Advances in the Accurate and Efficient Solution of Quantum Optimal Control Problems

Alfio Borzì

Dipartimento e Facoltà di Ingegneria, Università degli Studi del Sannio, Italy
Institute for Mathematics and Scientific Computing
Karl-Franzens-University, Graz, Austria

Joint work with

Greg von Winckel (University of Graz), quantum modeling, Newton and quasi-Newton methods, sparse grids, uncertainty.

Stefan Volkwein (University of Konstanz), theory of quantum control problems.

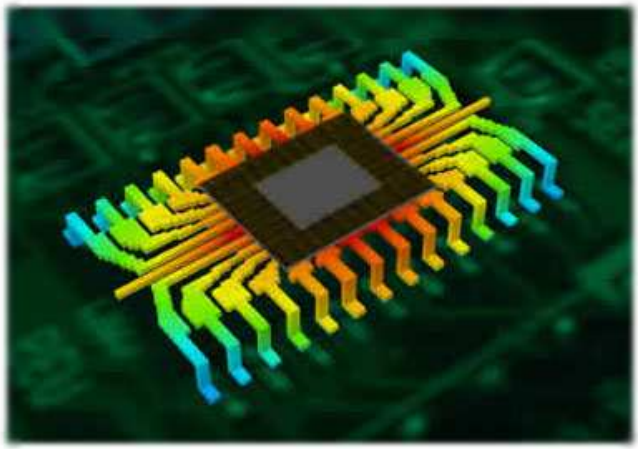
Ulrich Hohenester (University of Graz), theoretical physics, quantum optics, nanophysics.

Julien Salomon (University of Dauphine, Paris), monotonic schemes.

Partially funded by
FWF Austrian Science Fund SFB Project Fast multigrid methods for inverse problems and FWF Project Quantum optimal control of semiconductor nanostructures

Application fields of quantum control methodologies

1. Quantum control: state transitions, laser induced chemistry, magnetic and optical trapping.
2. Quantum computing: qubits, data operations.
3. Quantum transport, superfluids of atoms, vortices.
4. Construction of barriers, channels, etc. for few atoms.
5. Amplification of material waves: laser of atoms.
6. Semiconductor nanostructures.



New challenges from quantum control problems

The possibility to manipulate states of atoms and molecules by means of laser pulses or magnetic fields opens new technological perspectives.

The solution of quantum control problems poses new challenges involving optimal control theory, numerical analysis, and scientific computing.

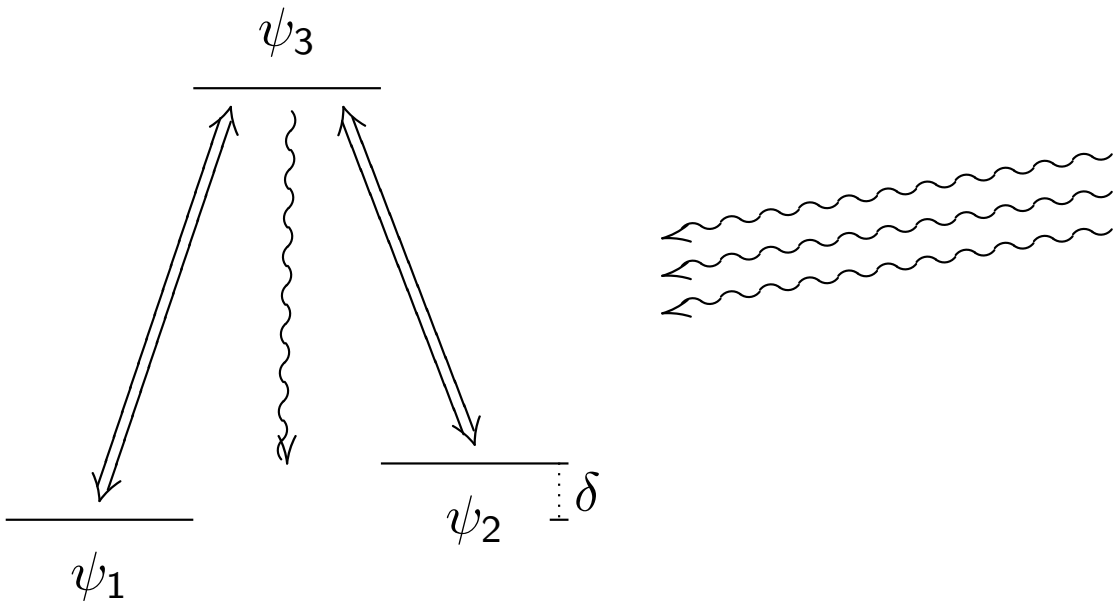
Quantum control models define an important class of bilinear control mechanisms.
They represent a class of nonlinear control strategies with the aim to obtain better system response than possible with linear control.

Optimal control of finite-level quantum systems

Optimal control of finite-level quantum systems

Quantum systems with a finite number of states model artificial atoms (semiconductor quantum dots) and quantum devices (quantum gates)

Consider a three-level system with two stable states ψ_1 and ψ_2 (conservative), and one unstable state ψ_3 (dissipative).



Today's trend is to use optimal control theory with sophisticated experimental technology and high-fidelity laser systems.

Finite-level quantum models

Governed by Schrödinger-type equations for a n-component wave function $\psi : [0, T] \to \mathbb{C}^n$ as follows

$$i\dot{\psi}(t) = H(\epsilon(t))\,\psi(t), \qquad \psi(0) = \psi_0,$$

for $t \in (0, T]$ and $T > 0$ is a given terminal time.

The funcion $\epsilon(t)$ represents the external control field.

The linear Hamiltonian $H = H_0 + H_1(\epsilon)$, consists of

A free Hamiltonian $H_0 \in \mathbb{C}^{n\times n}$ describing the unperturbed (uncontrolled) system;

A control Hamiltonian $H_1 : \mathbb{C} \to \mathbb{C}^{n\times n}$ modeling the coupling of the quantum state to the control field ϵ.

The objective of the quantum control

Very often, control is applied to reach a target state at $t = T$.
Moreover, one needs to avoid population of dissipative states during the control process, and we have physically motivated constraints like limited laser resources.
These modeling requirements may result in the following

$$\begin{aligned} J(\psi,\epsilon) &= \frac{1}{2}\left|\psi(T)-\psi_d\right|^2_{\mathbb{C}^n} + \frac{\gamma}{2}\left\|\epsilon\right\|^2_{L^2(0,T;\mathbb{C})} + \frac{\mu}{2}\left\|\dot{\epsilon}\right\|^2_{L^2(0,T;\mathbb{C})} \\ &\quad + \frac{1}{2}\sum_{j\in J}\alpha_j\left\|\psi_j\right\|^2_{L^2(0,T;\mathbb{C})} \end{aligned}$$

where ψ_d is the desired terminal state; $\gamma > 0$ and μ, $\alpha_i \geq 0$ are weighting factors; ψ_j denotes the j-th (dissipative) component of ψ.

First-order necessary optimality conditions

For the quantum optimal control problem

$$\min J(\psi,\epsilon),\ \text{subject to } i\dot{\psi}(t) = H(\epsilon(t))\,\psi(t), \quad \psi(0) = \psi_0$$

Theorem

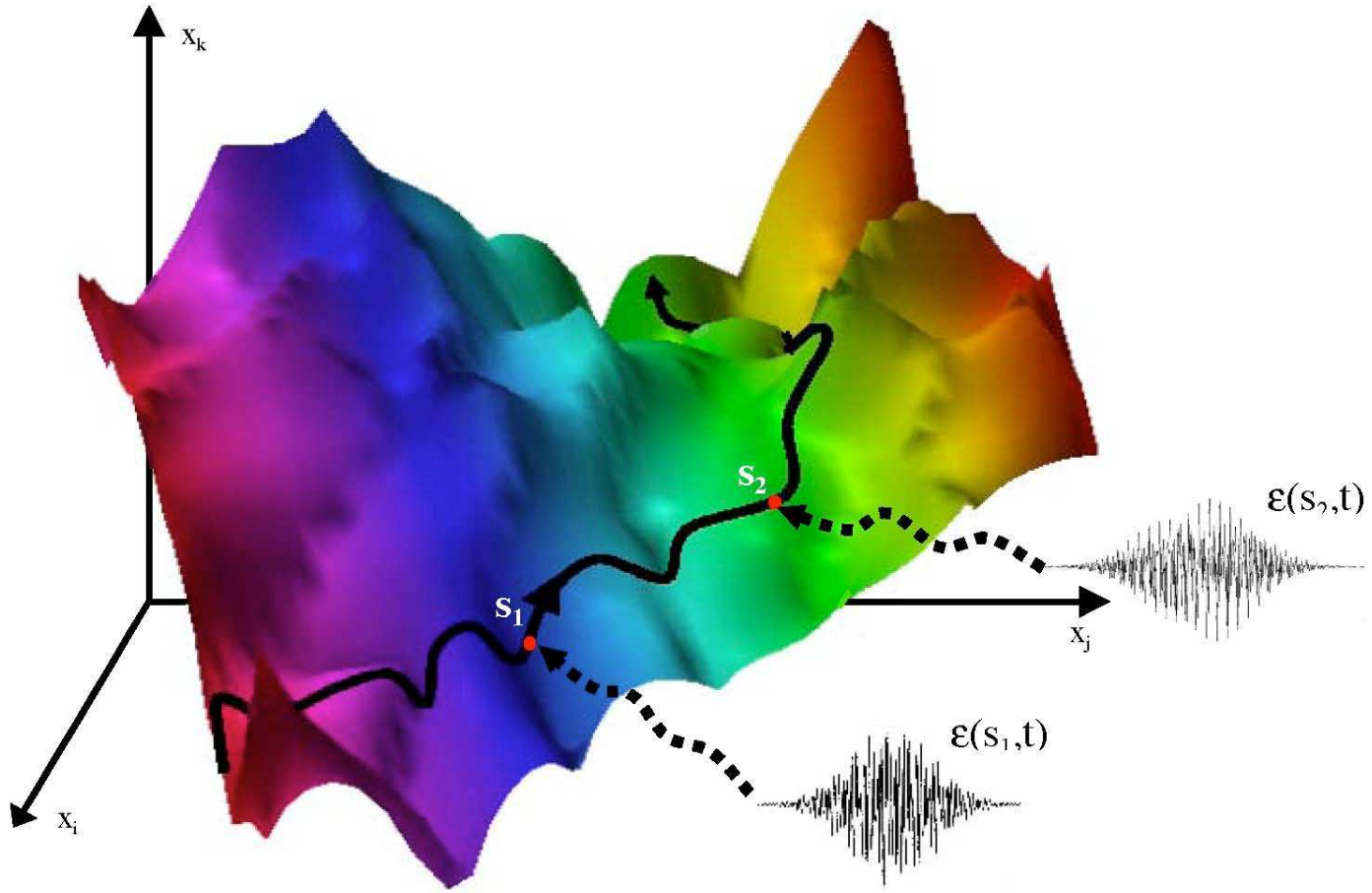
Suppose that $x = (\psi,\epsilon) \in X$ is a local solution to the optimal control problem. Then there exist (unique) Lagrange multipliers $p \in H^1(0,T;\mathbb{C}^n)$ $(\mu > 0)$ satisfying

