Telling optimal control how to maximize the entangling power of two-qubit gates

Christiane P. Koch*

joint work with

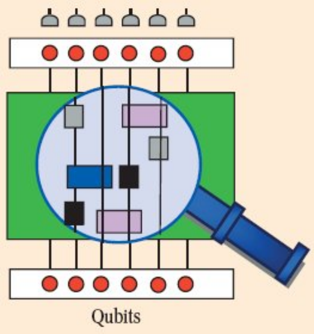
Daniel Reich*, Michael Goerz*, Mamadou Ndong*,

Matthias M. Müller,

Jiri Vala, Haidong Yuan, Birgitta Whaley & Tommaso Calarco

* Institut für Theoretische Physik, Freie Universität Berlin, Germany

introduction: optimal control & QIP

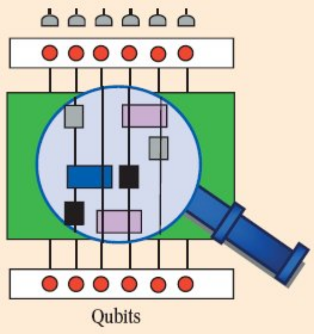


$$\mathrm{Tr}\left\{\hat{\mathbf{O}}^{+}\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}$$

- desired gate operation : $\hat{\mathbf{O}}$
- actual evolution : $\hat{\mathbf{U}}(T,0;\varepsilon)$
- desired fidelity :

  $1-\epsilon$ where $\epsilon < 10^{-4}$

introduction:
optimal control & QIP



$$\mathrm{Tr}\left\{\hat{\mathbf{O}}^{+}\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}$$

- desired gate operation : $\hat{\mathbf{O}}$
- actual evolution : $\hat{\mathbf{U}}(T,0;\varepsilon)$
- desired fidelity :

  $1-\epsilon$ where $\epsilon < 10^{-4}$

- **choice of $\hat{\mathbf{O}}$?**

outline

1. introduction: optimal control
    - terminology & basic concepts
    - functionals: how to convey the physics to the algorithm
2. a functional to optimize the entangling power of two-qubit gates
    - local equivalence classes
    - the new functional
    - 2nd order Krotov algorithm
3. application to a Rydberg gate in cold atoms
4. where can we go from here?

some terminology & basics of optimal control

coherent control

quantum mechanics $\triangleq$ probabilistic, but
deterministic theory

$$|\psi(t=0)\rangle \quad \overset{\text{Schrödinger equation}}{\rightsquigarrow} \quad |\psi(t>0)\rangle$$

given an initial state,
$|\psi(t=0)\rangle$ or $\hat{\boldsymbol{\rho}}(t=0)$,
which dynamics ($\triangleq$ which $\hat{\mathbf{H}}$)
guarantees a particular outcome,
$|\psi(t>0)\rangle$ or $\hat{\boldsymbol{\rho}}(t>0)$?

principle of coherent control

wave properties of matter
(superposition principle)

variation of phase between different, but
indistinguishable quantum pathways

↓

constructive
interference in
desired channel

destructive
interference in all
other channels

final goal: understanding the intricate workings of quantum interferences

principle of coherent control (rev'd)

$\left[\begin{array}{c}\text{wave properties of matter}\\ \text{(superposition principle)}\end{array}\right]$ *interacting* quantum (sub)systems

variation of phase between different, but indistinguishable quantum pathways

↓

constructive *interference* in desired channel

destructive *interference* in all other channels

final goal: understanding the intricate workings of interferences & entanglement

coherent control & optimal control

- goal: improve outcome of process
- vary some parameters

↓

simple, intuitive schemes

bichromatic c., pump-dump, STIRAP

in time or in frequency domain

coherent control & optimal control

- goal: improve outcome of process
- vary some parameters

↓

simple, intuitive schemes

bichromatic c., pump-dump, STIRAP

in time or in frequency domain

- goal: obtain maximum control over process
- tune ‘all’ available parameters

↓

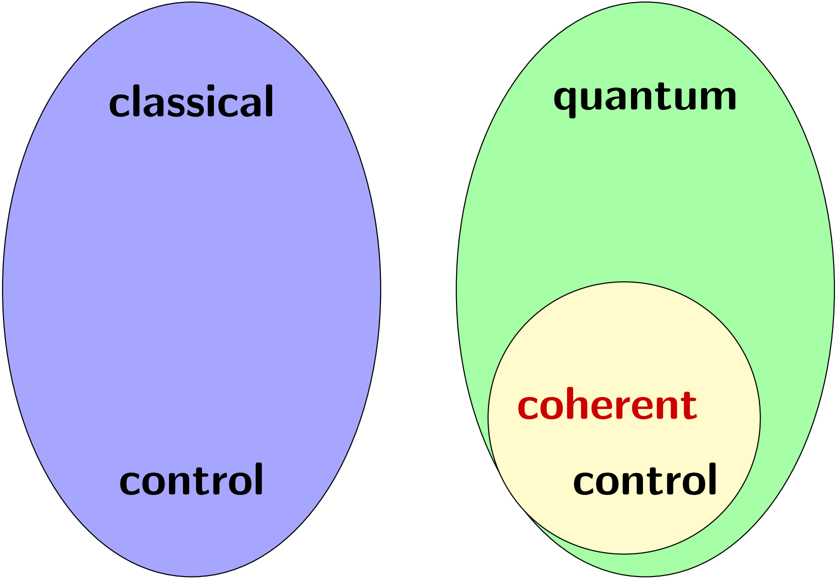
complex outcome →
discovery of new schemes

not necessarily accessible by intuition

in time / frequency
“phase space”

global or local in time

classical vs quantum control



classical

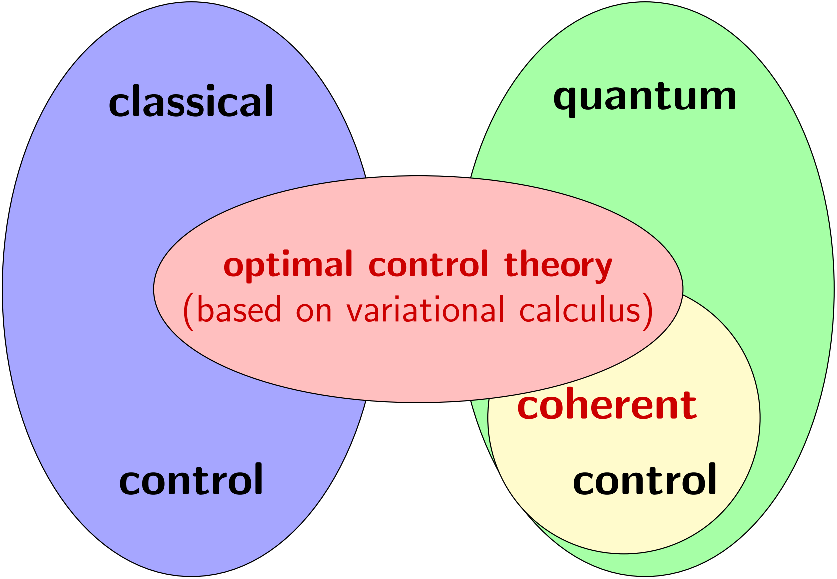
control

quantum

coherent

control

classical vs quantum control



classical

control

quantum

coherent

control

optimal control theory

(based on variational calculus)

let’s put philosophy aside

. . .

work out the tools

- they will be needed anyhow -

optimal control theory

time/frequency 'phasespace' picture

$$\begin{array}{ccc} t = 0 & & t = T \\ |\varphi_i\rangle & \rightsquigarrow & |\varphi_f\rangle \end{array}$$

inverse problem:
given the target and the equations of motion,
calculate the field

optimal control theory

time/frequency 'phasespace' picture

$$\begin{array}{ccc} t=0 & & t=T \\ |\varphi_i\rangle & \rightsquigarrow & |\varphi_f\rangle \end{array}$$

inverse problem:
given the target and the equations of motion,
calculate the field

local in time	*global in time*
impose 2 conditions (for phase and amplitude) → derive equations for field	information from dynamics throughout time interval to reach desired target at final T
also: Lyapunov control / tracking	*OCT*

optimal control theory

$$\begin{array}{ccc} t = 0 & & t = T \\ |\varphi_i\rangle & \rightsquigarrow & |\varphi_f\rangle \end{array}$$

define the objective :

$$\boxed{\mathrm{GOAL} \equiv \| \langle \varphi_i | \hat{\mathbf{U}}^{+}(T, 0; \varepsilon) | \varphi_f \rangle \|^2 = -F}$$

as a functional of the field ε

optimal control theory: variants

variational approach

- ‘guess’ the right functional, including eqs. of motion & phasefactors
- do the variations to obtain eqs. of motion and eq. for the field
- ‘guess’ the correct time discretization s.t. method converges

W. Zhu, J. Botina, H. Rabitz, JCP 108, 1953 (1998)

optimal control theory: variants

variational approach

- 'guess' the right functional, including eqs. of motion & phasefactors
- do the variations to obtain eqs. of motion and eq. for the field
- 'guess' the correct time discretization s.t. method converges

W. Zhu, J. Botina, H. Rabitz, JCP 108, 1953 (1998)

Krotov method

- ingredients: objective + constraint(s) + eqs. of motion
- construct auxiliary functional L with auxil. potential to guarantee monoton. convergence
- derive the eq. for $\varepsilon(t)$ from the minimization of L

Sklarz & Tannor, PRA 66, 053619 (2002), Palao & Kosloff, PRA 68, 062308 (2003)

optimal control theory: schemes

improve the field by

$$\Delta\varepsilon^{j+1}(t) =$$

$$\frac{S(t)}{\lambda_0}\,\mathfrak{Im}\left[\underbrace{\langle\varphi_i|\,\hat{\mathbf{U}}^{+}(T,0;\varepsilon^{j})}_{\substack{\text{forward}\\\text{propagation (1)}}}\,|\varphi_f\rangle\,\underbrace{\langle\varphi_f|\,\hat{\mathbf{U}}^{+}(t,T;\varepsilon^{j})}_{\substack{\text{backward}\\\text{propagation (2)}}}\,\hat{\boldsymbol{\mu}}\,\underbrace{\hat{\mathbf{U}}(t,0;\varepsilon^{j+1})\,|\varphi_i\rangle}_{\substack{\text{forward}\\\text{propagation (3)}}}\right]$$

optimal control theory: schemes

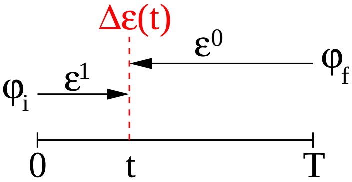
improve the field by

$$\Delta\varepsilon^{j+1}(t) =$$

$$\frac{S(t)}{\lambda_0}\,\Im\mathfrak{m}\left[\underbrace{\langle\varphi_i|\,\hat{\mathbf{U}}^{+}(T,0;\varepsilon^{j})}_{\substack{\text{forward}\\ \text{propagation (1)}}}\,|\varphi_f\rangle\,\underbrace{\langle\varphi_f|\,\hat{\mathbf{U}}^{+}(t,T;\varepsilon^{j})}_{\substack{\text{backward}\\ \text{propagation (2)}}}\,\hat{\mu}\,\underbrace{\hat{\mathbf{U}}(t,0;\varepsilon^{j+1})\,|\varphi_i\rangle}_{\substack{\text{forward}\\ \text{propagation (3)}}}\right]$$

