Methods for solution of large optimal control problems that bypass open-loop model reduction

Thomas Bewley (bewley@ucsd.edu)
UCSD Flow Control Lab

Methods for solution of large optimal control problems that bypass open-loop model reduction

Thomas Bewley (bewley@ucsd.edu)
UCSD Flow Control & Coordinated Robotics Labs

http://robotics.ucsd.edu

Recent IPAM talks by UCSD Flow Control Lab

http://ipam.ucla.edu

9/8 **Recent work in the control and estimation of flows of engineering significance**

9/26 **Fundamental performance limitations for NS systems**

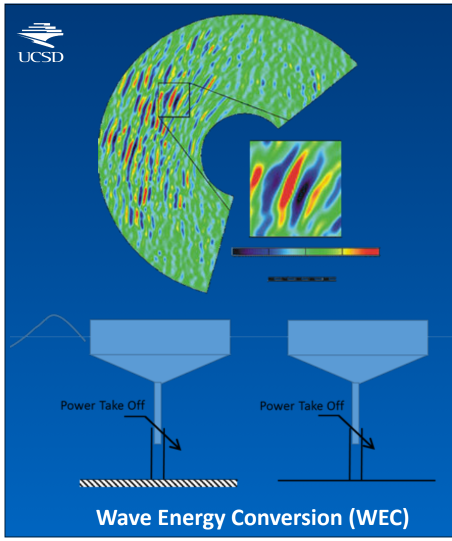
10/22 **Hybrid (Variational / Kalman) ensemble methods for state estimation in NS Systems**

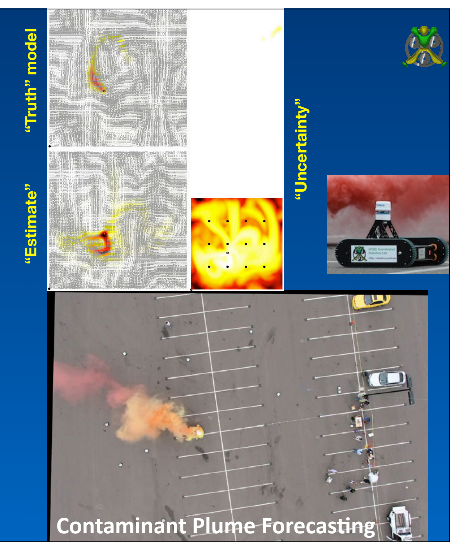
11/4 **On the regularity of spatial convolution kernels for linear feedback control & estimation of perturbations to nearly-parallel flows**

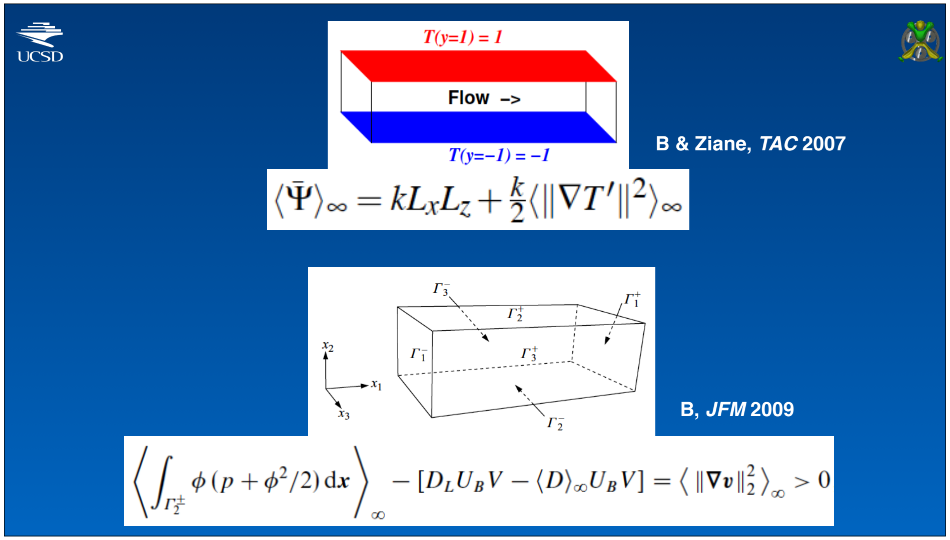
11/13 (also APS-DFD) *Daniele Cavaglieri* **Low-storage implicit/explicit Runge-Kutta schemes for the simulation of turbulent flows**