$$\begin{aligned} i\dot{\psi} &= H(\epsilon(\cdot))\psi \\ i\dot{p} &= H(\epsilon(\cdot))^* p - \alpha_j(\psi)_j \\ -\mu\ddot{\epsilon} + \gamma\epsilon &= \Re e(p\cdot(H'_{1r}(\epsilon)\psi)^*) + i\,\Re e(p\cdot(H'_{1i}(\epsilon)\psi)^*) \end{aligned}$$

where

$$\psi(0) = \psi_0, \qquad ip(T) = \psi(T) - \psi_d, \qquad \epsilon(T) = \epsilon(0) = 0.$$

Multiple minima



Courtesy H. Rabitz, Univ. Princeton

Second-order optimality conditions

Consider the following optimal control problem

$$\begin{cases} \min_\epsilon J(\psi,\epsilon) & := & \frac{1}{2}|\psi(T)-\psi_d|^2+\frac{\gamma}{2}\|\dot{\lambda}\|^2 \\ c(\psi,\epsilon) & := & i\dot{\psi}-a\,\psi-\epsilon\,\psi=0 \end{cases}$$

We obtain

$$(\nabla^2\hat{J}\,\delta\epsilon,\delta\epsilon)=(W\delta\epsilon)(W\delta\epsilon)^*+2\Re e(p\,\delta\epsilon,W\delta\epsilon)+\gamma(\dot{\delta\epsilon},\dot{\delta\epsilon}).$$

where $W=W(\psi(\epsilon),\epsilon)=c_\psi(\psi(\epsilon),\epsilon)^{-1}\,c_\epsilon(\psi(\epsilon),\epsilon)$.
Since $p(t)$ is unitary, we have $|p(t)|=|p(T)|=|\psi(T)-\psi_d|$.

Using this result we obtain that $|\Re e(p\,\delta\epsilon,W\delta\epsilon)|\le C\,|\psi(T)-\psi_d|\,\|\delta\epsilon\|^2$, for some $C>0$ depending on λ.
For sufficiently small values of $|\psi(T)-\psi_d|$ positiveness of the reduced Hessian is obtained.

A Λ-type three-level model

Free Hamiltonian

$$H_0 = \frac{1}{2}\begin{pmatrix} -\delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & -i\Gamma_o \end{pmatrix},$$

where the term $-i\Gamma_o$ accounts for environment losses (spontaneous photon emissions, scattering of gamma rays from crystals).
The coupling to the external field reads

$$H_1(\epsilon) = -\frac{1}{2}\begin{pmatrix} 0 & 0 & \mu_1\epsilon \\ 0 & 0 & \mu_2\epsilon \\ \mu_1\epsilon^* & \mu_2\epsilon^* & 0 \end{pmatrix},$$

where μ_1 and μ_2 describe the coupling strengths of states ψ_1 and ψ_2 to the inter-connecting state ψ_3 (e.g., optical dipole matrix elements).
Initial and final states are given by

$$\psi_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ and } \psi_d = \begin{pmatrix} 0 \\ e^{-i\delta T} \\ 0 \end{pmatrix}$$

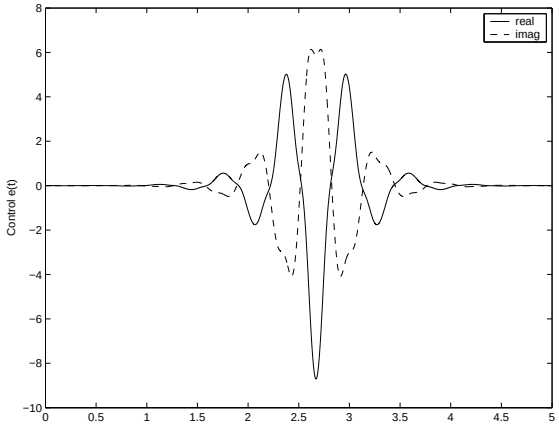
Importance of optimization parameters

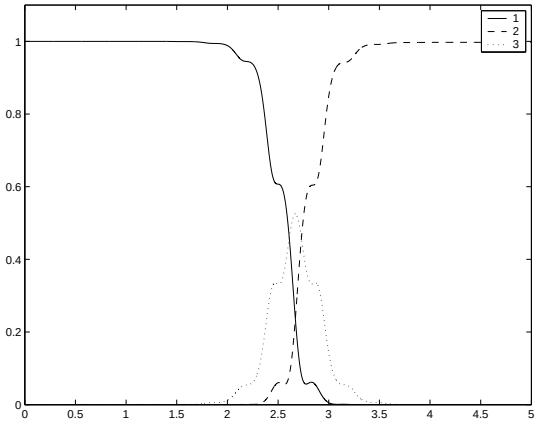
Smaller values of γ result in smaller $|\psi(T) - \psi_d|_{\mathbb{C}^3}$.
As μ increases, $|\psi(T) - \psi_d|_{\mathbb{C}^3}$ increases: additional smoothness of the control function (slightly) reduces the capability of tracking.
Larger μ makes the problem behaving better, resulting in a smaller number of iterations.
By taking $\alpha = \alpha_3 > 0$ dissipation is reduced and therefore better tracking is possibly achieved.

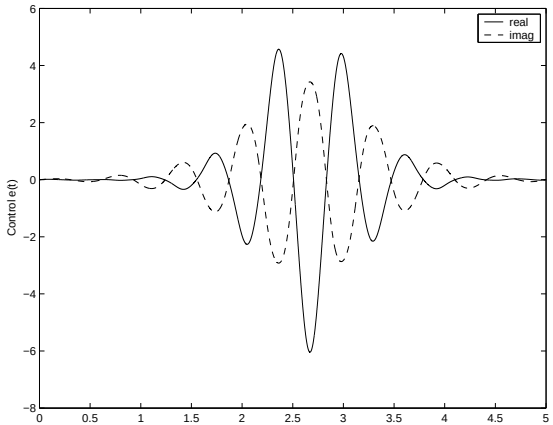
γ	μ	α	$\lvert\psi(T) - \psi_d\rvert_{\mathbb{C}^3}$	J	CPU
10^{-7}	10^{-7}	0.05	$8.6 \cdot 10^{-4}$	$2.37 \cdot 10^{-3}$	19.6
10^{-7}	10^{-9}	0.05	$3.7 \cdot 10^{-4}$	$5.46 \cdot 10^{-4}$	55.6
10^{-7}	0	0.05	$6.9 \cdot 10^{-5}$	$1.41 \cdot 10^{-4}$	424.8
10^{-7}	0	0	$1.2 \cdot 10^{-3}$	$2.33 \cdot 10^{-6}$	763.1
10^{-4}	10^{-4}	0.05	$3.3 \cdot 10^{-2}$	$6.52 \cdot 10^{-2}$	47.3
10^{-4}	10^{-6}	0.05	$4.4 \cdot 10^{-3}$	$9.03 \cdot 10^{-3}$	42.3
10^{-4}	0	0.05	$2.7 \cdot 10^{-3}$	$5.68 \cdot 10^{-3}$	17.2
10^{-4}	0	0	$8.3 \cdot 10^{-3}$	$3.34 \cdot 10^{-4}$	5.5

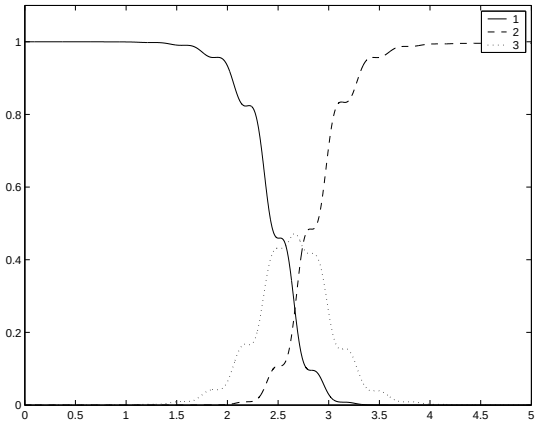
Optimal solutions

for $\gamma = 10^{-4}$, $\alpha_3 = 0.01$ and $\mu = 0$ (top); for $\gamma = 10^{-4}$, $\alpha_3 = 0.01$ and $\mu = 10^{-6}$.









