and optimality system

$$J(\psi, u) = \frac{1}{2}\left(1 - \left|\langle \psi_d | \psi(T) \rangle\right|^2\right) + \frac{\gamma}{2} \int_0^T (\dot{u}(t))^2 \, dt$$

Optimal control problem: Minimize the cost function $J(\psi, u)$ subject to the condition that ψ fulfills the Gross-Pitaevskii equation.
The optimal solution is characterized by the optimality system

$$i\dot{\psi} = \left(-\frac{1}{2}\nabla^2 + V_u + g|\psi|^2\right)\psi$$

$$i\dot{p} = \left(-\frac{1}{2}\nabla^2 + V_u + 2g|\psi|^2\right)p + g\,\psi^2\,p^*$$

$$\gamma\ddot{u} = -\Re e\langle\psi|\frac{\partial V_u}{\partial u}|p\rangle\,,$$

with the initial and terminal conditions

$$\psi(0) = \psi_0 \text{ and } ip(T) = -\langle\psi_d|\psi(T)\rangle\,\psi_d$$

$$u(0) = 0\,, \quad u(T) = 1\,.$$

Control potential and the gradient

Consider the double-well potential

$$V(x,u) = -\frac{u^2\, d^2}{8c}\, x^2 + \frac{1}{c}\, x^4$$

where $c = 40$ and d is a parameter corresponding to twice the distance of the two minima in the double well potential.
In terms of u we have the reduced cost functional

$$\hat{J}(u) = J(\psi(u), u)$$

where $\psi(u)$ denotes the unique solution to the GPE for u.
The gradient of $\hat{J}$ is given by

$$\nabla \hat{J}(u) = -\gamma\, \frac{d^2 u}{dt^2} - \Re e \langle \psi | \frac{\partial V_u}{\partial u} | p \rangle,$$

where ψ and p solve the state and the adjoint equations with u.

Discretization scheme

To discretize the state and adjoint equations we use a unconditionally stable second-order norm-preserving time-splitting spectral scheme

$$\psi^{m+1} = e^{-i\,\frac{\delta t}{2}\,V^{m+1}}\,e^{-i\delta t\,H_0}\,e^{-i\,\frac{\delta t}{2}\,V^m}\,\psi^m$$

The presence of the term $g\,\psi^2\,p^*$ in the adjoint equation requires additional work. We find

$$\begin{pmatrix} p_r \\ p_i \end{pmatrix}(t+\delta t) = \exp\left(i\,\bar{u}\cdot\bar{\sigma}\,\delta t\right)\begin{pmatrix} p_r \\ p_i \end{pmatrix}(t)$$

where $p = p_r + i\,p_i$ and $\bar{u} = (ia_r, A, -ia_i)$. We set $A = V_u + 2g|\psi|^2$ and $g\,\psi^2 = a_r + i\,a_i$.
Here $\bar{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ denotes the vector of the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Discrete gradient and ground states

Evaluation of the gradient of the reduced cost functional is given by the following

$$\nabla \hat{J}(u)^m = -\gamma \, \frac{u^{m+1} - 2u^m + u^{m-1}}{\delta t^2} - \Re e \sum_{j=1}^{N} h(p_j^m)^* \frac{\partial V_u}{\partial u}(u^m) \psi_j^m .$$

The initial state ψ_0 and the target state ψ_d are the groundstate wavefunctions of the Gross–Pitaevskii equation with single- $(u = 0)$ and double- $(u = 1)$ well potential, respectively.
To determine these states we consider the evolution of the Gross–Pitaevskii equation with δt replaced by $-i\delta t$ and at each step the wavefunction is normalized.

Computational performance of CNCG and MGOPT

		CNCG		MGOPT	
γ	T	$\frac{1}{2}\left(1 - \left\|\langle\psi_d, \psi(T)\rangle\right\|^2\right)$	CPU	$\frac{1}{2}\left(1 - \left\|\langle\psi_d, \psi(T)\rangle\right\|^2\right)$	CPU
10^{-1}	5	$1.49 \cdot 10^{-1}$	112	$4.23 \cdot 10^{-2}$	941
10^{-3}	5	$1.40 \cdot 10^{-2}$	825	$2.97 \cdot 10^{-3}$	515
10^{-5}	5	$1.29 \cdot 10^{-2}$	205	$4.56 \cdot 10^{-3}$	213
10^{-1}	10	$3.23 \cdot 10^{-3}$	473	$4.38 \cdot 10^{-4}$	625
10^{-3}	10	$1.39 \cdot 10^{-3}$	239	$1.19 \cdot 10^{-4}$	930
10^{-5}	10	$3.63 \cdot 10^{-3}$	65	$2.27 \cdot 10^{-4}$	425

Table: Computational performance of the CNCG and MGOPT schemes. Mesh 256×2500.