interference between

past and future events



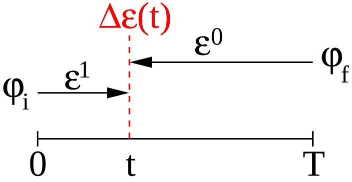
optimal control theory: schemes

improve the field by

$$\Delta\varepsilon^{j+1}(t) = \frac{S(t)}{\lambda_0}\,\mathfrak{Im}\left[\underbrace{\langle\varphi_i|\,\hat{\mathbf{U}}^{+}(T,0;\varepsilon^{j})}_{\substack{\text{forward}\\ \text{propagation (1)}}}\,|\varphi_f\rangle\,\underbrace{\langle\varphi_f|\,\hat{\mathbf{U}}^{+}(t,T;\varepsilon^{j})}_{\substack{\text{backward}\\ \text{propagation (2)}}}\,\hat{\mu}\,\underbrace{\hat{\mathbf{U}}(t,0;\varepsilon^{j+1})\,|\varphi_i\rangle}_{\substack{\text{forward}\\ \text{propagation (3)}}}\right]$$

interference between

past and future events



variational approach & Krotov method lead to similar schemes

Maday & Turinici, JCP 118, 8191 (2003) & work by N.C. Nielsen group

optimal control theory: schemes

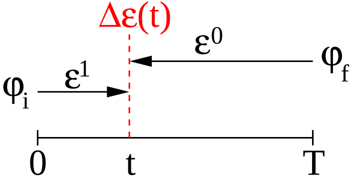
improve the field by

$$\Delta\varepsilon^{j+1}(t) =$$

$$\frac{S(t)}{\lambda_0}\,\mathfrak{Im}\left[\underbrace{\langle\varphi_i|\,\hat{\mathbf{U}}^{+}(T,0;\varepsilon^{j})}_{\substack{\text{forward}\\ \text{propagation (1)}}}\,|\varphi_f\rangle\,\underbrace{\langle\varphi_f|\,\hat{\mathbf{U}}^{+}(t,T;\varepsilon^{j})}_{\substack{\text{backward}\\ \text{propagation (2)}}}\,\hat{\mu}\,\underbrace{\hat{\mathbf{U}}(t,0;\varepsilon^{j+1})\,|\varphi_i\rangle}_{\substack{\text{forward}\\ \text{propagation (3)}}}\right]$$

interference between

past and future events



variational approach & Krotov method lead to similar schemes

Maday & Turinici, JCP 118, 8191 (2003) & work by N.C. Nielsen group

to date: reduced Krotov method only (1st order variant)

functionals

or

how to convey the desired physics
to the OCT algorithm

objective functionals / costs

$$J\left[\{\varphi_k(t),\varphi_k^*(t)\},\varepsilon(t)\right] =$$

$$J_T\left[\{\varphi_k(T),\varphi_k^*(T)\}\right] + J_t\left[\{\varphi_k(t),\varphi_k^*(t)\},\varepsilon(t)\right]$$

final-time target	intermediate-time target time-dependent cost state-dependent cost

functionals of the field $\boldsymbol{\varepsilon}(\mathbf{t})$

- explicitly
- implicitly through $\varphi_k(t)$, $\varphi_k(T)$

final-time objectives J_T

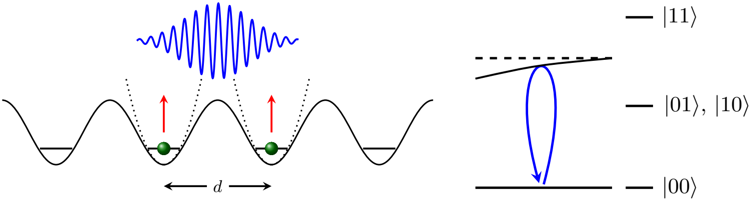
(i) state-to-state transfer

(ii) unitary transformation

final-time objectives J_T

(i) state-to-state transfer

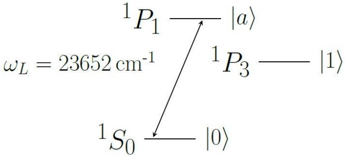
(ii) unitary transformation



goal: perform a two-qubit gate on the logical basis

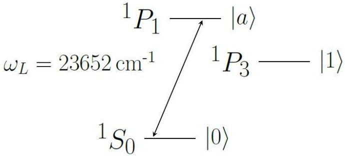
example: phasegate for two Ca atoms

qubit encoding:



example: phasegate for two Ca atoms

qubit encoding:

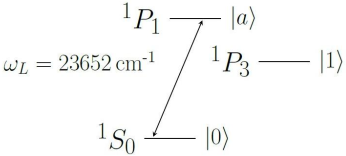


two-atom Hamiltonian:

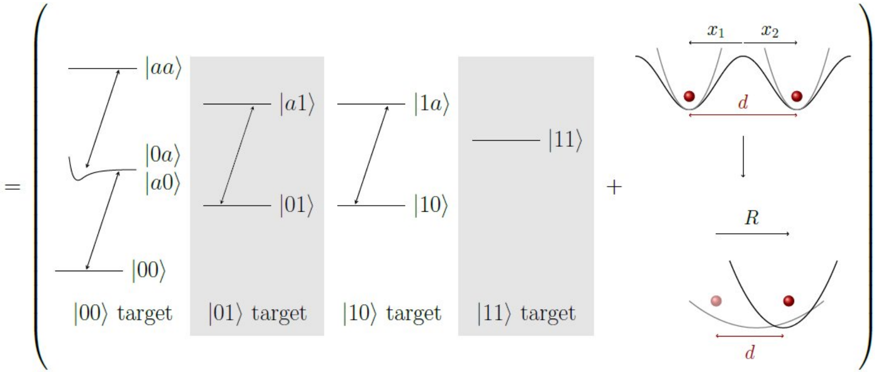
$$\hat{H}_{2q} = \begin{pmatrix} E_0 & 0 & \mu\epsilon(t) \\ 0 & E_1 & 0 \\ \mu\epsilon(t) & 0 & E_a \end{pmatrix} \otimes \mathbb{1}_{1q} \otimes \mathbb{1}_{x_1} + \\ + \mathbb{1}_{x_2} \otimes \mathbb{1}_{1q} \otimes \begin{pmatrix} E_0 & 0 & \mu\epsilon(t) \\ 0 & E_1 & 0 \\ \mu\epsilon(t) & 0 & E_a \end{pmatrix} + \\ + \sum_{ij} \hat{V}^{(ij)}_{BO}(|x_2 - x_1|) + \sum_{ij} \hat{V}^{(ij)}_{\mathrm{trap}}(x_1, x_2)$$

example: phasegate for two Ca atoms

qubit encoding:



two-atom Hamiltonian:

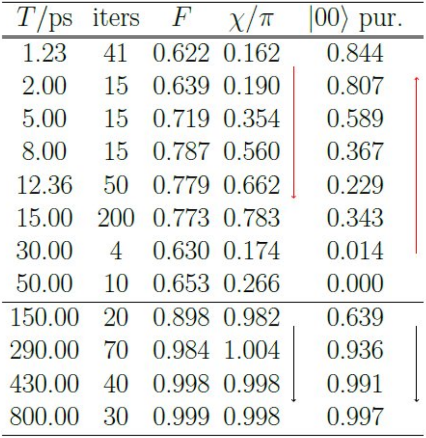


example: phasegate for two Ca atoms

target: true two-qubit phase

$$\chi = \phi_{00} - \phi_{01} - \phi_{10} + \phi_{11}$$

$d = 5\,\mathrm{nm}$



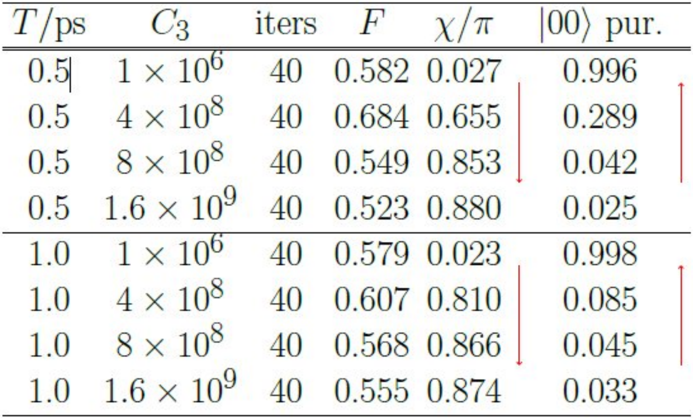
T/ps	iters	F	χ/π	$\lvert 00\rangle$ pur.
1.23	41	0.622	0.162	0.844
2.00	15	0.639	0.190	0.807
5.00	15	0.719	0.354	0.589
8.00	15	0.787	0.560	0.367
12.36	50	0.779	0.662	0.229
15.00	200	0.773	0.783	0.343
30.00	4	0.630	0.174	0.014
50.00	10	0.653	0.266	0.000
150.00	20	0.898	0.982	0.639
290.00	70	0.984	1.004	0.936
430.00	40	0.998	0.998	0.991
800.00	30	0.999	0.998	0.997

example: phasegate for two Ca atoms

target: true two-qubit phase

$$\chi = \phi_{00} - \phi_{01} - \phi_{10} + \phi_{11}$$

$d = 200\,\mathrm{nm}$: scaling up interaction C_3



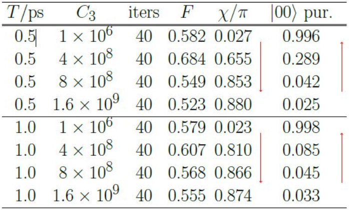
T/ps	C_3	iters	F	χ/π	$\|00\rangle$ pur.
0.5	1×10^6	40	0.582	0.027	0.996
0.5	4×10^8	40	0.684	0.655	0.289
0.5	8×10^8	40	0.549	0.853	0.042
0.5	1.6×10^9	40	0.523	0.880	0.025
1.0	1×10^6	40	0.579	0.023	0.998
1.0	4×10^8	40	0.607	0.810	0.085
1.0	8×10^8	40	0.568	0.866	0.045
1.0	1.6×10^9	40	0.555	0.874	0.033

example: phasegate for two Ca atoms

target: true two-qubit phase

$$\chi = \phi_{00} - \phi_{01} - \phi_{10} + \phi_{11}$$

d = 200 nm: scaling up interaction C_3



T/ps	C_3	iters	F	χ/π	$\|00\rangle$ pur.
0.5	1×10^6	40	0.582	0.027	0.996
0.5	4×10^8	40	0.684	0.655	0.289
0.5	8×10^8	40	0.549	0.853	0.042
0.5	1.6×10^9	40	0.523	0.880	0.025
1.0	1×10^6	40	0.579	0.023	0.998
1.0	4×10^8	40	0.607	0.810	0.085
1.0	8×10^8	40	0.568	0.866	0.045
1.0	1.6×10^9	40	0.555	0.874	0.033

limiting factor for fast gate:

sufficient time to resolve GS motional dynamics

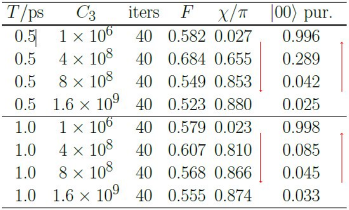
(will play a role in any scheme based on resonant excitation)

example: phasegate for two Ca atoms

target: true two-qubit phase

$$\chi = \phi_{00} - \phi_{01} - \phi_{10} + \phi_{11}$$

d = 200 nm: scaling up interaction C_3



T/ps	C_3	iters	F	χ/π	$\|00\rangle$ pur.
0.5	1×10^6	40	0.582	0.027	0.996
0.5	4×10^8	40	0.684	0.655	0.289
0.5	8×10^8	40	0.549	0.853	0.042
0.5	1.6×10^9	40	0.523	0.880	0.025
1.0	1×10^6	40	0.579	0.023	0.998
1.0	4×10^8	40	0.607	0.810	0.085
1.0	8×10^8	40	0.568	0.866	0.045
1.0	1.6×10^9	40	0.555	0.874	0.033

limiting factor for fast gate:

sufficient time to resolve GS motional dynamics

(will play a role in any scheme based on resonant excitation)

what about flexibility in single-qubit phases?