Monotonic schemes for quantum control problems

Monotonic schemes have been initially introduced by Krotov. They are a special case of gradient based methods.
Monotonic schemes have been further developed by Tannor, Zhu & Rabitz, Maday, Turinici & Salomon.

Consider two fields ϵ and ϵ', and the corresponding ψ and ψ', and p and p'. We have

$$\begin{aligned}\widetilde{J}(\epsilon') - \widetilde{J}(\epsilon) &= \int_0^T \Re e(p(s)^* \cdot \big(H_1(\epsilon'(s)) - H_1(\epsilon(s))\big)\psi'(s)) \\ &\quad -\frac{\gamma}{2}\left(|\epsilon'(s)|^2 - |\epsilon(s)|^2\right) \mathrm{d}s + \frac{1}{2}\left\langle \psi' - \psi, \Lambda(\psi' - \psi)\right\rangle_{L^2(0,T;\mathbb{C}^n)}.\end{aligned}$$

This formula is the starting point for the design of monotonic schemes.

Crank-Nicholson monotonic schemes

CNMS schemes represent a recent advance in monotonic schemes, supported by theoretical results on convergence properties.

Algorithm (CNMS)

Given an initial control amplitude ϵ^0 and its associated state ψ^0 and Lagrange multiplier p^0, suppose that ψ^k, p^k, ϵ^k, have already been computed. The derivation of ψ^{k+1}, p^{k+1}, ϵ^{k+1}, is done as follows.
Forward propagation: Given $\psi_0^{k+1} = \psi_0$, compute $\psi_{\ell+1}^{k+1}$ from ψ_ℓ^{k+1} by
Step 1. (Newton iteration) Compute ϵ_ℓ^{k+1} by

$$\begin{pmatrix} \epsilon_{\Re e,\ell}^{k+1} \\ \epsilon_{\Im m,\ell}^{k+1} \end{pmatrix} = \begin{pmatrix} \epsilon_{\Re e,\ell}^{k} \\ \epsilon_{\Im m,\ell}^{k} \end{pmatrix} + \frac{1}{2}\left(-\frac{\delta t}{4}\Im m B_\ell^k + \frac{\gamma}{2} I_2\right)^{-1}\left(\frac{1}{2}\Re e A_\ell^k - \gamma \begin{pmatrix} \epsilon_{\Re e,\ell}^{k} \\ \epsilon_{\Im m,\ell}^{k} \end{pmatrix}\right).$$

Step 2. Compute $\psi_{\ell+1}^{k+1}$ by

$$\psi_{\ell+1} = \left(I_d - \frac{\delta t H_\ell}{2i}\right)^{-1}\left(I_d + \frac{\delta t H_\ell}{2i}\right)\psi_\ell \qquad (1)$$

Backward propagation: Given $p_N^{k+1} = i\psi_d$, compute p_ℓ^{k+1} from $p_{\ell+1}^{k+1}$.

Nonlinear conjugate gradient schemes

Gradient schemes are the most commonly used methods in computational physics. A first advance in gradient-based schemes has been the extension of NCG methods to problems defined on complex Hilbert spaces.
A sufficient descent condition

$$\Re e \langle g_k, d_k \rangle \leq -c \|g_k\|^2$$

Step 1. Given $k = 1$, ϵ_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 $\epsilon_{k+1} = \epsilon_k + \tau_k d_k$.

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

Step 5. Compute β_k using Dai-Yuan or Hager-Zhang schemes.

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.
ϵ_{k_0} given starting approximation on the coarsest grid.
I_k^{k+1} interpolation operator from k to $k+1$.
$NCG_k(\epsilon_k)$ the basic iteration; $*$ denotes the resulting solution.

Algorithm (Cascadic NCG method)

Step 1. *Given* $k = k_0$, $\epsilon^*_{k_0}$.

Step 2. *Compute* $\epsilon_k = NCG_k(\epsilon^*_k)$.

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

Step 4. *Else if* $k < k_f$ *then interpolate* $\epsilon^*_{k+1} = I_k^{k+1}\epsilon_k$.

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

Performance of NCG,Cascadic-NCG, and CNMS schemes

The NCG scheme provides better performance while refining the computational mesh. There is a lack of robustness of the CNMS scheme forsufficiently small γ.

	$N = 2048$		$N = 4096$	
tol_{abs}	CPU(NCG)	CPU(CNMS)	CPU(NCG)	CPU(CNMS)
10^{-4}	1.17	1.28	2.32	1.39
10^{-5}	4.32	12.63	9.26	15.92
10^{-6}	5.01	48.00	17.21	*no conv*

We see a dramatic improvement with the Cascadic-NCG version

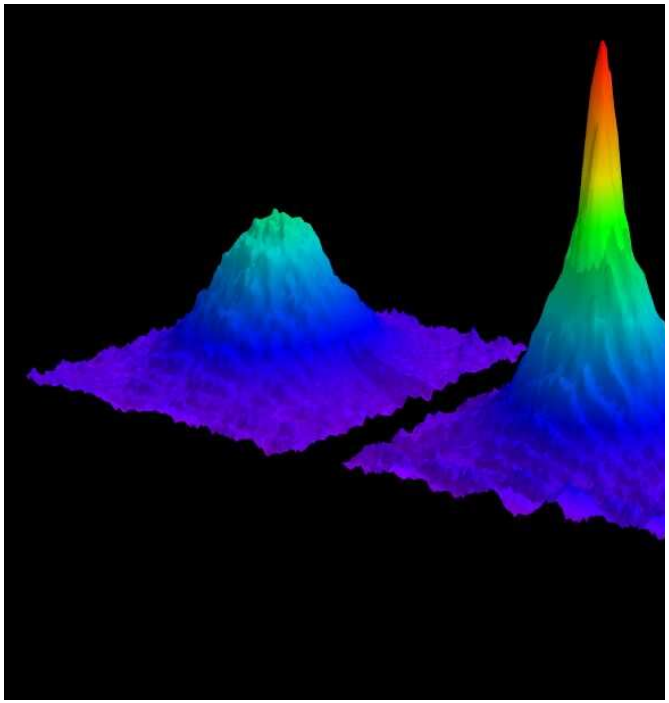
	$\gamma = 10^{-4}$		$\gamma = 10^{-6}$	
N	CPU(NCG)	CPU(C-NCG)	CPU(NCG)	CPU(C-NCG)
4096	40.54	6.26	254.70	58.10
8192	112.57	12.71	319.46	134.00
16384	312.17	27.42	626.84	279.46

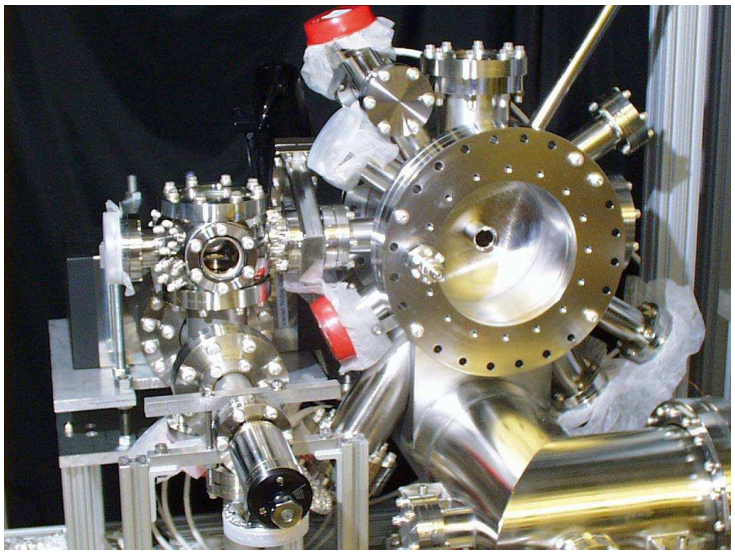
Computational effort to solve for $tol_{abs} = 10^{-6}$; $\gamma_0 = 0.01$, $\alpha_3 = 0.05$; in C-NCG coarsest level $N = 1024$.

Optimal control of extended quantum systems: transport and trapping

Bose Einstein condensates

Consider a bosonic gas (e.g. Rubidium) trapped in a magnetic field. By lowering the confining potential, atoms with higher energy escape (evaporation) while the remaining atoms condensate to a lower temperature (ca. 10 sec, 50 micron).





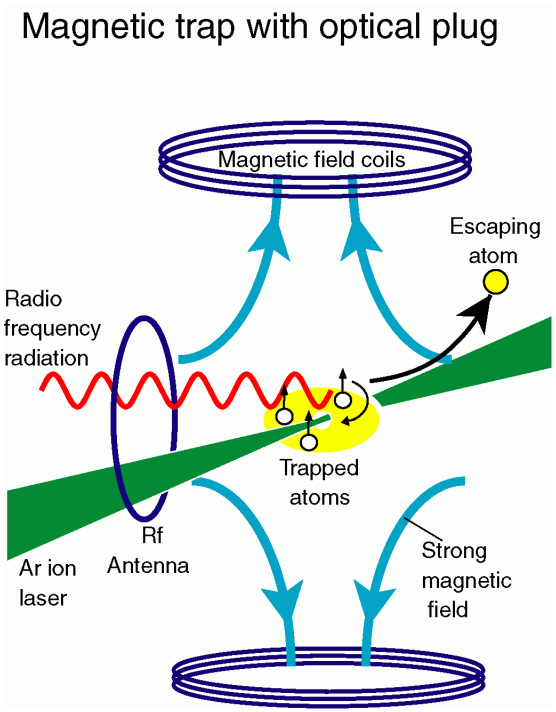
Courtesy W. Ketterle, MIT

Bose Einstein condensates model

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

$$i\,\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \left(-\frac{1}{2}\nabla^2 + V(\mathbf{r},\lambda(t)) + g\,|\psi(\mathbf{r},t)|^2\right)\psi(\mathbf{r},t)$$

We consider $V(\mathbf{r},\lambda(t))$ is a three-dimensional potential produced by a magnetic microtrap. The control function $\lambda(t)$ describes the variation of the confining potential with time.