Computational performance of CNCG and MGOPT

		CNCG		MGOPT	
γ	*Mesh*	$\frac{1}{2}\left(1 - \left\|\langle\psi_d, \psi(T)\rangle\right\|^2\right)$	CPU	$\frac{1}{2}\left(1 - \left\|\langle\psi_d, \psi(T)\rangle\right\|^2\right)$	CPU
10^{-2}	f	$1.26 \cdot 10^{-2}$	580	$6.70 \cdot 10^{-4}$	695
10^{-4}	f	$5.13 \cdot 10^{-4}$	90	$5.47 \cdot 10^{-4}$	299
10^{-6}	f	$6.47 \cdot 10^{-4}$	77	$4.54 \cdot 10^{-4}$	758
10^{-2}	c	$2.23 \cdot 10^{-2}$	17	$9.69 \cdot 10^{-4}$	116
10^{-4}	c	$4.54 \cdot 10^{-4}$	202	$6.01 \cdot 10^{-4}$	82
10^{-6}	c	$1.38 \cdot 10^{-2}$	14	$8.78 \cdot 10^{-4}$	78

Table: Computational performance of the CNCG and MGOPT schemes; $T = 7.5$. $c = 128 \times 1250$, $f = 256 \times 2500$.

Computational performance of CNCG and MGOPT

	CNCG		MGOPT	
g	$\frac{1}{2}\left(1-\left|\langle\psi_d,\psi(T)\rangle\right|^2\right)$	CPU	$\frac{1}{2}\left(1-\left|\langle\psi_d,\psi(T)\rangle\right|^2\right)$	CPU
25	$3.89 \cdot 10^{-4}$	53	$7.08 \cdot 10^{-4}$	149
50	$2.35 \cdot 10^{-3}$	80	$9.84 \cdot 10^{-3}$	76
75	$5.54 \cdot 10^{-3}$	90	$1.85 \cdot 10^{-3}$	163
100	$4.93 \cdot 10^{-1}$	13	$2.47 \cdot 10^{-1}$	27
100	$4.94 \cdot 10^{-1}$	50	$5.44 \cdot 10^{-3}$	257

Table: Computational performance of the CNCG and MGOPT schemes for different values of g; $T = 7.5$, $\gamma = 10^{-4}$, mesh 128×1250.

Time evolution for linear and optimized control

The linear $u(t) = t/T$ is a guess for the optimal control iterative process.

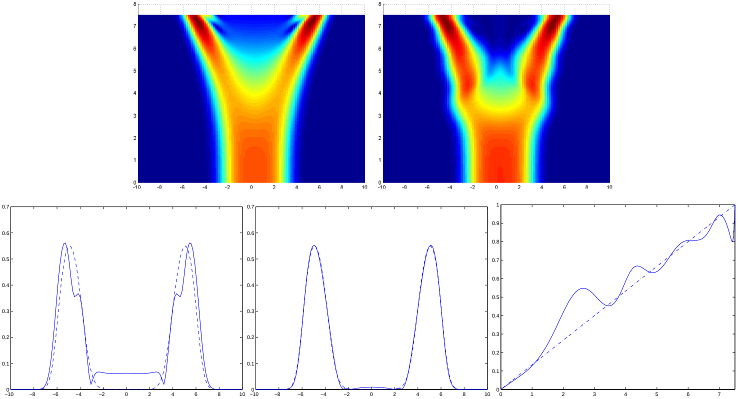
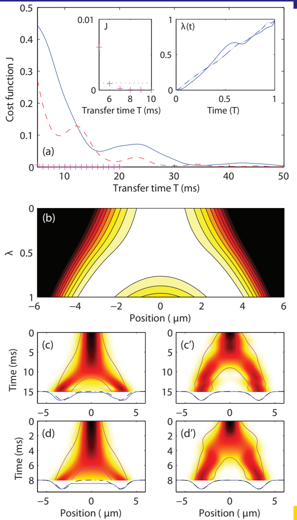


Figure: The function $|\psi(x,t)|$ on the space-time domain (top) for the linear (left) and optimized (right) control. The corresponding profiles at $t = T$ (bottom, continuous line) compared to the desired state (dashed line). The tracking error $\frac{1}{2}\left(1 - |\langle \psi_d, \psi(T)\rangle|^2\right)$ results $6.26\,10^{-2}$ (lin) and $1.22\,10^{-3}$ (opt). MGOPT, Mesh 128×1250; [illegible] 10^{-4}

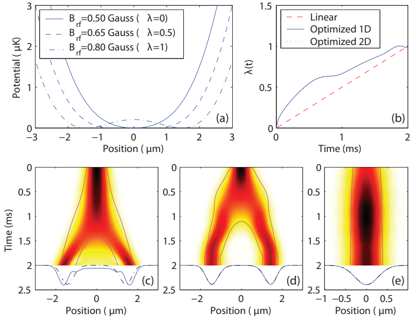
Magnetic confinement of Hänsel et a



Wavefunction splitting with a control scheme that favours adiabatic transport by minimizing excitations to higher states.