final-time objectives J_T

$$J_T = -\frac{\lambda_0}{N}\mathfrak{Re}\left[\mathrm{Tr}\left\{\hat{\mathbf{O}}^{+}\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}\right]$$

real-valued, phase-sensitive functional

- $\hat{\mathbf{O}}$ target operator
- λ_0 weight
- $N = \dim\{\hat{\mathbf{O}}\}$

- $\hat{\mathbf{P}}_N$ projector on subspace of $\hat{\mathbf{O}}$
- $\hat{\mathbf{U}}(T,0;\varepsilon)$ actual time evolution

final-time objectives J_T

$$J_T = -\frac{\lambda_0}{N}\mathfrak{Re}\left[\mathrm{Tr}\left\{\hat{\mathbf{O}}^{+}\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}\right]$$

real-valued, phase-sensitive functional

- $\hat{\mathbf{O}}$ target operator
- λ_0 weight
- $N = \dim\{\hat{\mathbf{O}}\}$
- $\hat{\mathbf{P}}_N$ projector on subspace of $\hat{\mathbf{O}}$
- $\hat{\mathbf{U}}(T,0;\varepsilon)$ actual time evolution

- state-to-state transfer: $\hat{\mathbf{O}} = |\varphi_{\text{target}}\rangle\langle\varphi_{\text{target}}|$, $N = 1$
- single-qubit gate: $N = 2$, two-qubit gate: $N = 4$

cf. Palao & Kosloff, PRA 68, 062308 (2003)

intermediate-time objectives J_t

assumption: additive costs

$$J_t = \int_0^T \left\{ g_a\left[\varepsilon(t)\right] + g_b\left[\varphi(t), \varphi^*(t)\right] \right\} dt$$

intermediate-time objectives J_t

assumption: additive costs

$$J_t = \int_0^T \left\{ g_a\left[\varepsilon(t)\right] + g_b\left[\varphi(t), \varphi^*(t)\right] \right\} dt$$

examples

$$g_a\left[\varepsilon(t)\right] = \lambda_a S(t) \left[\varepsilon(t) - \varepsilon_{\mathrm{ref}}(t)\right]^2$$

- minimization of field intensity ($\varepsilon_{\mathrm{ref}}(t) = 0$) or change in field intensity ($\varepsilon_{\mathrm{ref}}(t) = \varepsilon_{\mathsf{old}}$)
- choice of $\varepsilon_{\mathrm{ref}}(t)$ determines update vs replacement rule !

intermediate-time objectives J_t

assumption: additive costs

$$J_t = \int_0^T \{g_a[\boldsymbol{\varepsilon}(t)] + g_b[\varphi(t), \varphi^*(t)]\}\, dt$$

examples

$$g_a[\boldsymbol{\varepsilon}(t)] = \lambda_a S(t)\,[\boldsymbol{\varepsilon}(t) - \varepsilon_{\mathrm{ref}}(t)]^2$$

- minimization of field intensity ($\varepsilon_{\mathrm{ref}}(t) = 0$) or change in field intensity ($\varepsilon_{\mathrm{ref}}(t) = \varepsilon_{\mathsf{old}}$)
- choice of $\varepsilon_{\mathrm{ref}}(t)$ determines update vs replacement rule !

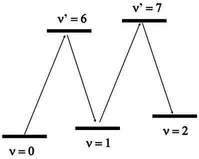
$$g_b[\varphi(t), \varphi^*(t)] = \lambda_b \langle \varphi(t) | \hat{\mathbf{D}}(t) | \varphi(t) \rangle$$

- $\hat{\mathbf{D}}(t)$ target operator
- λ_a, λ_b weights, $S(t)$ switch/shape function

time-dependent targets

$$g_b\left[\varphi(t),\varphi^*(t)\right] = \lambda_b\langle\varphi(t)|\hat{\mathbf{D}}(t)|\varphi(t)\rangle$$

prescribing a desired evolution



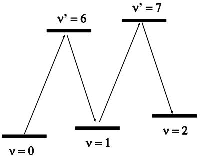
$$\begin{aligned}\hat{\mathbf{D}}(t) &= |6\rangle\langle 6|\Theta(T_1-t)+\\ &\quad |1\rangle\langle 1|\Theta(t-T_1)\Theta(T_2-t)+\\ &\quad |7\rangle\langle 7|\Theta(t-T_2)\Theta(T_3-t)+\\ &\quad |2\rangle\langle 2|\Theta(t-T_3)\Theta(T-t)\end{aligned}$$

Ndong, Tal-Ezer, Kosloff, CPK,
JCP 130, 124108 (2009)

time-dependent targets

$$g_b\left[\varphi(t), \varphi^*(t)\right] = \lambda_b \langle \varphi(t) | \hat{\mathbf{D}}(t) | \varphi(t) \rangle$$

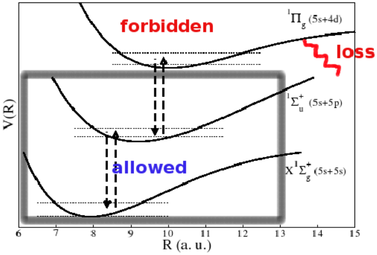
prescribing a desired evolution



$$\begin{aligned}\hat{\mathbf{D}}(t) \;=\; & |6\rangle\langle 6|\Theta(T_1 - t) + \\ & |1\rangle\langle 1|\Theta(t - T_1)\Theta(T_2 - t) + \\ & |7\rangle\langle 7|\Theta(t - T_2)\Theta(T_3 - t) + \\ & |2\rangle\langle 2|\Theta(t - T_3)\Theta(T - t)\end{aligned}$$

Ndong, Tal-Ezer, Kosloff, CPK,
JCP 130, 124108 (2009)

keeping the dynamics in a subspace



$$\hat{\mathbf{D}}(t) = \hat{\mathbf{P}}_{\text{allow}}$$

Palao, Kosloff, CPK,
PRA 77, 063412 (2008)

where are we? outline!

1. introduction: optimal control
    - terminology & basic concepts
    - functionals: how to convey the physics to the algorithm

2. a functional to optimize the entangling power of two-qubit gates
    - local equivalence classes
    - the new functional
    - 2nd order Krotov algorithm

3. application to a Rydberg gate in cold atoms

4. where can we go from here?

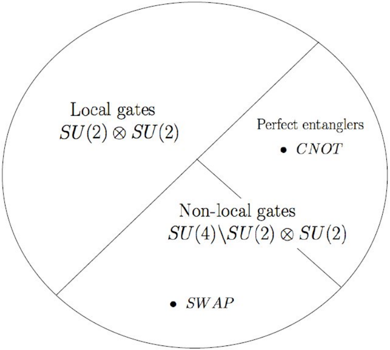
local equivalence classes

classification of two-qubit gates

$G = SU(4)$ group of all two-qubit gates
$K = SU(2) \otimes SU(2)$ local gates
$G/K = SU(4)/SU(2) \otimes SU(2)$
non-local gates

$$\mathfrak{su}(4) = \mathfrak{k} \oplus \mathfrak{p}$$

Cartan decomposition of Lie algebras



$$\hat{\mathbf{U}} = \hat{\mathbf{k}}_1 e^{-\frac{i}{2}\sum_{j=x,y,z} c_j \hat{\boldsymbol{\sigma}}_j^1 \hat{\boldsymbol{\sigma}}_j^2} \hat{\mathbf{k}}_2$$

$$\hat{\mathbf{U}}_2 = \hat{\mathbf{k}}_1 \hat{\mathbf{U}}_2 \hat{\mathbf{k}}_2$$

Zhang, Vala, Sastry, Whaley,
PRA 67, 042313 (2003)

Yuan & Khaneja, PRA 72, 040301 (2005)

local invariants

$$\hat{\mathbf{m}} = \hat{\mathbf{U}}_B^T \hat{\mathbf{U}}_B$$

$$\hat{\mathbf{U}}_B = \hat{\mathbf{Q}}^+ \hat{\mathbf{U}} \hat{\mathbf{Q}} \quad \text{(i.e. } \hat{\mathbf{U}} \text{ in Bell basis)}$$

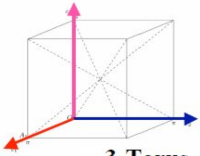
$$\begin{aligned} g_1 &= \mathfrak{Re}\,\mathrm{Tr}[\hat{\mathbf{m}}]^2/16\det(\hat{\mathbf{U}}) \\ g_2 &= \mathfrak{Im}\,\mathrm{Tr}[\hat{\mathbf{m}}]^2/16\det(\hat{\mathbf{U}}) \\ g_3 &= \mathrm{Tr}[\hat{\mathbf{m}}]^2 - \mathrm{Tr}[\hat{\mathbf{m}}^2]/4\det(\hat{\mathbf{U}}) \end{aligned}$$

g_1, g_2, g_3 define local equivalence class $[\hat{\mathbf{U}}]$,

i.e. a class of two-qubit gates that are equivalent up to local (single-qubit) operations

Zhang, Vala, Sastry, Whaley, PRA 67, 042313 (2003)

Weyl chamber



3-Torus

Cartan decomposition

$$U = k_1 A k_2 = k_1 \, exp[\frac{i}{2}(c_1 \sigma_x^1\sigma_x^2 + c_2 \sigma_y^1\sigma_y^2 + c_3 \sigma_z^1\sigma_z^2)] \, k_2$$

Local invariants

$$g_1 = \cos c_1 \cos c_2 \cos c_3$$

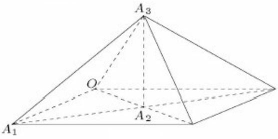
$$g_2 = \sin c_1 \sin c_2 \sin c_3$$

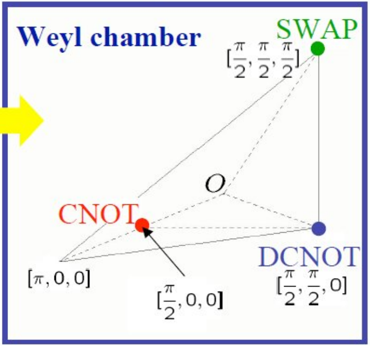
$$g_3 = 2\,(\cos^2 c_1 + \cos^2 c_2 + \cos^2 c_3) - 3$$