Total control of matter waves

Trapping and coherent manipulation of cold neutral atoms in microtraps near surfaces of atomic chips is the focus of the present research towards total control of matter waves at small scales.

This achievement has boosted developments in the atomic interferometry, the construction of quantum gates, the microscopic magnetic field imaging, quantum data encoding, etc..

At the base of all these developments is the ability to manipulate the Bose-Einstein condensates (BEC) subject to a control potential $V(x, \lambda(t))$ where $\lambda(t)$ parameterizes the trapping field.

We consider the problem to split and transport a BEC being confined in a single well $V(x, 0)$ at $t = 0$ to a double well $V(x, 1)$ at time T. We have

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

where $c = 40$ and d is the width of the double well potential.

Optimal control formulation and optimality system

Consider a BEC at the initial state ψ_0 and a target state ψ_d.

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

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

$$i\,\frac{\partial}{\partial t}\psi=\left(-\frac{1}{2}\nabla^2+V_\lambda+g|\psi|^2\right)\psi$$

$$i\,\frac{\partial}{\partial t}p=\left(-\frac{1}{2}\nabla^2+V_\lambda+2g|\psi|^2\right)p+g\,\psi^2\,p^*$$

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

where $\langle u,v\rangle=\int_\Omega u(x)^*v(x)\,dx$. We have the initial and terminal conditions

$$\psi(0)=\psi_0 \text{ and } ip(T)=-\langle\psi_d|\psi(T)\rangle\,\psi_d$$
$$\lambda(0)=0\,,\quad \lambda(T)=1\,.$$

Discretization schemes for quantum control problems

The discretization of quantum state and adjoint equations is a delicate issue. For BEC we use unconditionally stable second-order norm-preserving time-splitting spectral schemes.

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

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 $A = V_\lambda + 2g|\psi|^2$ 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

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

The gradient

For a given potential $V(x, \lambda(t))$, we have a unique $\psi(\lambda) = \psi(x, t)$. In terms of λ we have the reduced objective $\hat{J}(\lambda) = J(\psi(\lambda), \lambda)$.
The Taylor series of $\hat{J}(\lambda)$ in a Hilbert space X is

$$\hat{J}(\lambda + \epsilon\varphi) = \hat{J}(\lambda) + \epsilon \left(\nabla \hat{J}(\lambda), \varphi\right)_X + \frac{\epsilon^2}{2} \left([\nabla^2 \hat{J}(\lambda)]\varphi, \varphi\right)_X + O(\epsilon^3)$$

For $X = L^2(0, T; \mathbb{R})$, the reduced gradient is given

$$\nabla \hat{J}_{L^2}(\lambda) = -\gamma\, \ddot{\lambda} - \Re e\langle\psi| \frac{\partial V_\lambda}{\partial \lambda} |p\rangle,$$

In the case $X = H^1(0, T; \mathbb{R})$ formulation, we have that

$$-\frac{d^2}{dt^2}[\nabla \hat{J}_{H^1}(\lambda)] = -\gamma \ddot{\lambda} - \Re e\langle\psi, \frac{\partial V_\lambda}{\partial \lambda} p\rangle,$$

with $[\nabla \hat{J}(\lambda)](0) = 0$ and $[\nabla \hat{J}(\lambda)](T) = 0$.
The H^1 gradient $\nabla \hat{J}_{H^1}(\lambda)$ has the same regularity as λ, while the L^2 gradient does not.

Optimal controls obtained on differente X spaces

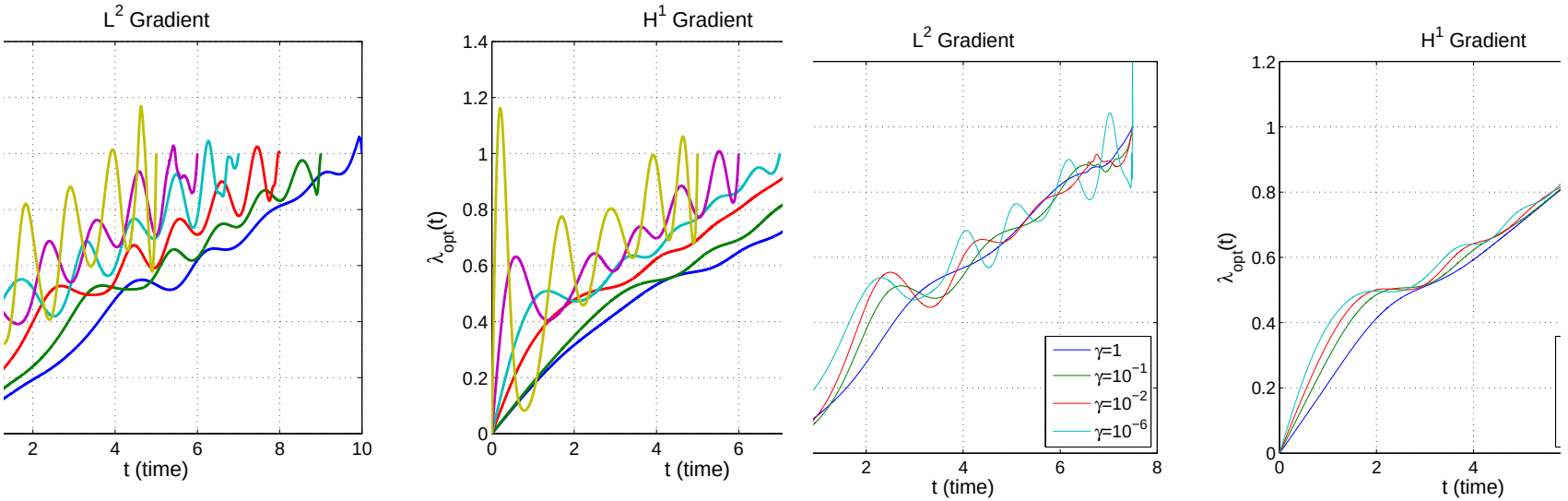


Figure: The optimal control function in the L^2 and H^1 settings for decreasing time horizons. As T becomes smaller, λ becomes a more oscillating function.
Dependence of the optimal control function on the regularization parameter γ for the L^2 and H^1 spaces. More oscillating controls are obtained with smaller γ. $M = 3200$ time steps with $g = 10$ and $T = 7.5$.

Hager-Zhang Nonlinear conjugate gradient on X space

Step 1. Given $k = 1$, λ_1, $d_1 = -g_1$, if $\|g_1\|_X < tol$ then stop.

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

$$\begin{aligned} \hat{J}(\lambda_k + \tau_k d_k) &\leq \hat{J}(\lambda_k) + \delta\, \tau_k\, (g_k, d_k)_X \\ (g(\lambda_k + \tau_k d_k), d_k)_X &> \sigma\, (g_k, d_k)_X, \qquad 0 < \delta < \sigma < 1/2 \end{aligned}$$

Step 3. Let $\lambda_{k+1} = \lambda_k + \tau_k\, d_k$.

Step 4. Compute $g_{k+1} = \nabla \hat{J}_X(\lambda_{k+1})$.

If $\|g_{k+1}\|_X < tol_{abs}$ or $\|g_{k+1}\|_X < tol_{rel}\, \|g_1\|_X$ or $k = k_{\max}$ then stop.

Step 5. Compute β_k by

$$\beta_k = \frac{(\sigma_k, g_{k+1})_X}{(d_k, y_k)_X}, \qquad \sigma_k = y_k - 2 d_k \frac{(y_k, y_k)_X}{(y_k, d_k)_X}, \quad y_k = g_{k+1} - g_k$$

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

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

BFGS on X space

With BFGS the search direction is given by $p_k = -H_k \nabla \hat{J}(\lambda_k)$.
By the Sherman-Morrison-Woodbury formula, we establish a recurrence for H.

$$H_{k+1} = H_k + \frac{s_k^\top y_k + y_k^\top H_k y_k}{(s_k^\top y_k)^2}(s_k s_k^\top) - \frac{H_k y_k s_k^\top + s_k y_k^\top H_k}{s_k^\top y_k}$$