Magnetic confinement of Lesanovsky et al.

Radio-frequency double-well confinement proposed by Lesanovsky *et al.* which is produced by a surface-mounted dc four-wire structure on an atom chip.



Choice of spaces: L^2 vs. H^1 control space

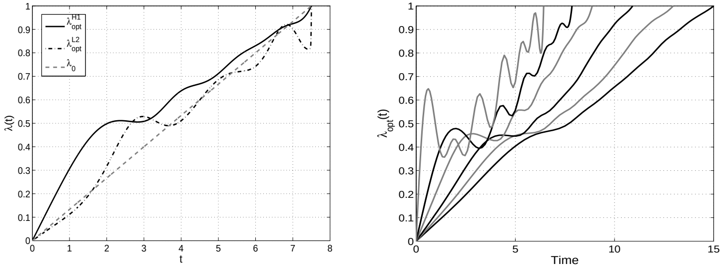


Figure: Left: Optimal control function $\lambda(t)$ for $T = 7.5$, $\gamma = 1 \times 10^{-4}$, on 2500 grid points. Right: The H^1 optimal control function for the linear Schrödinger equation for various control intervals.

Theoretical analysis of MGOPT

We assume that for each k, $\hat{J}_k$ is twice Frechét differentiable and $\nabla^2 \hat{J}_k$ is (locally) positive definite and satisfies the conditions

$$(\nabla^2 \hat{J}_k(u)y, y)_k \geq \beta \|y\|_k^2 \qquad \text{and} \qquad \|\nabla^2 \hat{J}_k(u) - \nabla^2 \hat{J}_k(y)\| \leq \rho \, \|u - y\|_k$$

uniformly for some positive constants β and ρ.
We remark that

$$\nabla \left(\hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\lambda_{k-1} \right) |_{\lambda_{k-1} = I_k^{k-1}\tilde{\lambda}_k} = I_k^{k-1} \nabla \hat{J}_k(\tilde{\lambda}_k),$$

We use the expansion

$$\hat{J}_k(\lambda + z) = \hat{J}_k(\lambda) + (\nabla \hat{J}_k(\lambda), z)_k + \frac{1}{2} \int_0^1 (\nabla^2 \hat{J}_k(\lambda + tz)z, z)_k \, dt.$$

Lemma

For $v, x, y \in V_k$ assume $(\nabla \hat{J}_k(\lambda), y)_k \leq 0$ and let γ be such that

$$0 \leq \gamma \leq -2\delta(\nabla \hat{J}_k(\lambda), y)_k \left[\int_0^1 (\nabla^2 \hat{J}_k(\lambda + t\gamma y)y, y)_k dt\right]^{-1} \quad \text{for some } \delta \in [0, 1$$

Then $-(1-\delta)\gamma(\nabla \hat{J}_k(\lambda), y)_k \leq \hat{J}_k(\lambda) - \hat{J}_k(\lambda + \gamma y) \leq -\gamma(\nabla \hat{J}_k(\lambda), y)_k$.

We can find $0 < \alpha \leq 2$ in Step 5. of Algorithm MGOPT such that an Armijo-type condition of sufficient decrease is satisfied.

Lemma

For $v, x, y \in V_k$ assume $(\nabla \hat{J}_k(\lambda), y)_k \leq 0$ and let

$$\alpha(\lambda, y) = \min\{2, \frac{-(\nabla \hat{J}_k(\lambda), y)_k}{(\nabla^2 \hat{J}_k(\lambda)y, y)_k + \rho \|y\|_k^3}\}$$

Then

$$\hat{J}_k(\lambda + \alpha(\lambda, y)y) \leq \hat{J}_k(\lambda) + \frac{1}{2}\alpha(\lambda, y)(\nabla \hat{J}_k(\lambda), y)_k.$$

The following lemma states that the coarse-to-fine minimization step with step-length α given by Lemma 2 is a minimizing step (without requiring exact solution of the coarse minimization problem).