J. Zhang, J. Vala, S.Sastry, K.B. Whaley
Phys. Rev. A 67, 042313 (2003)

symmetry reduction

g_1,g_2 and g_3 are invariant with permutation of c_1, c_2, and c_3 with/without sign flips





There is a one-to-one correspondence between the points inside the Weyl chamber and local equivalence classes

the new functional

optimization target [Ô] instead of Ô

(old) functional to obtain Ô

$$J_T = -\frac{\lambda_0}{N}\mathfrak{Re}\left[\operatorname{Tr}\left\{\hat{\mathbf{O}}^{+}\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}\right]$$

(new) functional to obtain [Ô]

$$J_T = \Delta g_1{}^2 + \Delta g_2{}^2 + \Delta g_3{}^2$$

with $\Delta g_i{}^2 = \left|g_i(\hat{\mathbf{O}}) - g_i(\hat{\mathbf{U}})\right|^2$ and $g_i(\hat{\mathbf{O}})$ the local invariants of $\hat{\mathbf{O}}$

optimization target [Ô] instead of Ô

(old) functional to obtain Ô

$$J_T = -\frac{\lambda_0}{N}\mathfrak{Re}\left[\mathrm{Tr}\left\{\hat{\mathbf{O}}^+\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}\right]$$

(new) functional to obtain [Ô]

$$J_T = \Delta g_1{}^2 + \Delta g_2{}^2 + \Delta g_3{}^2$$

with $\Delta g_i{}^2 = \left|g_i(\hat{\mathbf{O}}) - g_i(\hat{\mathbf{U}})\right|^2$ and $g_i(\hat{\mathbf{O}})$ the local invariants of $\hat{\mathbf{O}}$

remember:

$$J = J\left[\{\varphi_k(t), \varphi_k^*(t)\}, \varepsilon(t)\right]$$

to carry out variations, we need to express g_i in terms of $\varphi_k(t)$

functional based on local invariants

using the definition of the invariants and of the Bell basis

functional based on local invariants

using the definition of the invariants and of the Bell basis
and after quite some algebra

functional based on local invariants

using the definition of the invariants and of the Bell basis and after quite some algebra

$$\boxed{J_T = f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5}$$

functional based on local invariants

using the definition of the invariants and of the Bell basis
and after quite some algebra

$$J_T = f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5$$

$$\begin{aligned}
f_1 &= \mathfrak{Re}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) \\
f_2 &= \mathfrak{Im}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) \\
f_3 &= \mathfrak{Re}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - \left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)^2 - \left(\vec{\beta}_k\cdot\vec{\beta}_l\right)^2 \\
&\quad +2\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right) \\
f_4 &= \mathfrak{Im}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\beta}_k\cdot\vec{\beta}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right)
\end{aligned}$$

functional based on local invariants

using the definition of the invariants and of the Bell basis
and after quite some algebra

$$\boxed{J_T = f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5}$$

$$\begin{aligned}
f_1 &= \mathfrak{Re}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) \\
f_2 &= \mathfrak{Im}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) \\
f_3 &= \mathfrak{Re}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - \left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)^2 - \left(\vec{\beta}_k\cdot\vec{\beta}_l\right)^2 \\
&\quad +2\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right) \\
f_4 &= \mathfrak{Im}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\beta}_k\cdot\vec{\beta}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right)
\end{aligned}$$

with $a_0 = \mathrm{Tr}^2\left(\hat{\mathbf{m}}_O\right)/16\det(\hat{\mathbf{O}})$ and $b_0 = \left[\mathrm{Tr}^2\left(\hat{\mathbf{m}}_O\right) - \mathrm{Tr}\left(\hat{\mathbf{m}}_O^2\right)\right]/4\det(\hat{\mathbf{O}})$

$(\alpha_k)_m = \mathfrak{Re}\left[\langle m|\varphi_k(T)\rangle\right]$, $(\beta_k)_m = \mathfrak{Im}\left[\langle m|\varphi_k(T)\rangle\right]$, $m = 1,\dots,\dim(\mathcal{H})$

functional based on local invariants

$$J_T = f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5$$

$$\begin{aligned}
f_1 &= \mathfrak{Re}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right)\\
f_2 &= \mathfrak{Im}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right)\\
f_3 &= \mathfrak{Re}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - \left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)^2 - \left(\vec{\beta}_k\cdot\vec{\beta}_l\right)^2\\
&\quad +2\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right)\\
f_4 &= \mathfrak{Im}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\beta}_k\cdot\vec{\beta}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right)
\end{aligned}$$

with $a_0 = \mathrm{Tr}^2\left(\hat{\mathbf{m}}_O\right)/16\det(\hat{\mathbf{O}})$ and $b_0 = \left[\mathrm{Tr}^2\left(\hat{\mathbf{m}}_O\right) - \mathrm{Tr}\left(\hat{\mathbf{m}}_O^2\right)\right]/4\det(\hat{\mathbf{O}})$

$(\alpha_k)_m = \mathfrak{Re}\left[\langle m|\varphi_k(T)\rangle\right]$, $(\beta_k)_m = \mathfrak{Im}\left[\langle m|\varphi_k(T)\rangle\right]$, $m = 1, \ldots, \dim(\mathcal{H})$

problem:

functional based on local invariants

$$J_T = f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5$$

$$\begin{aligned}
f_1 &= \mathfrak{Re}\left[a_0 \det(\hat{\mathbf{U}}\right] - \frac{1}{16}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right)\\
f_2 &= \mathfrak{Im}\left[a_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{16}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right)\\
f_3 &= \mathfrak{Re}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} \vec{\alpha}_k^2\vec{\alpha}_l^2 + \vec{\beta}_k^2\vec{\beta}_l^2 - 2\vec{\alpha}_k^2\vec{\beta}_l^2 - 4\left(\vec{\alpha}_k\cdot\vec{\beta}_k\right)\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - \left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)^2 - \left(\vec{\beta}_k\cdot\vec{\beta}_l\right)^2\\
&\quad +2\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\beta}_k\cdot\vec{\beta}_l\right)\\
f_4 &= \mathfrak{Im}\left[b_0 \det(\hat{\mathbf{U}})\right] - \frac{1}{4}\sum_{k,l} 4\vec{\alpha}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\vec{\beta}_k^2\left(\vec{\alpha}_l\cdot\vec{\beta}_l\right) - 4\left(\vec{\alpha}_k\cdot\vec{\alpha}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right) + 4\left(\vec{\beta}_k\cdot\vec{\beta}_l\right)\left(\vec{\alpha}_k\cdot\vec{\beta}_l\right)
\end{aligned}$$

with $a_0 = \mathrm{Tr}^2\left(\hat{\mathbf{m}}_O\right)/16\det(\hat{\mathbf{O}})$ and $b_0 = \left[\mathrm{Tr}^2\left(\hat{\mathbf{m}}_O\right) - \mathrm{Tr}\left(\hat{\mathbf{m}}_O^2\right)\right]/4\det(\hat{\mathbf{O}})$

$(\alpha_k)_m = \mathfrak{Re}\left[\langle m|\varphi_k(T)\rangle\right]$, $(\beta_k)_m = \mathfrak{Im}\left[\langle m|\varphi_k(T)\rangle\right]$, $m = 1, \ldots, \dim(\mathcal{H})$

problem: J_T is 8th degree polynomial in $\{\vec{\alpha}_k, \vec{\beta}_k\}$, resp. $\{|\varphi_k\rangle\}$ $\curvearrowright$ non-convex

optimization of non-convex functionals

(old) functional to obtain Ô

$$J_T = -\frac{\lambda_0}{N}\mathfrak{Re}\left[\mathrm{Tr}\left\{\hat{\mathbf{O}}^{+}\hat{\mathbf{P}}_N\hat{\mathbf{U}}(T,0;\varepsilon)\hat{\mathbf{P}}_N\right\}\right]$$

quadratic

(new) functional to obtain [Ô]

$$J_T = \Delta g_1{}^2 + \Delta g_2{}^2 + \Delta g_3{}^2$$

non-convex

for non-convex functionals

- local optima may exist
- how to ensure monotonic convergence?

2nd order
Krotov algorithm

basic concept

- ingredients:
    - final-time target $J_T[\varphi_T, \varphi_T^*]$
    - time-dep. targets / costs $g_a[\epsilon] + g_b[\varphi(t), \varphi^*(t)]$
    - equations of motion

$$i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle = \hat{\mathbf{H}}(t)|\varphi(t)\rangle \quad |\varphi(t_0)\rangle = |\varphi_0\rangle$$

Sklarz & Tannor, PRA 66, 053619 (2002)

basic concept

- ingredients:
  - final-time target $J_T[\varphi_T, \varphi_T^*]$
  - time-dep. targets / costs $g_a[\epsilon] + g_b[\varphi(t), \varphi^*(t)]$
  - equations of motion

$$i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle = \hat{\mathbf{H}}(t)|\varphi(t)\rangle \quad |\varphi(t_0)\rangle = |\varphi_0\rangle$$

- construction of **auxiliary functional** L

$$L[\varphi, \varphi^*, \epsilon, \Phi] = J[\varphi, \varphi^*, \epsilon]$$

choose **arbitrary scalar potential** $\Phi[\varphi, \varphi^*, t]$ such that

$$L[\varphi^i, \varphi^{*,i}, \epsilon^i, \Phi] \geq L[\varphi^{i+1}, \varphi^{*,i+1}, \epsilon^{i+1}, \Phi]$$

$\longrightarrow$ *building in monotonic convergence*

Sklarz & Tannor, PRA 66, 053619 (2002)

auxiliary functional L

$$\begin{aligned} L[\varphi, \varphi^*, \epsilon, \Phi] &= G[\varphi(T), \varphi^*(T)] - \Phi[\varphi(0), \varphi^*(0), 0] \\ &\quad - \int_0^T R[\varphi(t), \varphi^*(t), \epsilon(t), t] dt \end{aligned}$$

final-time contribution:

$$G\left[\varphi(T), \varphi^*(T)\right] = J_T[\varphi(T), \varphi^*(T)] + \Phi[\varphi(T), \varphi^*(T), T]$$

auxiliary functional L

$$\begin{aligned} L[\varphi,\varphi^*,\epsilon,\Phi] &= G[\varphi(T),\varphi^*(T)] - \Phi[\varphi(0),\varphi^*(0),0] \\ &\quad - \int_0^T R[\varphi(t),\varphi^*(t),\epsilon(t),t]dt \end{aligned}$$

final-time contribution:

$$G\,[\varphi(T),\varphi^*(T)] = J_T[\varphi(T),\varphi^*(T)] + \Phi[\varphi(T),\varphi^*(T),T]$$

intermediate-time contribution:

$$\begin{aligned} R\,[\varphi(t),\varphi^*(t),\epsilon(t),t] &= -\Big(g_a[\epsilon(t)] + g_b[\varphi(t),\varphi^*(t)]\Big) \\ &\quad + \frac{\partial\Phi}{\partial t} + \sum_{k=1}^{N}\Big[\nabla_{\varphi_k}\Phi\cdot f_k[\varphi,\varphi^*,\epsilon,t] \\ &\quad + \nabla_{\varphi_k^*}\Phi\cdot {f_k}^*[\varphi,\varphi^*,\epsilon,t]\Big] \end{aligned}$$