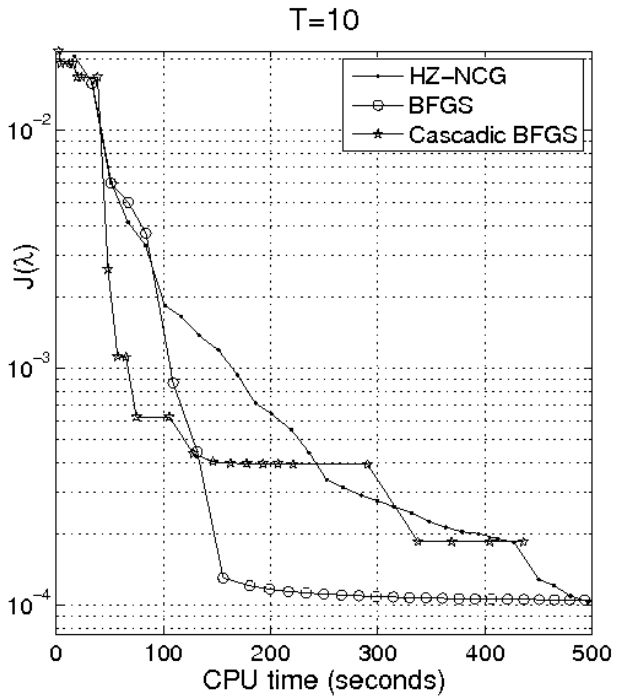
where $s_k = \tau_k p_k$. Supposing X is either $L^2(0, T; \mathbb{R})$ or $H^1(0, T; \mathbb{R})$, the function space analog of the outer product is a dyadic operator $x \otimes y : X \to X$. The action of this operator on a third element $v \in X$ can be expressed in terms of the inner product $(x \otimes y)\, v = (y, z)_X\, v$.
From the recursion relation for H, we obtain

$$p_k = -H_0 g_k - \sum_{j=0}^{k-1} c_j [d_j (s_j, g_k)_X\, s_j - (z_j, g_k)_X\, s_j - (s_j, g_k)_X\, z_j]$$

where $c_j = (s_j, y_j)_X^{-1}$, $d_j = 1 + c_j (y_j, z_j)$, and we also have a similar sum formula for the term $z_k = H_k y_k$.

$$z_k = H_0 y_k + \sum_{j=0}^{k-1} c_j \left\{ [d_j (s_j, y_k)_X - (z_j, y_k)_X]\, s_j - (s_j, y_k)_X\, z_j \right\}.$$

Results with HZ-NCG and BFGS on H^1 space



mesh	$\hat{J}_{min}$	$\|\nabla \hat{J}_{min}\|$	iterations	CPU time (sec)
400	1.6605×10^{-2}	1.4288×10^{-1}	15	3.8407×10^{1}
800	5.5963×10^{-4}	4.5284×10^{-2}	62	2.8107×10^{2}
1600	2.9634×10^{-4}	1.0733×10^{-2}	30	3.6334×10^{2}
3200	1.0562×10^{-4}	3.6378×10^{-3}	37	9.6153×10^{2}

MultiGrid OPTimization framework

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

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

where $\lambda_k \in X_k$ and $\hat{J}_k(\cdot)$ is the objective.

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

We also choose an optimization scheme as 'smoother'

$$\lambda_k^\ell = O_k\left(\lambda_k^{\ell-1}\right)$$

That provides sufficient reduction

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

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

MGOPT Algorithm

Initialize λ_k^0. If $k = 1$, solve $\min_{\lambda_k} \hat{J}_k(\lambda_k) - (f_k, \lambda_k)_k$ and return. Else if $k > 1$,

1. *Pre-optimization*: $\lambda_k^\ell = O_k(\lambda_k^{\ell-1}, f_k)$, $\ell = 1, 2, \ldots, \gamma_1$
2. *Coarse grid problem*
   Restrict the solution: $\lambda_{k-1}^{\gamma_1} = I_k^{k-1} \lambda_k^{\gamma_1}$
   Fine-to-coarse correction: $\tau_{k-1} = \nabla \hat{J}_{k-1}(\lambda_{k-1}^{\gamma_1}) - I_k^{k-1} \nabla \hat{J}_k(\lambda_k^{\gamma_1})$
   $f_{k-1} = I_k^{k-1} f_k + \tau_{k-1}$
   Apply MGOPT to the coarse grid problem:

$$\min_{\lambda_{k-1}} \hat{J}_{k-1}(\lambda_{k-1}) - (f_{k-1}, \lambda_{k-1})_{k-1}$$

3. *Coarse grid correction*
   Prolongate the error: $d = I_{k-1}^k(\lambda_{k-1} - \lambda_{k-1}^{\gamma_1})$
   Perform a line search in the direction d to obtain a step length α_k.
   Coarse grid correction: $\lambda_k^{\gamma_1+1} = \lambda_k^{\gamma_1} + \alpha_k d$
4. *Post-optimization*: $\lambda_k^\ell = O_k(\lambda_k^{\ell-1}, f_k)$, $\ell = \gamma_1 + 2, \ldots, \gamma_1 + \gamma_2 + 1$

Computational performance of CNCG and MGOPT

		CNCG		MGOPT
γ	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU	$\frac{1}{2}\left(1-\left\|\langle\psi_d,\psi(T)\rangle\right\|^2\right)$	CPU
10^{-2}	$2.23\cdot10^{-2}$	17	$9.69\cdot10^{-4}$	116
10^{-4}	$4.54\cdot10^{-4}$	202	$6.01\cdot10^{-4}$	82
10^{-6}	$1.38\cdot10^{-2}$	14	$8.78\cdot10^{-4}$	78

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

	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	$3.89\cdot10^{-4}$	53	$7.08\cdot10^{-4}$	149
50	$2.35\cdot10^{-3}$	80	$9.84\cdot10^{-3}$	76
75	$5.54\cdot10^{-3}$	90	$1.85\cdot10^{-3}$	163
100	$4.94\cdot10^{-1}$	50	$5.44\cdot10^{-3}$	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 λ control

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

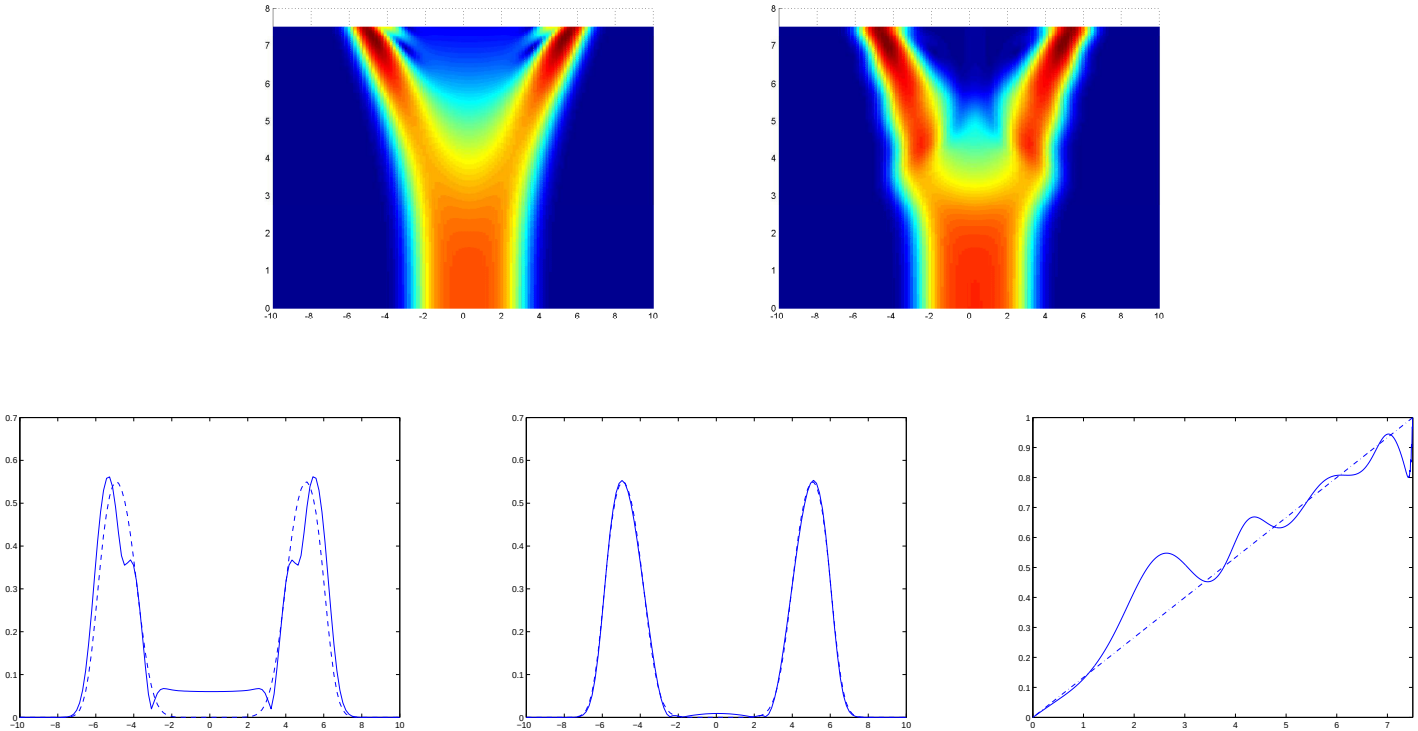


Figure: The function $|\psi(x,t)|$ on the space-time domain (top) for the linear (left) and optimized (right) control. The corresponding profiles at $t = T$ (bottom, continuous line) compared to the desired state (dashed line). The tracking error $\frac{1}{2}\left(1 - \left|\langle \psi_d, \psi(T)\rangle\right|^2\right)$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

Dipole quantum control uncertainty on the potential

Electronic states of a charged particle in a well potential

The control of quantum electronic states has a host of applications such as control of photochemical processes and semiconductor lasers.

Consider a confining potential $V_0(x)$ with a 'well' envelope. The eigenproblem

$$\left\{-\partial_x^2 + V_0(x) - \lambda_j\right\}\phi_j(x) = 0, \quad j = 1, 2, \ldots,$$

defines eigenfunctions representing the eigenstates with energy λ_j.