Lemma

Take $\lambda_k \in V_k$. Denote with $\tilde{J}_{k-1}(\lambda_{k-1}) = \hat{J}_{k-1}(\lambda_{k-1}) - \phi_{k-1}\,\lambda_{k-1}$ where $\phi_{k-1} = \nabla \hat{J}_{k-1}(I_k^{k-1}\lambda_k) - I_k^{k-1}\nabla \hat{J}_k(\lambda_k)$. Let $\tilde{\lambda}_{k-1} \in V_{k-1}$ be such that $\tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) \leq \tilde{J}_{k-1}(I_k^{k-1}\lambda_k)$ and define $y_k = I_{k-1}^k(\tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)$. Then

$$\hat{J}_k(\lambda_k + \alpha(\lambda_k, y_k) y_k) \leq \hat{J}_k(\lambda_k) + \frac{1}{2}\alpha(\lambda_k, y_k)(\nabla \hat{J}_k(\lambda_k), y_k)_k,$$

where $\alpha(\lambda_k, y_k)$ is defined in Lemma 2 (strict inequality holds if $\tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) < \tilde{J}_{k-1}(I_k^{k-1}\lambda_k)$).

Proof.

The proof follows from Lemma 2 after showing that $(\nabla \hat{J}_k(\lambda_k), y_k)_k \leq 0$.
From the expansion, we obtain

$$(\nabla \tilde{J}_{k-1}(I_k^{k-1}\lambda_k), \tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)_k \leq \tilde{J}_{k-1}(\tilde{\lambda}_{k-1}) - \tilde{J}_{k-1}(I_k^{k-1}\lambda_k) \leq 0.$$

Now we have

$$\begin{aligned}
(\nabla \hat{J}_k(\lambda_k), y_k)_k &= (\nabla \hat{J}_k(\lambda_k), I_{k-1}^k(\tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k))_k \\
&= (I_k^{k-1}\nabla \hat{J}_k(\lambda_k), \tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)_{k-1} \\
&= (\nabla \tilde{J}_{k-1}(I_k^{k-1}\lambda_k), \tilde{\lambda}_{k-1} - I_k^{k-1}\lambda_k)_{k-1} \leq 0.
\end{aligned}$$

For the last equality recall the remark above. □

Theorem

For each k, let λ_k^* be the minimizing solution. Further, let $\hat{J}_k$ be twice Frechét differentiable and let $\nabla^2 \hat{J}_k$ be locally Lipschitz-continuous and satisfies $(\nabla^2 \hat{J}_k(\lambda_k^*)y, y)_k \geq \beta \|y\|_k^2$ together with $\|\nabla^2 \hat{J}_k(\lambda) - \nabla^2 \hat{J}_k(y)\| \leq \rho \,\|\lambda - y\|_k$ uniformly for some positive constants β and ρ in a Neighborhood V_k^ϵ of λ_k^*. Then the MGOPT scheme provides a minimizing step.

Proof.

Let $\lambda_k^{(0)} \in V_k^\epsilon$. Then $A = \{\lambda \in V_k : \hat{J}_k(\lambda) \leq \hat{J}_k(\lambda_k^{(0)})\}$ is a compact set. For $k = 2$, let λ_k the result of the MGOPT step. We have $\tilde{\lambda}_{k-1} = \text{argmin}_{\lambda \in V_{k-1}} \hat{J}_{k-1}(\lambda)$ and from Lemma 3 it follows that

$$\begin{aligned} \hat{J}_k(\lambda_k) &= \hat{J}_k(O_k^{m_2}(\lambda_k^{m_1+1})) \leq \hat{J}_k(\lambda_k^{(m_1)} + \alpha\, I_{k-1}^k(\lambda_{k-1} - I_k^{k-1}\lambda_k^{(m_1)})) \\ &\leq \hat{J}_k(O_k^{m_1}(\lambda_k^0)) \leq \hat{J}_k(\lambda_k^0) \end{aligned}$$

where strict inequality occurs in all steps whenever $\nabla \hat{J}_k$ is non zero. For $k > 2$, due to the induction hypothesis and because of Lemma 3 the theorem holds.

References

- U. Ascher and E. Haber, *A multigrid method for distributed parameter estimation problems*, ETNA, 15(2003), pp. 1–12.
- L. Borcea, *Nonlinear multigrid for imaging electrical conductivity and permittivity at low frequency*, Inverse Problems, 17 (2001), pp. 329–359.
- A. Borzì, *On the convergence of the MG/OPT method*, PAMM, 5(1)(2005), pp. 735–736.
- A. Borzì and U. Hohenester, *Multigrid optimization schemes for solving Bose–Einstein condensates control problems*, SISC
- B. Kaltenbacher, *V-cycle convergence of some multigrid methods for ill-posed problems*, Math. Comp., 72 (2003), pp. 1711–1730.
- R.M. Lewis and S.G. Nash, *Model problems for the multigrid optimization of systems governed by differential equations*, SIAM J. Sci. Comp., 26 (2005), pp. 1811–1837.