auxiliary functional L

$$\begin{aligned}L[\varphi,\varphi^*,\epsilon,\Phi] &= G[\varphi(T),\varphi^*(T)] - \Phi[\varphi(0),\varphi^*(0),0] \\ &\quad - \int_0^T R[\varphi(t),\varphi^*(t),\epsilon(t),t]dt\end{aligned}$$

final-time contribution:

$$G\left[\varphi(T),\varphi^*(T)\right] = J_T[\varphi(T),\varphi^*(T)] + \Phi[\varphi(T),\varphi^*(T),T]$$

intermediate-time contribution:

$$\begin{aligned}R\left[\varphi(t),\varphi^*(t),\epsilon(t),t\right] &= -\left(g_a[\epsilon(t)] + g_b[\varphi(t),\varphi^*(t)]\right) \\ &\quad + \frac{\partial\Phi}{\partial t} + \sum_{k=1}^{N}\Big[\nabla_{\varphi_k}\Phi\cdot f_k[\varphi,\varphi^*,\epsilon,t] \\ &\quad + \nabla_{\varphi_k^*}\Phi\cdot f_k{}^*[\varphi,\varphi^*,\epsilon,t]\Big]\end{aligned}$$

the choice of $\Phi[\varphi(t),\varphi^*(t),t]$ completely determines G, R, L

central idea of Krotov's method

- goal: minimization of L, resp. J_T

two-step solution

1. we need an extremum in $\vec{\varphi}^i$ $\curvearrowright$

$$\nabla_{\vec{\varphi}} G\Big|_{\vec{\varphi}^{(i)}} = 0 \quad \text{and} \quad \nabla_{\vec{\varphi}} R\Big|_{\vec{\varphi}^{(i)}} = 0$$

$\curvearrowright$ equation for backward propagation

$$\begin{aligned} \frac{d}{dt}\vec{\chi}(t) &= -\mathbf{J}^T(t)\cdot\vec{\chi}(t) + \nabla_{\vec{\varphi}}\, g\left(t, \vec{\varphi}^{(i)}, \epsilon^{(i)}\right) \\ \vec{\chi}(T) &= -\nabla_{\vec{\varphi}} J_T\left(\vec{\varphi}^{(i)}(T)\right) \end{aligned}$$

central idea of Krotov's method

2. we want a minimum of L, i.e. minimum of G & maximum of R
   but L is changed by both changes in $\vec{\varphi}$
   and changes in ϵ

Krotov's solution

Konnov & Krotov, Automation Remote Control 60, 10 (1999)

central idea of Krotov's method

2. we want a minimum of L, i.e. minimum of G & maximum of R
but L is changed by both changes in $\vec{\varphi}$
and changes in ϵ

Krotov's solution

(i) choose Φ at the extremum, $\vec{\varphi}^{i}$, such that it is the **worst** possible choice with respect to any change in the states $\curvearrowright$ maximize L when going from $\vec{\varphi}^{i}$ to $\vec{\varphi}^{i+1}$ for fixed ϵ^{i}

(ii) then any change in the field from ϵ^{i} to ϵ^{i+1} will lead to a minimization of L

$$\epsilon^{(i+1)}(t) = \arg\max_{\epsilon(t)} R\left(\vec{\varphi}(t)^{(i+1)}, \epsilon(t), t\right) \quad \text{or}$$

$$\frac{\partial R}{\partial \epsilon}\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) = 0 \quad , \quad \frac{\partial^2 R}{\partial \epsilon^2}\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) < 0$$

Konnov & Krotov, Automation Remote Control 60, 10 (1999)

central idea of Krotov's method

2. we want a minimum of L, i.e. minimum of G & maximum of R **but** L is changed by both changes in $\vec{\varphi}$ and changes in ϵ

Krotov's solution

(i) choose **Φ** at the extremum, $\vec{\varphi}^i$, such that it is the **worst** possible choice with respect to any change in the states $\curvearrowright$ maximize L when going from $\vec{\varphi}^i$ to $\vec{\varphi}^{i+1}$ for fixed ϵ^i

(ii) then any change in the field from ϵ^i to ϵ^{i+1} will lead to a minimization of L

$$\epsilon^{(i+1)}(t) = \arg\max_{\epsilon(t)} R\left(\vec{\varphi}(t)^{(i+1)}, \epsilon(t), t\right) \quad \text{or}$$

$$\frac{\partial R}{\partial \epsilon}\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) = 0 \quad , \quad \frac{\partial^2 R}{\partial \epsilon^2}\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) < 0$$

Konnov & Krotov, Automation Remote Control 60, 10 (1999)

central idea of Krotov's method

Krotov's solution

(i) optimization with respect to change in states is translated into construction of Φ at the extremum $\vec{\varphi}^i$

(ii) *convergence* for step in field, $\epsilon^{(i)} \longrightarrow \epsilon^{(i+1)}$, assured *globally* for R by

$$\frac{\partial R}{\partial \epsilon}\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) = 0$$

$$\Delta_\epsilon R = R\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) - R\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i)}, t\right) \geq 0$$

a *global optimum* would be found, if we could actually implement

$$\epsilon^{(i+1)}(t) = \arg\max_{\epsilon(t)} R\left(\vec{\varphi}(t)^{(i+1)}, \epsilon(t), t\right)$$

Krotov's step (i)

second order construction of Φ

Krotov's ansatz

construct Φ to second order in the states $|\varphi_k^{(i+1)}\rangle$

$$\Phi(t, \varphi^{(i+1)}, \varphi^{*,(i+1)}) = \frac{1}{2}\sum_{k=1}^{N}\left(\langle\chi_k^{(i)}|\varphi_k^{(i+1)}\rangle + \langle\varphi_k^{(i+1)}|\chi_k^{(i)}\rangle\right) + \frac{1}{2}\sum_{k,l=1}^{N}\langle\varphi_k^{(i+1)} - \varphi_k^{(i)}|\hat{\boldsymbol{\sigma}}_{kl}(t)|\varphi_l^{(i+1)} - \varphi_l^{(i)}\rangle$$

choose $\hat{\boldsymbol{\sigma}}_{kl}(t)$ such that

maximum condition for G and minimum condition for R are fulfilled

- Krotov: constructive proof for global conditions

$$\hat{\boldsymbol{\sigma}}(t) = \left(\alpha\left(e^{\gamma(T-t)} - 1\right) + \beta\right)\cdot \mathbb{1} \equiv \sigma(t)\cdot\mathbb{1}$$

Konnov & Krotov, Automation Remote Control 60, 10 (1999)

Krotov's ansatz

construct Φ to second order in the states $|\varphi_k^{(i+1)}\rangle$

$$\begin{aligned}\Phi(t,\varphi^{(i+1)},\varphi^{*,(i+1)}) \;&=\; \frac{1}{2}\sum_{k=1}^{N}\left(\langle\chi_k^{(i)}|\varphi_k^{(i+1)}\rangle + \langle\varphi_k^{(i+1)}|\chi_k^{(i)}\rangle\right)\\ &\quad+\frac{1}{2}\sum_{k,l=1}^{N}\langle\varphi_k^{(i+1)}-\varphi_k^{(i)}|\hat{\boldsymbol{\sigma}}_{kl}(t)|\varphi_l^{(i+1)}-\varphi_l^{(i)}\rangle\end{aligned}$$

choose $\hat{\boldsymbol{\sigma}}_{kl}(t)$ such that

maximum condition for G and minimum condition for R are fulfilled

- Krotov: constructive proof for global conditions

$$\hat{\boldsymbol{\sigma}}(t) = \left(\alpha\left(e^{\gamma(T-t)}-1\right)+\beta\right)\cdot\mathbb{1} \equiv \sigma(t)\cdot\mathbb{1}$$

Konnov & Krotov, Automation Remote Control 60, 10 (1999)

construction of $\hat{\sigma}_{kl}(t)$

- can be done locally or globally
- Sklarz/Tannor's discussion local
  (but results coincide with global derivation due to choice of J_T)
- Krotov: constructive proof for global conditions
- derivation for global conditions leads to much simpler solution for fourth-degree tensor $\hat{\boldsymbol{\sigma}}$

$$\hat{\boldsymbol{\sigma}}(t) = \left(\alpha\left(e^{\gamma(T-t)} - 1\right) + \beta\right) \cdot \mathbb{1} \equiv \sigma\left(t\right) \cdot \mathbb{1}$$

Krotov's proof

main idea: assure that nothing goes wrong
for very large & very small $\vec{\varphi}$ and very large ϵ

If:

1. The right-hand side of the equation of motion, $f\left(t, \vec{\varphi}, \epsilon\right)$, is bounded. Specifically, for large values of the state vector, $\vec{\varphi}$, the right-hand side of the equations of motion does not grow faster than quadratically with respect to $\|\vec{\varphi}\|$ for all t and possible fields ϵ.
2. The Jacobian of the right-hand side of the equations of motion, $\mathbf{J}$, is bounded for any time t, field ϵ and state vector $\vec{\varphi}$.
3. The functionals $J_T(\vec{\varphi})$ and $g\left(\epsilon, \vec{\varphi}, t\right)$ are twice differentiable and bounded. In particular, for large values of the state vector $\vec{\varphi}$, the functionals J_T and g do not grow faster than quadratically with respect to $\|\vec{\varphi}\|$.

then

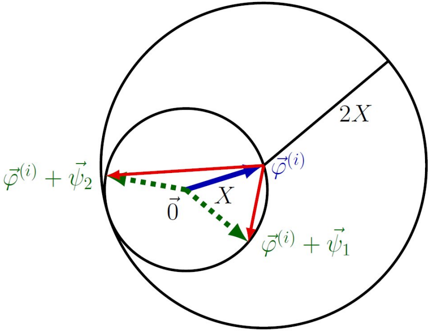
$$\hat{\boldsymbol{\sigma}}(t) = \left(\alpha\left(e^{\gamma(T-t)} - 1\right) + \beta\right) \cdot \mathbb{1} \equiv \sigma\left(t\right) \cdot \mathbb{1}$$

quantum control

- state vectors $\vec{\varphi}^{(i)}$, $\vec{\varphi}^{(i+1)}$ inherently bounded
- $\curvearrowright$ boundedness conditions already guaranteed if $f\,(t, \vec{\varphi}, \epsilon)$, $\mathbf{J}$, $J_T(\vec{\varphi})$ and $g\,(\epsilon, \vec{\varphi}, t)$ *regular*
- change in states $\in$ compact subset of $\mathbb{R}^{2NM}$

quantum control

- state vectors $\vec{\varphi}^{(i)}$, $\vec{\varphi}^{(i+1)}$ inherently bounded
- $\curvearrowright$ boundedness conditions already guaranteed if $f\,(t, \vec{\varphi}, \epsilon)$, $\mathbf{J}$, $J_T(\vec{\varphi})$ and $g\,(\epsilon, \vec{\varphi}, t)$ *regular*
- change in states $\in$ compact subset of $\mathbb{R}^{2NM}$



fulfilling $\Delta G(\vec{\psi}) \leq 0$

$$\Delta G\left(\vec{\psi}\right) \;=\; \left(\vec{\chi}(T)\cdot\vec{\psi}\right) + \frac{1}{2}\sigma(T)\cdot\left(\vec{\psi}\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)$$