A representative potential with applications in semiconductor nanostructures is the infinite barrier well potential where $V_0(x) = 0$ for $x \in (0, \ell)$ and $V_0(0) = +\infty$ and $V_0(\ell) = +\infty$.
The infinite barrier condition is equivalent to homogeneous Dirichlet boundary conditions for the wavefunction and thus we have

$$\lambda_j = \frac{j^2\pi^2}{\ell^2} \text{ and } \phi_j(x) = \sin(j\pi x/\ell).$$

Electric dipole control

Consider a control field modeling a laser pulse. Using the dipole approximation results in the following

$$V(x,t) = V_0(x) + u(t)\,x$$

where $u : (0,T) \to \mathbb{R}$ is the modulating control amplitude.
The quantum state of a charged particle subject to this potential is governed by the time-dependent Schrödinger equation ($c(\psi,u)=0$)

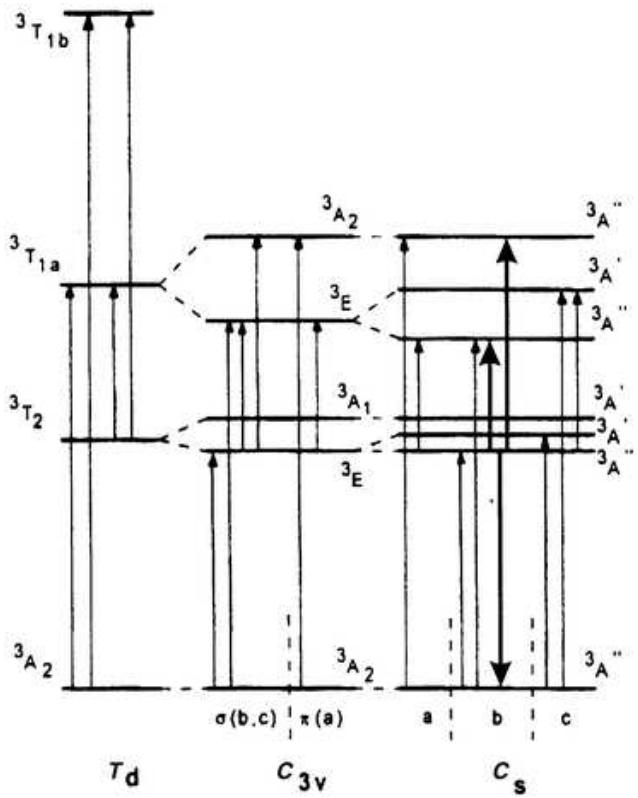
$$i\frac{\partial}{\partial t}\psi(x,t) = \left\{-\frac{\partial}{\partial x^2} + V(x,t)\right\}\psi(x,t), \qquad (x,t) \in Q = \Omega \times (0,T),$$

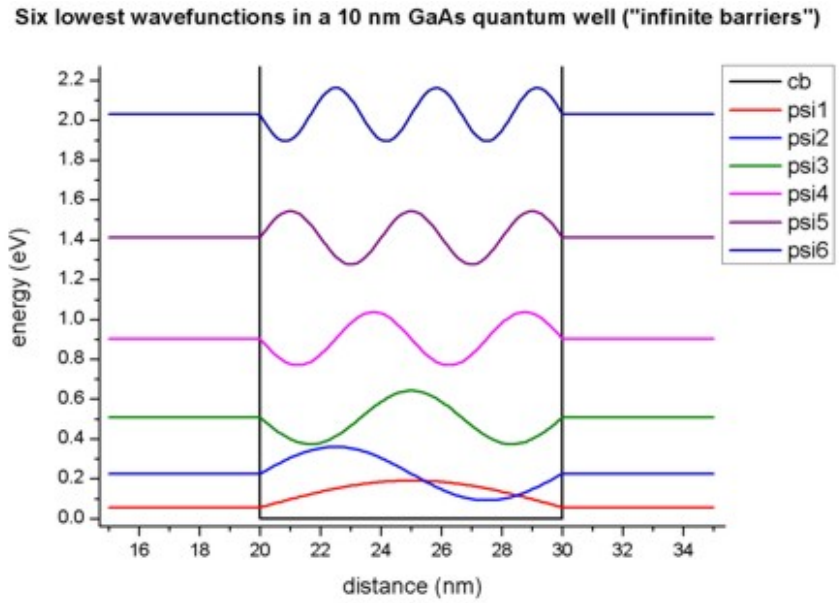
Objective of the control

$$J(\psi,u) := \frac{1}{2}\left(1 - \|P\psi(\cdot,T)\|^2_{\mathcal{H}}\right) + \frac{\gamma}{2}\|u\|^2_{\mathcal{U}}$$

where the projector $P\psi = (\psi_d,\psi)_{\mathcal{H}}\,\psi_d$ and $\|u\|^2_{\mathcal{U}} = \|u\|^2 + \alpha\,\|\dot{u}\|^2$

Electric dipole transitions and a GaAs quantum well





Transitions $\phi_j \longrightarrow \phi_k$

Dipole control optimality system

Introduce the Lagrangian

$$L(\psi, u, p) = J(\psi, u) + \Re e \int_0^T \int_\Omega p^*(x,t) c(\psi, u)(x,t)\, dxdt$$

where p is the Lagrange multiplier. The following first-order optimality system characterizes the optimal solution

$$\left\{ i\partial_t + \partial_x^2 - V_0(x) - u(t)\, x \right\} \psi(x,t) = 0$$

$$\left\{ i\partial_t + \partial_x^2 - V_0(x) - u(t)\, x \right\} p(x,t) = 0$$

$$-\gamma\, u + \gamma\alpha\, \ddot{u} + \Re e \int_\Omega p^*(x,t)\, x\, \psi(x,t)\, dx = 0$$

with homogeneous Dirichlet boundary conditions, and initial and terminal conditions given by

$$\psi(x,0) = \psi_0(x)$$

$$p(x,T) = i\, (\psi_d(\cdot), \psi(\cdot,T))_{\mathcal{H}}\, \psi_d(x)$$

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

Formulation and solution of the discrete control problem

We develop a Newton scheme

with the following requirements

1. appropriate discretization scheme that is norm-preserving and second-order accurate also in the case of time-varying potential;
2. discretize-before-optimize approach to avoid any inconsistency between the optimality condition and its discrete approximation;
3. appropriate construction of the Hessian in complex Hilbert spaces;
4. robust linesearch based on bisection – bilinear control problems have very flat minima;

Modified Crank-Nicholson scheme

Our MCN scheme results in the following

$$\psi_k - \psi_{k-1} = -\frac{i\delta t}{4}[H(t_k) + H(t_{k-1})][\psi_k + \psi_{k-1}].$$

Spatial discretization $\mathbf{H}_k$ of the Hamiltonian $H(t_k)$ is by linear FEM. We have that $\mathbf{H}_k = \mathbf{H}_k^\top$, which is important for preserving unitarity of the time-stepping method. Let $\mathbf{A}_k = \frac{\delta t}{4}[\mathbf{H}_k + \mathbf{H}_{k-1}]$.

$$\mathbf{B}_k = \begin{pmatrix} \mathbf{I} & \mathbf{A}_k \\ -\mathbf{A}_k & \mathbf{I} \end{pmatrix}.$$

This gives the following representation of the equality constraint

$$\mathbf{c}_k(\mathbf{y}, \mathbf{u}) = \mathbf{B}_k \mathbf{y}_k - \mathbf{B}_k^\top \mathbf{y}_{k-1}, \quad \mathbf{y}_k = \begin{pmatrix} \Re e[\psi_k] \\ \Im m[\psi_k] \end{pmatrix},$$

where $\mathbf{y}$ is a compact notation for the set of state vectors at each time step $\mathbf{y}_1, \ldots, \mathbf{y}_{N_t}$ and similarly for $\mathbf{u}$.

Discrete optimality system

Let $\mathbf{S}$ corresponds to multiplication by i. We have that

$$\mathbf{S} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}, \quad (\psi_d, \psi)_{\mathcal{H}} \text{ corresponds to } \begin{pmatrix} \mathbf{y_d}^\top \\ \mathbf{y_d}^\top \mathbf{S} \end{pmatrix} \mathbf{y}$$

In this representation, we can rewrite the objective in the form

$$J(\mathbf{y}, \mathbf{u}) = \frac{1}{2}\left[1 - \mathbf{y}_{N_t}^\top \begin{pmatrix} \mathbf{y_d} & -\mathbf{S}\mathbf{y_d} \end{pmatrix} \begin{pmatrix} \mathbf{y_d}^\top \\ \mathbf{y_d}^\top \mathbf{S} \end{pmatrix} \mathbf{y}_{N_t}\right] + \frac{\gamma}{2}\mathbf{u}^\top \mathbf{K}\mathbf{u}$$