for $\vec{\psi} = \vec{0}$: $\Delta G\left(\vec{0}\right) = 0$

for $\vec{\psi} \neq \vec{0}$:

fulfilling $\Delta G(\vec{\psi}) \leq 0$

$$\Delta G\left(\vec{\psi}\right) = \left(\vec{\chi}(T)\cdot\vec{\psi}\right) + \frac{1}{2}\sigma(T)\cdot\left(\vec{\psi}\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)$$

for $\vec{\psi} = \vec{0}$: $\Delta G\left(\vec{0}\right) = 0$

for $\vec{\psi} \neq \vec{0}$:

$$\Delta G\left(\vec{\psi}\right) = \left(\vec{\psi}\cdot\vec{\psi}\right)\left[\frac{1}{2}\sigma(T) + \frac{\left(\vec{\chi}(T)\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)}{\left(\vec{\psi}\cdot\vec{\psi}\right)}\right]$$

$$A = \sup_{\vec{\psi}} \frac{\left(\vec{\chi}_T\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)}{\left(\vec{\psi}\cdot\vec{\psi}\right)}$$

fulfilling $\Delta G(\vec{\psi}) \leq 0$

$$\Delta G\left(\vec{\psi}\right) = \left(\vec{\chi}(T)\cdot\vec{\psi}\right) + \frac{1}{2}\sigma(T)\cdot\left(\vec{\psi}\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)$$

for $\vec{\psi} = \vec{0}$: $\Delta G\left(\vec{0}\right) = 0$

for $\vec{\psi} \neq \vec{0}$:

$$\Delta G\left(\vec{\psi}\right) = \left(\vec{\psi}\cdot\vec{\psi}\right)\left[\frac{1}{2}\sigma(T) + \frac{\left(\vec{\chi}(T)\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)}{\left(\vec{\psi}\cdot\vec{\psi}\right)}\right]$$

$$A = \sup_{\vec{\psi}} \frac{\left(\vec{\chi}_T\cdot\vec{\psi}\right) + J_T\left(\vec{\varphi}(T)^{(i)} + \vec{\psi}\right) - J_T\left(\vec{\varphi}(T)^{(i)}\right)}{\left(\vec{\psi}\cdot\vec{\psi}\right)}$$

$$\boxed{\sigma(T) < -2A}$$

fulfilling $\Delta R(\vec{\psi}) \geq 0$

$$\begin{aligned} \Delta R\left(\vec{\psi}(t), t\right) &= \left(\dot{\vec{\chi}}(t) \cdot \vec{\psi}(t)\right) + \frac{1}{2}\dot{\sigma}(t) \cdot \left(\vec{\psi}(t) \cdot \vec{\psi}(t)\right) \\ &\quad + \left(\left(\vec{\chi}(t) + \sigma(t)\vec{\psi}(t)\right) \cdot \Delta\vec{f}\left(\vec{\psi}(t), t\right)\right) - \Delta g\left(\vec{\psi}(t), t\right) \end{aligned}$$

for $\vec{\psi} = \vec{0}$: $\Delta R\left(\vec{0}, t\right) = 0\,\forall t$

for $\vec{\psi} \neq \vec{0}$:

$$\frac{1}{2}\dot{\sigma}\left(t\right) - \left|\sigma\left(t\right) \cdot B\right| + C > 0$$

fulfilling $\Delta R(\vec{\psi}) \geq 0$

$$\begin{aligned}\Delta R\left(\vec{\psi}(t),t\right) &= \left(\dot{\vec{\chi}}(t)\cdot\vec{\psi}(t)\right) + \frac{1}{2}\dot{\sigma}(t)\cdot\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)\\ &\quad + \left(\left(\vec{\chi}(t)+\sigma(t)\vec{\psi}(t)\right)\cdot\Delta\vec{f}\left(\vec{\psi}(t),t\right)\right) - \Delta g\left(\vec{\psi}(t),t\right)\end{aligned}$$

for $\vec{\psi}=\vec{0}$: $\Delta R\left(\vec{0},t\right)=0\,\forall t$

for $\vec{\psi}\neq\vec{0}$:

$$\begin{aligned}\Delta R\left(\vec{\psi}(t),t\right) &= \left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)\\ &\quad \left[\frac{1}{2}\dot{\sigma}(t)+\sigma(t)\frac{\vec{\psi}(t)\cdot\Delta\vec{f}}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}+\frac{\dot{\vec{\chi}}(t)\cdot\vec{\psi}(t)+\vec{\chi}(t)\cdot\Delta\vec{f}-\Delta g}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}\right]\end{aligned}$$

$$\begin{aligned}B &= \sup_{\vec{\psi}(t)\in\mathbb{R}^{2NM};t\in[0,T]}\left|\frac{\vec{\psi}(t)\cdot\Delta\vec{f}}{\vec{\psi}(t)\cdot\vec{\psi}(t)}\right|\\ C &= \inf_{\vec{\psi}(t)\in\mathbb{R}^{2NM};t\in[0,T]}\frac{\dot{\vec{\chi}}(t)\cdot\vec{\psi}(t)+\vec{\chi}(t)\cdot\Delta\vec{f}-\Delta g}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}\end{aligned}$$

fulfilling $\Delta R(\vec{\psi}) \geq 0$

$$\begin{aligned}\Delta R\left(\vec{\psi}(t), t\right) &= \left(\dot{\vec{\chi}}(t)\cdot\vec{\psi}(t)\right) + \frac{1}{2}\dot{\sigma}(t)\cdot\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)\\ &\quad + \left(\left(\vec{\chi}(t)+\sigma(t)\vec{\psi}(t)\right)\cdot\Delta\vec{f}\left(\vec{\psi}(t), t\right)\right) - \Delta g\left(\vec{\psi}(t), t\right)\end{aligned}$$

for $\vec{\psi} = \vec{0}$: $\Delta R\left(\vec{0}, t\right) = 0\,\forall t$

for $\vec{\psi} \neq \vec{0}$:

$$\begin{aligned} B &= \sup_{\vec{\psi}(t)\in\mathbb{R}^{2NM};t\in[0,T]} \left|\frac{\vec{\psi}(t)\cdot\Delta\vec{f}}{\vec{\psi}(t)\cdot\vec{\psi}(t)}\right| \\ C &= \inf_{\vec{\psi}(t)\in\mathbb{R}^{2NM};t\in[0,T]} \frac{\dot{\vec{\chi}}(t)\cdot\vec{\psi}(t)+\vec{\chi}(t)\cdot\Delta\vec{f}-\Delta g}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}\end{aligned}$$

$$\frac{1}{2}\dot{\sigma}(t) - \left|\sigma(t)\cdot B\right| + C > 0$$

maximizing L wrt $\Delta\vec{\varphi}$

$$\begin{aligned} \sigma(T) &< -2A \\ \frac{1}{2}\dot{\sigma}(t) - |\sigma(t)\cdot B| + C &> 0 \end{aligned}$$

one solution

$$\sigma(t) = e^{\bar{B}(T-t)}\left(\frac{\bar{C}}{\bar{B}} - \bar{A}\right) - \frac{\bar{C}}{\bar{B}}$$

with $\bar{B} = 2B + \delta$, $\bar{C} = \min(-\delta, 2C - \delta)$ and $\bar{A} = \max(\varepsilon, 2A + \varepsilon)$

or more generally

$$\sigma(t) = \alpha\left(e^{\gamma(T-t)} - 1\right) + \beta$$

how to get A, B, C ?

A, B, C are Taylor expansions of certain quantities starting at the first or second order

$\curvearrowright$ **estimate the remainder (Lagrange's form)**

$$W(\vec{\varphi}) = \sum_{|\alpha|\leq n-1} \frac{1}{\alpha!} \left(\partial^{\alpha} W\right) \left(\vec{\varphi}^{(i)}\right) \cdot \vec{\psi}^{\alpha} + \mathcal{R}_{\vec{\varphi}^{(i)},n} \left(\vec{\psi}\right)$$

how to get A, B, C ?

A, B, C are Taylor expansions of certain quantities starting at the first or second order

$\curvearrowright$ estimate the remainder (Lagrange's form)

$$W(\vec{\varphi}) = \sum_{|\alpha| \leq n-1} \frac{1}{\alpha!} \left(\partial^\alpha W\right) \left(\vec{\varphi}^{(i)}\right) \cdot \vec{\psi}^{\alpha} + \mathcal{R}_{\vec{\varphi}^{(i)},n} \left(\vec{\psi}\right)$$

$$\mathcal{R}_{\vec{\varphi}^{(i)},n} \left(\vec{\psi}\right) \leq \frac{1}{\alpha!} M_n^W \left(\vec{\varphi}^{(i)}\right) \vec{\psi}^{\alpha} \, , |\alpha| = n$$

$$M_n^W \left(\vec{\varphi}^{(i)}\right) = \sup_{\vec{\psi}, |\alpha| = n} \left| \partial^\alpha W \left(\vec{\varphi}^{(i)} + \vec{\psi}\right) \right| \, ,$$

how to get A, B, C ?

A, B, C are Taylor expansions of certain quantities starting at the first or second order

↷ estimate the remainder (Lagrange's form)

$$W(\vec{\varphi}) = \sum_{|\alpha| \leq n-1} \frac{1}{\alpha!} \left(\partial^\alpha W\right) \left(\vec{\varphi}^{(i)}\right) \cdot \vec{\psi}^\alpha + \mathcal{R}_{\vec{\varphi}^{(i)},n} \left(\vec{\psi}\right)$$

$$\mathcal{R}_{\vec{\varphi}^{(i)},n} \left(\vec{\psi}\right) \leq \frac{1}{\alpha!} M_n^W \left(\vec{\varphi}^{(i)}\right) \vec{\psi}^\alpha \ , |\alpha| = n$$

$$M_n^W \left(\vec{\varphi}^{(i)}\right) = \sup_{\vec{\psi}, |\alpha| = n} \left| \partial^\alpha W \left(\vec{\varphi}^{(i)} + \vec{\psi}\right) \right| \ ,$$

estimate that is independent of $\vec{\varphi}^{(i)}$

$$M_n^W = \sup_{\vec{\varphi}^{(i)} \in \mathbb{X}} M_n \left(\vec{\varphi}^{(i)}\right)$$

estimate of A

$$A = \sup_{\vec{\psi}} \frac{J_{T,2}}{\left(\vec{\psi}\cdot\vec{\psi}\right)}\,.$$

estimate $J_{T,2}$ by its Lagrange remainder:

$$A \leq \frac{1}{2} M_2^{J_T} = \frac{1}{2} \sup_{\vec{\psi},|\alpha|=2} \left|\partial^\alpha J_T\left(\vec{\psi}\right)\right|$$

estimate of A

$$A = \sup_{\vec{\psi}} \frac{J_{T,2}}{\left(\vec{\psi} \cdot \vec{\psi}\right)} .$$

estimate $J_{T,2}$ by its Lagrange remainder:

$$A \leq \frac{1}{2} M_2^{J_T} = \frac{1}{2} \sup_{\vec{\psi}, |\alpha| = 2} \left| \partial^{\alpha} J_T \left(\vec{\psi}\right) \right|$$

for functionals J_T that are linear or quadratic in $\vec{\varphi}$

$A = 0$

estimate of B

$$
\begin{aligned}
B \;\leq\;& X \sup_{\vec{\psi}} \left\| \partial \omega \left(\vec{\psi} \right) \right\| \\
&+ \sup_{\vec{\psi}(t);\, t \in [0,T]} \left\| \omega \left(\vec{\psi}(t), \epsilon^{(i)}, t \right) \right\|_{mat}
\end{aligned}
$$

estimate of B

$$\begin{aligned} B &\leq X \sup_{\vec{\psi}} \left\| \partial \omega \left(\vec{\psi} \right) \right\| \\ &\quad + \sup_{\vec{\psi}(t); t \in [0,T]} \left\| \omega \left(\vec{\psi}(t), \epsilon^{(i)}, t \right) \right\|_{mat} \end{aligned}$$

for Hamiltonians that do not depend on the state

$$\begin{aligned} B &= \sup_{\vec{\psi}; t \in [0,T]} \left| \frac{\vec{\psi}(t) \cdot \omega \left(\epsilon^{(i)}, t \right) \cdot \vec{\psi}(t)}{\vec{\psi}(t) \cdot \vec{\psi}(t)} \right| \\ &= \sup_{t \in [0,T]} \left\| \omega \left(\epsilon^{(i)}, t \right) \right\|_{mat} \end{aligned}$$