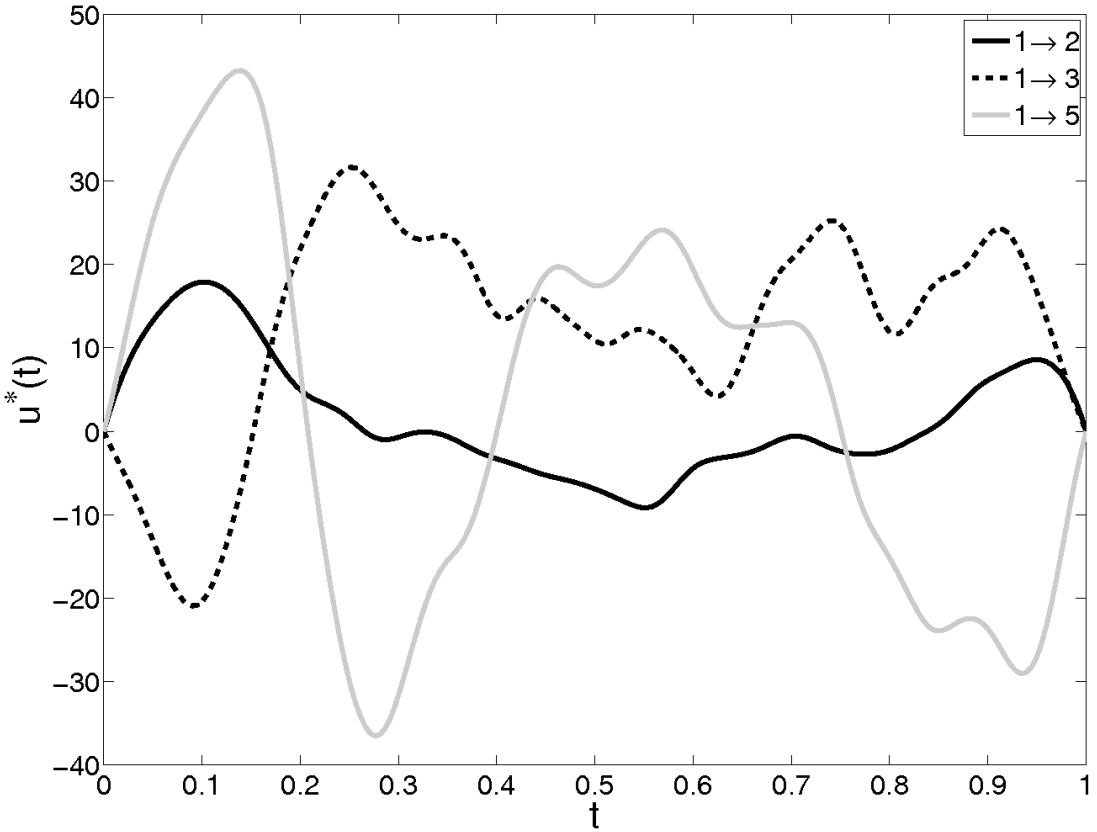
The matrix $\mathbf{K}$ is the discretization of $I - \alpha\partial_x^2$. We have the Lagrangian

$$\mathbf{L}(\mathbf{y}, \mathbf{u}, \mathbf{p}) = J(\mathbf{y}, \mathbf{u}) + \sum_{k=1}^{N_t} \mathbf{p}_k^\top \mathbf{c}_k(\mathbf{y}, \mathbf{u})$$

Differentiating this Lagrangian with respect to its arguments and setting the derivatives to zero gives the discrete optimality system

$$\begin{aligned} \mathbf{B}_k \mathbf{y}_k &= \mathbf{B}_k^\top \mathbf{y}_{k-1} \\ \mathbf{B}_k^\top \mathbf{p}_k &= \mathbf{B}_{k+1} \mathbf{p}_{k+1} \\ \gamma \mathbf{K}\mathbf{u} &= \mathbf{f} \end{aligned}$$

Results with globalized Newton method: optimal controls



Optimal controls for transitions from the first state to the second, the third, and the fifth states.

Results with globalized Newton method: minimization

Iteration	$J_{SD} - J^*$	$J_{NCG} - J^*$	$J_{KN} - J^*$
1	2.4969×10^{-1}	2.4969×10^{-1}	2.4969×10^{-1}
2	1.3070×10^{-2}	1.3070×10^{-2}	1.5346×10^{-2}
3	6.4184×10^{-3}	6.4184×10^{-3}	5.1099×10^{-3}
4	5.5337×10^{-3}	5.3438×10^{-3}	2.2381×10^{-4}
5	4.8170×10^{-3}	3.1011×10^{-3}	1.8383×10^{-4}
6	4.2081×10^{-3}	2.3384×10^{-3}	1.6253×10^{-5}
7	3.6768×10^{-3}	1.2475×10^{-3}	2.7534×10^{-6}
8	3.2177×10^{-3}	9.1869×10^{-5}	3.3921×10^{-7}
9	2.8141×10^{-3}	5.9258×10^{-5}	4.7022×10^{-9}

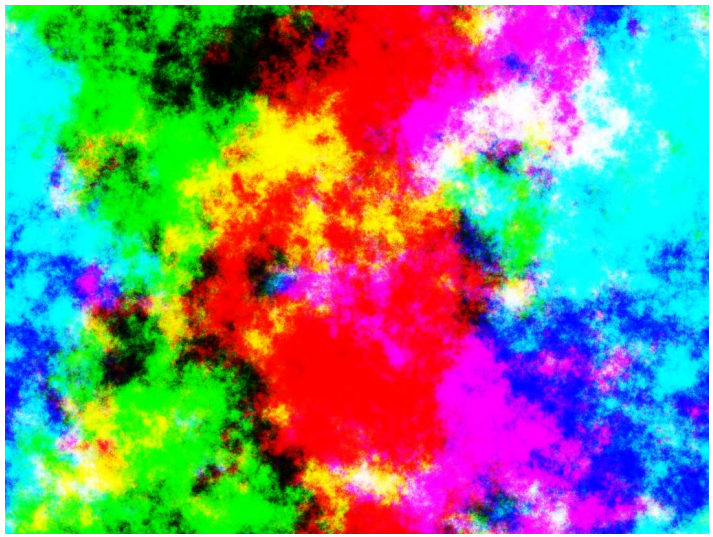
Table: Convergence of the SteepestDescent scheme, the NonlinearCG scheme, and the KrylovNewton scheme to reach the optimal cost $J^* = J(u^*)$.

Uncertainty on the potential

Consider quantum transition where the confining potential is subject to (thermal) space-time random field perturbation

$$V_0(x,t,\omega) = \bar{V}_0(x) + U(x,\omega)\, S(t,\omega)$$

where $x \in \Omega$, $t \in (0,T)$, and the event $\omega \in \mathcal{O}$.



The triple $(\mathcal{O}, \mathcal{A}, \mathcal{P})$ denotes a probability space where $\mathcal{O}$ is the space of elementary events, $\mathcal{A}$ is the sigma-algebra of subsets of $\mathcal{O}$, and $\mathcal{P}$ the probability measure on $\mathcal{O}$. Transitions $\phi_j \longrightarrow \phi_k$

Karhunen–Loève expansion

We use the Karhunen–Loève (KL) expansion of the random field $V_0(x,t,\omega)$, that is based on a spectral decomposition of the covariance kernel of the stochastic processes.
Assuming sufficient decay we consider the truncated KL expansion

$$U_N(x,\omega) = U_0(x) + \sum_{j=1}^{N} \sqrt{\lambda_j}\, z_j(x)\, Y_j(\omega)$$

and

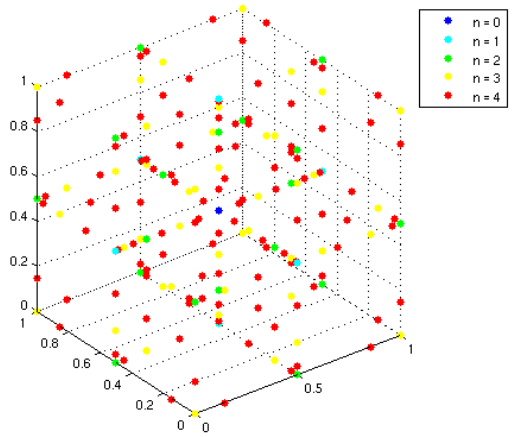
$$S_M(t,\omega) = S_0(t) + \sum_{j=1}^{M} \sqrt{\mu_j}\, \chi_j(t)\, W_j(\omega)$$

where N and M denote the number of terms in the truncation. The $(\lambda_j, z_j(x))$ and $(\mu_j, \chi_j(t))$ are the real positive eigenvalues and eigenvectors of the covariance Kernels.

Smolyak sparse-grids

To work on the $\mathcal{N} = N \times M$ full tensor-product probability space, we use sparse grids, thus drastically reduce the Curse of dimensionality.

A Smolyak scheme provides multi-variate interpolation as linear combination of tensor-product formulas.



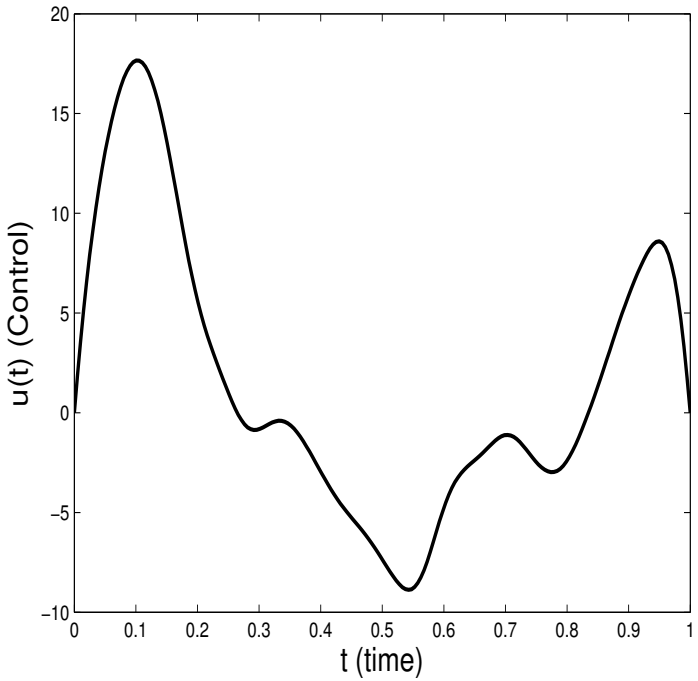
A full tensor-product $\mathcal{N}$-dimensional grid of order $\mathcal{J}$ with $j_1 = j_2 = \cdots j_{\mathcal{N}} = \mathcal{J}$ in each dimension is given by

$$\hat{z}^{\mathcal{J}} = \bar{z}^{j_1} \otimes \bar{z}^{j_2} \otimes \cdots \otimes \bar{z}^{j_N}$$