$\curvearrowright$ for unitary evolution: $B = 0$

$\curvearrowright$ for non-unitary evolution: max. eigenvalue

estimate of C

$$-C \leq \sup_{\vec{\psi}(t);t\in[0,T]} \frac{\vec{\chi}(t)\cdot\omega_1\cdot\vec{\psi}(t)}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)} + \sup_{\vec{\psi}(t);t\in[0,T]} \frac{g_2}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}$$

$$-C \leq \sup_{t\in[0,T]} \left(M_1^{\omega}\cdot\|\vec{\chi}(t)\|\right) + \frac{1}{2}\sup_{\vec{\psi},|\alpha|=2} \left|\partial^{\alpha} g\left(\vec{\psi}\right)\right| .$$

estimate of C

$$-C \leq \sup_{\vec{\psi}(t);t\in[0,T]} \frac{\vec{\chi}(t)\cdot\omega_1\cdot\vec{\psi}(t)}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)} + \sup_{\vec{\psi}(t);t\in[0,T]} \frac{g_2}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}$$

$$-C \leq \sup_{t\in[0,T]} \left(M_1^{\omega}\cdot\|\vec{\chi}(t)\|\right) + \frac{1}{2}\sup_{\vec{\psi},|\alpha|=2}\left|\partial^{\alpha} g\left(\vec{\psi}\right)\right| .$$

for state-independent $\hat{\mathbf{H}}$ and g depending on $\vec{\varphi}$ only up to linear order : $\omega_1 = 0$ and $g_2 = 0$

$$C = 0$$

estimate of C

$$-C \leq \sup_{\vec{\psi}(t);t\in[0,T]} \frac{\vec{\chi}(t)\cdot\omega_1\cdot\vec{\psi}(t)}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)} + \sup_{\vec{\psi}(t);t\in[0,T]} \frac{g_2}{\left(\vec{\psi}(t)\cdot\vec{\psi}(t)\right)}$$

$$-C \leq \sup_{t\in[0,T]} \left(M_1^{\omega}\cdot\|\vec{\chi}(t)\|\right) + \frac{1}{2}\sup_{\vec{\psi},|\alpha|=2} \left|\partial^{\alpha} g\left(\vec{\psi}\right)\right| .$$

for state-independent $\hat{\mathbf{H}}$ and g depending on $\vec{\varphi}$ only up to linear order : $\omega_1 = 0$ and $g_2 = 0$

$$C = 0$$

↷ for certain (!) functionals and EoMs: $A = 0$ & $B = 0$ & $C = 0$
↷ $\sigma(t) = 0$ and the second order contribution to Φ vanishes:

Palao-Kosloff version of Krotov method (Krotov-PK)

(still ensuring monotonic convergence globally)

Krotov's step (ii)

second order construction of ϵ

equation for the field

remember

$$\frac{\partial R}{\partial \epsilon}\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) = 0$$

$$\Delta_\epsilon R = R\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i+1)}, t\right) - R\left(\vec{\varphi}^{(i+1)}, \epsilon^{(i)}, t\right) \geq 0$$

equations of motion in basis set expansion: $2M \times 2M$ matrix

$$\Omega^k = \begin{pmatrix} H^{k,R} & -H^{k,I} \\ H^{k,I} & H^{k,R} \end{pmatrix}$$

first order condition yields:

$$\epsilon^{(i+1)}(t) = \epsilon_{\text{ref}}(t) + \frac{1}{2\lambda_a S(t)} \left\{ \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \chi_{km}^{(i)} \frac{\partial \Omega_{mn}^k}{\partial \epsilon} \varphi_{kn}^{(i+1)} + \sigma(t) \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \Delta\varphi_{km} \frac{\partial \Omega_{mn}^k}{\partial \epsilon} \Delta\varphi_{kn} \right\}$$

we now have an algorithm

that is monotonically convergent

for arbitary

targets/constraints

remark: combine Krotov with BFGS

1. Quasi-Newton algorithms are approximate solutions to an extremization problem using information from the second-order Taylor expansion of the function

2. BFGS is a quasi-Newton algorithm using a rank-two update formula involving only gradient information to approximate the (inverse) Hessian
   for convex functions it is globally and monotonically convergent if supplemented by a line search fulfilling the Wolfe conditions

3. L-BFGS uses only information from the gradients and state vectors of previous steps to solve the memory problem in storing the approximate inverse Hessian
   under certain additional assumptions convergence remains monotonic

remark: combine Krotov with BFGS

1. Quasi-Newton algorithms are approximate solutions to an extremization problem using information from the second-order Taylor expansion of the function
2. BFGS is a quasi-Newton algorithm using a rank-two update formula involving only gradient information to approximate the (inverse) Hessian
   for convex functions it is globally and monotonically convergent if supplemented by a line search fulfilling the Wolfe conditions
3. L-BFGS uses only information from the gradients and state vectors of previous steps to solve the memory problem in storing the approximate inverse Hessian
   under certain additional assumptions convergence remains monotonic

remark: combine Krotov with BFGS

1. Quasi-Newton algorithms are approximate solutions to an extremization problem using information from the second-order Taylor expansion of the function

2. BFGS is a quasi-Newton algorithm using a rank-two update formula involving only gradient information to approximate the (inverse) Hessian
   for convex functions it is globally and monotonically convergent if supplemented by a line search fulfilling the Wolfe conditions

3. L-BFGS uses only information from the gradients and state vectors of previous steps to solve the memory problem in storing the approximate inverse Hessian
   under certain additional assumptions convergence remains monotonic

remark: combine Krotov with BFGS

remember:

$$\Delta\epsilon(t)_{Krotov} = \frac{1}{2\lambda_a S(t)} \left\{ \sum_{k=1}^{N}\sum_{m,n=1}^{2M} \chi_{km}^{(i)} \frac{\partial \Omega_{mn}^{k}}{\partial \epsilon} \varphi_{kn}^{(i+1)} \right. \\ \left. + \sigma(t) \sum_{k=1}^{N}\sum_{m,n=1}^{2M} \Delta\varphi_{km} \frac{\partial \Omega_{mn}^{k}}{\partial \epsilon} \Delta\varphi_{kn} \right\}$$

remark: combine Krotov with BFGS

remember:

$$\Delta\epsilon(t)_{Krotov} = \frac{1}{2\lambda_a S(t)} \left\{ \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \chi_{km}^{(i)} \frac{\partial\Omega_{mn}^{k}}{\partial\epsilon} \varphi_{kn}^{(i+1)} + \sigma(t) \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \Delta\varphi_{km} \frac{\partial\Omega_{mn}^{k}}{\partial\epsilon} \Delta\varphi_{kn} \right\}$$

compare gradient ascent w/ BFGS (for linear dependence of $\hat{\mathbf{H}}$ on ϵ!):

$$\Delta\epsilon(t)_{BFGS} = \hat{\mathbf{B}}^{-1}(t) \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \chi_{km}^{(i)} \frac{\partial\Omega_{mn}^{k}}{\partial\epsilon} \varphi_{kn}^{(i)} \quad \text{with } \hat{\mathbf{B}}(t) \text{ approximated Hessian}$$

remark: combine Krotov with BFGS

remember:

$$\Delta\epsilon(t)_{Krotov} = \frac{1}{2\lambda_a S(t)} \left\{ \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \chi_{km}^{(i)} \frac{\partial\Omega_{mn}^{k}}{\partial\epsilon} \varphi_{kn}^{(i+1)} + \sigma(t) \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \Delta\varphi_{km} \frac{\partial\Omega_{mn}^{k}}{\partial\epsilon} \Delta\varphi_{kn} \right\}$$

compare gradient ascent w/ BFGS (for linear dependence of $\hat{\mathbf{H}}$ on ϵ!):

$$\Delta\epsilon(t)_{BFGS} = \hat{\mathbf{B}}^{-1}(t) \sum_{k=1}^{N} \sum_{m,n=1}^{2M} \chi_{km}^{(i)} \frac{\partial\Omega_{mn}^{k}}{\partial\epsilon} \varphi_{kn}^{(i)} \quad \text{with } \hat{\mathbf{B}}(t) \text{ approximated Hessian}$$

$\curvearrowright$ choose $\lambda_a S(t)$ according to L-BFGS

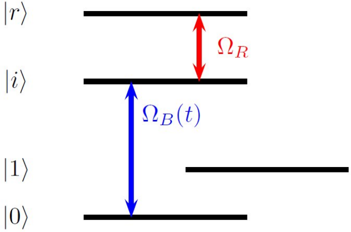
this does not affect the monotonicity that is ensured by Krotov's construction

where are we ? outline !

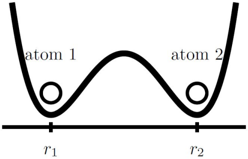
1. introduction: optimal control
   - terminology & basic concepts
   - functionals: how to convey the physics to the algorithm
2. a functional to optimize the entangling power of two-qubit gates
   - local equivalence classes
   - the new functional
   - 2nd order Krotov algorithm
3. application to a Rydberg gate in cold atoms
4. where can we go from here?