The sparse grid $\tilde{z}^{\mathcal{J}}$ of order $\mathcal{J}$ is composed of a strict subset of full grids. For example, for $\mathcal{N} = 2$ and $\mathcal{J} = 4$ we have

$$\{z^1 \otimes z^4\} \bigcup \{z^2 \otimes z^3\} \bigcup \{z^3 \otimes z^2\} \bigcup \{z^4 \otimes z^1\}$$

Optimal controls for different potentials



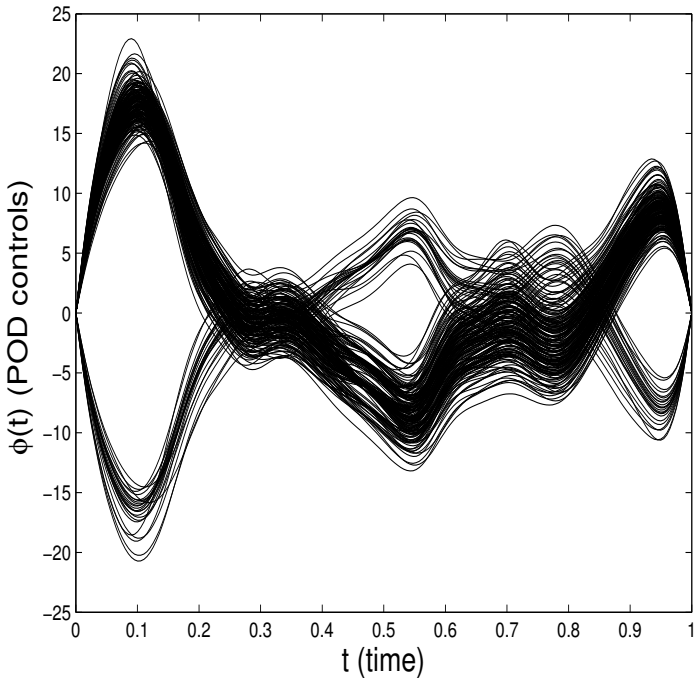


Figure: *Left*: The optimal control for the unperturbed potential problem. *Right*: The optimal controls for 181 potentials.

The quest for a unique robust control

A possible way to define a robust control is the following

$$\begin{cases} \min_{u\in\mathcal{U}} \int_{\mathcal{O}} J(\psi(\omega,u),u)\, d\mathcal{P}(\omega) \\ \\ i\frac{\partial}{\partial t}\psi = \left\{-\nabla^2 + V(\omega,u)\right\}\psi \end{cases} \qquad \omega \text{ in } \mathcal{O}$$

To approximate the solution to this problem we could

1. Construct the mean of the controls corresponding to each event ω, i.e. $\tilde{u} = \mathbb{E}(u)$; or use u of the unperturbed potential.
2. Define a reduced model of controls based on POD; minimize on the POD basis;
3. Solve the full optimality system discretized on the probaility space.

Two useful metrics to evaluate the quality of the control:
$\langle J(u,\omega)\rangle$ – The average cost functional;
$|J(u,\omega)|_\infty$ – the maximum value of the objective in the parameter space.

Reduced model of controls

Let $u_j(t)$, $j = 1, \ldots, \chi$ represent the controls for different $\omega \in \mathcal{O}$, then the POD basis functions are obtained from

$$A_{jk} = \int_0^T u_j(t) u_k(t) dt, \qquad A = V^\top \Sigma V.$$

The kth POD basis function is $\phi_k(t) = \frac{1}{\sqrt{\sigma_k}} \sum_{j=1}^{m} V_{jk} u_j(t)$, where $m \leq \chi$.

$$u(t) = \sum_{k=1}^{m} c_k \phi_k(t) \qquad \text{i.e.} \qquad \mathbf{u} = \Phi \mathbf{c} \qquad \text{with } \Phi_{ij} = (\phi_j(t_i)) \in \mathbb{R}^{n \times m}$$

The reduced gradient with respect to the POD expansion coefficients is

$$\nabla \hat{J}(\mathbf{c}) = \gamma \Phi^\top K \Phi \mathbf{c} - \Phi^\top f$$

The Hessian for the problem is $\nabla^2 \hat{J}(\mathbf{c}) = \Phi^\top \nabla^2 \hat{J}(\mathbf{c}) \Phi$. Assuming that $n >> m$, computing the search direction $\delta \mathbf{c}$ requires solving the small set of equations

$$\Phi^\top \nabla^2 \hat{J}(\mathbf{c}) \Phi \, \delta \mathbf{c} = -\Phi^\top \nabla \hat{J}(\mathbf{c}).$$

Full optimality system

We have the Lagrangian

$$L(\psi_1, ..., \psi_m, u, p_1, ..., p_m) = J(\psi_1, ..., \psi_m, u) + \sum_{k=1}^{m} w_k \, p_k^\top \, c_k(\psi_k, u)$$

Differentiating gives the full gradient

$$L_u = \gamma u^\top K + \sum_{k=1}^{n} w_k \, p_k^\top \frac{\partial c_k}{\partial u}$$

$$L_{\psi_k} = -w_k \, \psi_k^\top P_k + w_k \, p_k^\top \frac{\partial c_k}{\partial \psi_k}$$

$$L_{p_k} = c_k^\top (\psi_k, u)$$

where P_k is projector matrix onto the target.

Results on robust controls

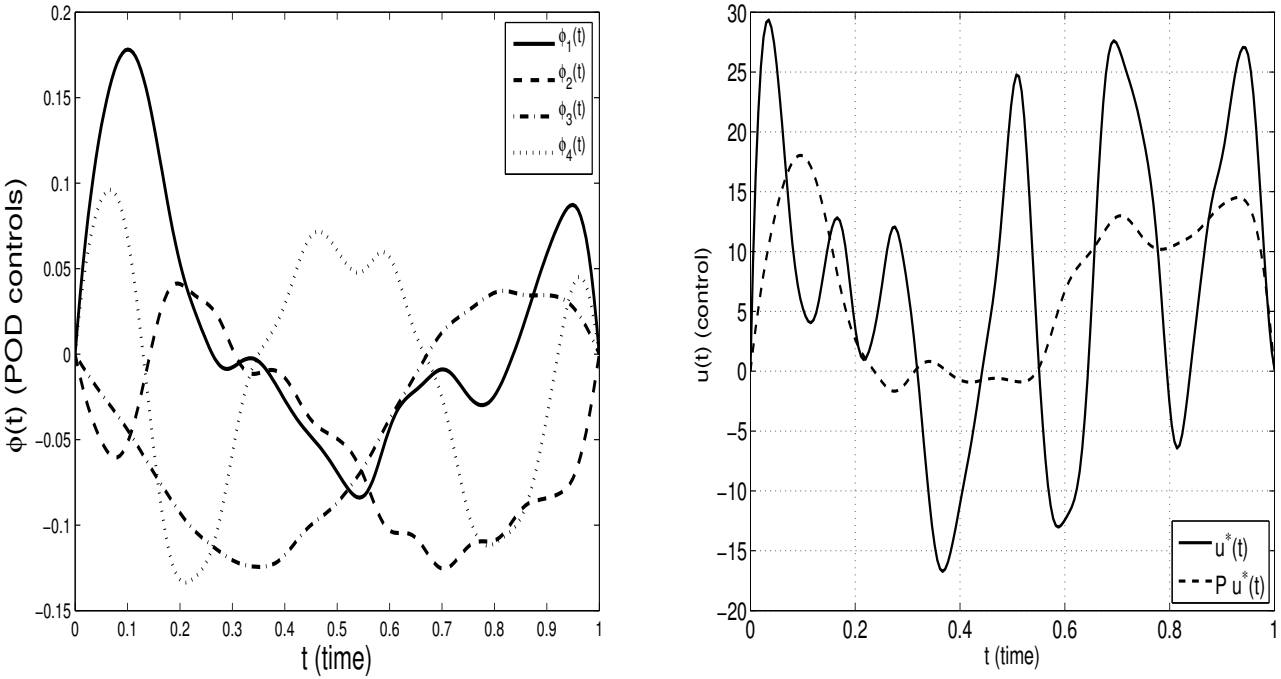


Figure: *Left*: The first four POD functions. *Right*: The optimal control for the coupled problem (solid) and its best L^2 approximation in the POD basis (dashed).

Control	$\langle J(\bar{u},\omega)\rangle$	$\lvert J(\bar{u},\omega)\rvert_\infty$
u_{unpert}	4.02×10^{-2}	1.06×10^{-1}
$u_{fullmin}$	2.56×10^{-2}	7.89×10^{-2}
$u_{5podmin}$	2.70×10^{-4}	5.56×10^{-3}

References

1. G. von Winckel and A.B., *Optimal control of quantum well transitions with uncertain potential*, in preparation.
2. G. von Winckel and A.B., *A globalized Newton method for the accurate solution of a dipole quantum control problem*, submitted.
3. G. von Winckel and A. B., *Computational techniques for a quantum control problem with H^1-cost*, Inverse Problems, 24 (2008), 034007.
4. A. B. and G. von Winckel, *Multigrid methods and sparse-grid collocation techniques for parabolic optimal control problems with random coefficients*, in SIAM J. Scientific Computing
5. A. B. and U. Hohennester, *Multigrid optimization schemes for solving Bose-Einstein condensate control problems*, SIAM J. Scientific Computing, 30 (2008), pp. 441–462.
6. A. B., J. Salomon, and S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, J. Comput. Appl. Math., 216 (2008), pp. 170–197.
7. A. B., G. Stadler, and U. Hohenester, *Optimal quantum control in nanostructures: Theory and application to a generic three-level system*, Phys. Rev. A 66, (2002) 053811.

Thanks for your attention