Rydberg qubits

one-atom level scheme



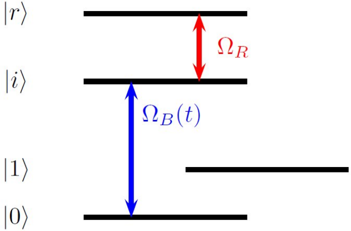
optical tweezers



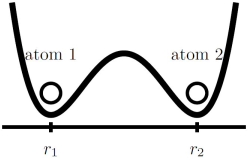
Gaetan et al., Nat. Phys. 5, 115 (2009)

Rydberg qubits

one-atom level scheme



optical tweezers



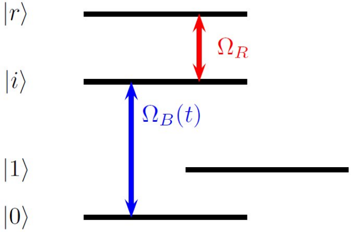
Gaetan et al., Nat. Phys. 5, 115 (2009)

one-atom Hamiltonian

$$\begin{aligned}\hat{\mathbf{H}}^{(1)} = \; & |0\rangle\langle 0| \otimes \left(\hat{\mathbf{T}}_{\hat{\mathbf{r}}} + V^{0}_{trap}(\hat{\mathbf{r}})\right) \\ & + |1\rangle\langle 1| \otimes \left(\hat{\mathbf{T}}_{\hat{\mathbf{r}}} + V^{1}_{trap}(\hat{\mathbf{r}})\right) \\ & + |i\rangle\langle i| \otimes \left(\hat{\mathbf{T}}_{\hat{\mathbf{r}}} + V^{i}_{trap}(\hat{\mathbf{r}})\right) \\ & + \varepsilon_{\mathbf{B}}(t)\left(|0\rangle\langle i| + |i\rangle\langle 0|\right) \otimes \mu(\hat{\mathbf{r}}) \\ & + |r\rangle\langle r| \otimes \left(\hat{\mathbf{T}}_{\hat{\mathbf{r}}} + V^{r}_{trap}(\hat{\mathbf{r}})\right) \\ & + \varepsilon_{\mathbf{R}}\left(|i\rangle\langle r| + |r\rangle\langle i|\right) \otimes \mu(\hat{\mathbf{r}})\end{aligned}$$

Rydberg qubits

one-atom level scheme



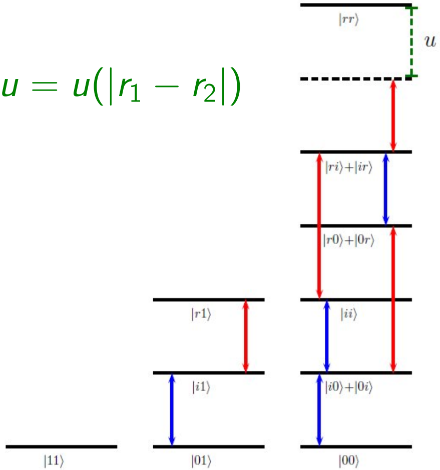
two-atom Hamiltonian

$$\hat{\mathsf{H}} = \hat{\mathsf{H}}_1^{(1)} \otimes \mathbb{1}_{4,2} \otimes \mathbb{1}_{\hat{\mathbf{r}}_2} + \mathbb{1}_{4,1} \otimes \mathbb{1}_{\hat{\mathbf{r}}_1} \otimes \hat{\mathsf{H}}_2^{(1)} + \hat{\mathsf{H}}_{int}^{(1,2)}$$

$$\hat{\mathsf{H}}_{int}^{(1,2)} = |rr\rangle\langle rr| \otimes \frac{u_0}{|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|^3}$$

two-atom level scheme

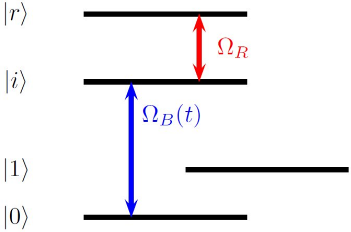
$$u = u(|r_1 - r_2|)$$



Jaksch et al., PRL 85, 2208 (2000)

Rydberg qubits

one-atom level scheme

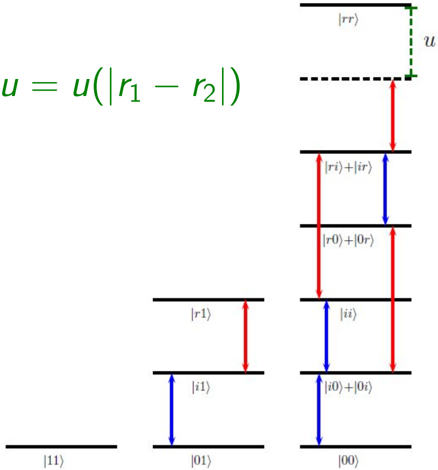


two-atom Hamiltonian

$$\begin{aligned}\hat{\mathbf{H}} &= \hat{\mathbf{H}}_1^{(1)} \otimes \mathbb{1}_{4,2} \otimes \mathbb{1}_{\hat{\mathbf{r}}_2} \\ &\quad + \mathbb{1}_{4,1} \otimes \mathbb{1}_{\hat{\mathbf{r}}_1} \otimes \hat{\mathbf{H}}_2^{(1)} \\ &\quad + \hat{\mathbf{H}}_{int}^{(1,2)} \\ \hat{\mathbf{H}}_{int}^{(1,2)} &= |rr\rangle\langle rr| \otimes \frac{u_0}{|\hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2|^3}\end{aligned}$$

two-atom level scheme

$$u = u(|r_1 - r_2|)$$

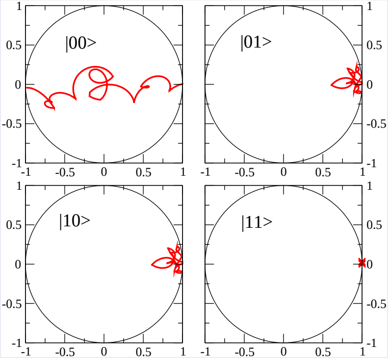


Jaksch et al., PRL 85, 2208 (2000)

controlled Rydberg phasegate

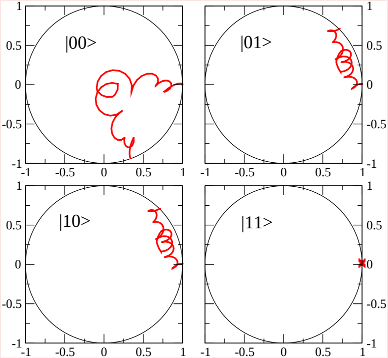
gate time T=20 ns

functional to obtain Ô



$\mathcal{F} = 0.993$

functional to obtain [Ô]

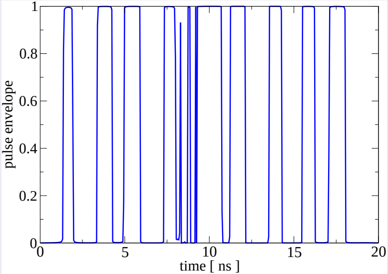


$\mathcal{F} = 0.996$

controlled Rydberg phasegate

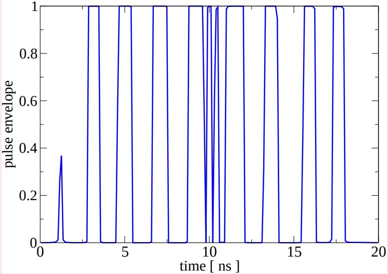
gate time T=20 ns

functional to obtain Ô



$\mathcal{F} = 0.993$

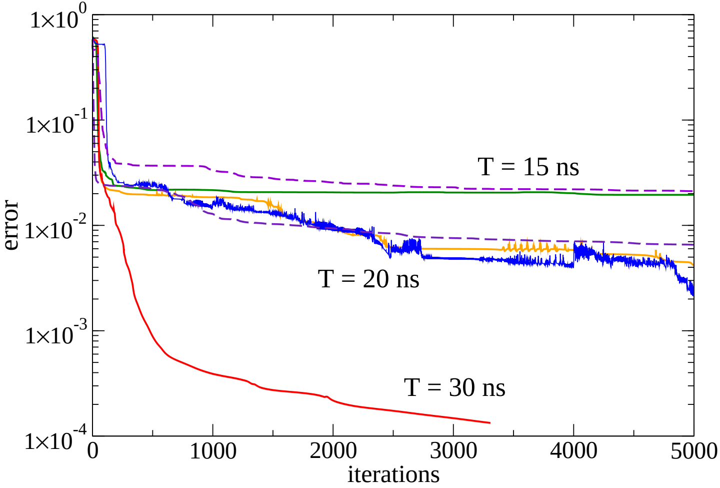
functional to obtain [Ô]



$\mathcal{F} = 0.996$

controlled Rydberg phasegate

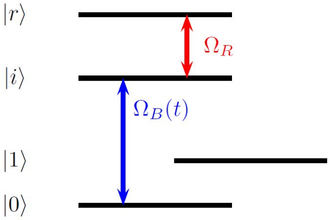
approaching the quantum speed limit

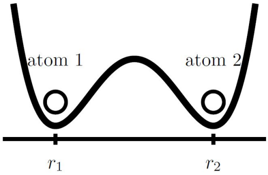


controlled Rydberg phasegate

to be continued

- analyse error sources
- check role of pulse parametrization
- test with Hamiltonian allowing for non-diagonal Ô





summary

optimal control is an extremely versatile tool but

summary

optimal control is an extremely versatile tool but you need to know how to ask questions!

- we derived a new class of optimization functionals suitable for quantum information purposes
- based on geometric classification of entangling operations (Cartan decomposition & representation in Weyl chamber)
- requires optimization algorithm ensuring monotonic convergence – 2nd order Krotov method
- first results for a Rydberg gate encouraging
- full power of approach still needs to be explored

summary

optimal control is an extremely versatile tool but you need to know how to ask questions!

- we derived a new class of optimization functionals suitable for quantum information purposes
- based on geometric classification of entangling operations (Cartan decomposition & representation in Weyl chamber)
- requires optimization algorithm ensuring monotonic convergence – 2nd order Krotov method
- first results for a Rydberg gate encouraging
- full power of approach still needs to be explored

summary

optimal control is an extremely versatile tool but you need to know how to ask questions!

- we derived a new class of optimization functionals suitable for quantum information purposes
- based on geometric classification of entangling operations (Cartan decomposition & representation in Weyl chamber)
- requires optimization algorithm ensuring monotonic convergence – 2nd order Krotov method
- first results for a Rydberg gate encouraging
- full power of approach still needs to be explored

summary

optimal control is an extremely versatile tool but you need to know how to ask questions!

- we derived a new class of optimization functionals suitable for quantum information purposes
- based on geometric classification of entangling operations (Cartan decomposition & representation in Weyl chamber)
- requires optimization algorithm ensuring monotonic convergence – 2nd order Krotov method
- first results for a Rydberg gate encouraging
- full power of approach still needs to be explored

summary

optimal control is an extremely versatile tool but you need to know how to ask questions!

- we derived a new class of optimization functionals suitable for quantum information purposes
- based on geometric classification of entangling operations (Cartan decomposition & representation in Weyl chamber)
- requires optimization algorithm ensuring monotonic convergence – 2nd order Krotov method
- first results for a Rydberg gate encouraging
- full power of approach still needs to be explored

summary

optimal control is an extremely versatile tool but you need to know how to ask questions!

- we derived a new class of optimization functionals suitable for quantum information purposes
- based on geometric classification of entangling operations (Cartan decomposition & representation in Weyl chamber)
- requires optimization algorithm ensuring monotonic convergence – 2nd order Krotov method
- first results for a Rydberg gate encouraging
- full power of approach still needs to be explored

where can we go from here?

where can we go from here?

1. optimize for the complete Weyl chamber, i.e. for an arbitrary perfect entangler
    - problem: no simple inversion of $g_1, g_2, g_3 \longrightarrow c_1, c_2, c_3$
    - solution: define ellipsoid in g-space containing almost all of the Weyl chamber
2. optimize for a specified trajectory in the Weyl chamber
3. . . .

thank you !