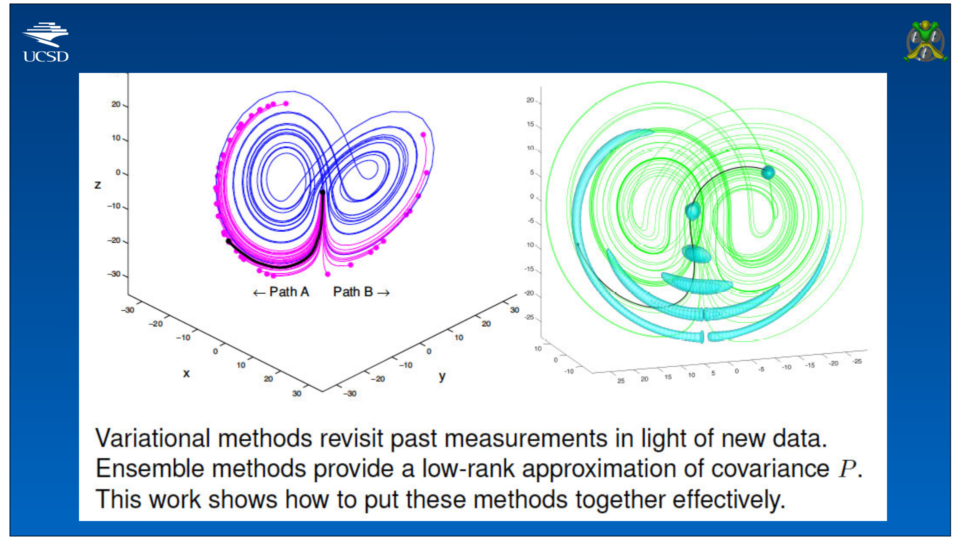
11/13 (also APS-DFD) *Pooriya Beyhaghi* **Uncertainty Quantification of an ergodic process, and its application in simulated-based optimization problems**

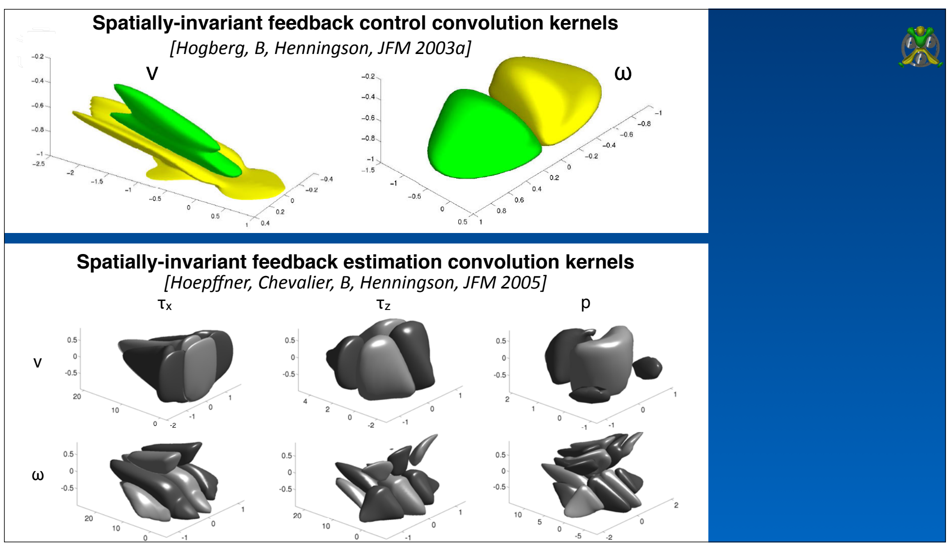
11/13 (also APS-DFD) *Gianluca Meneghello* **Coordinated in-situ observation of developing hurricanes using atmospheric balloons -- a Model Predictive Control approach**

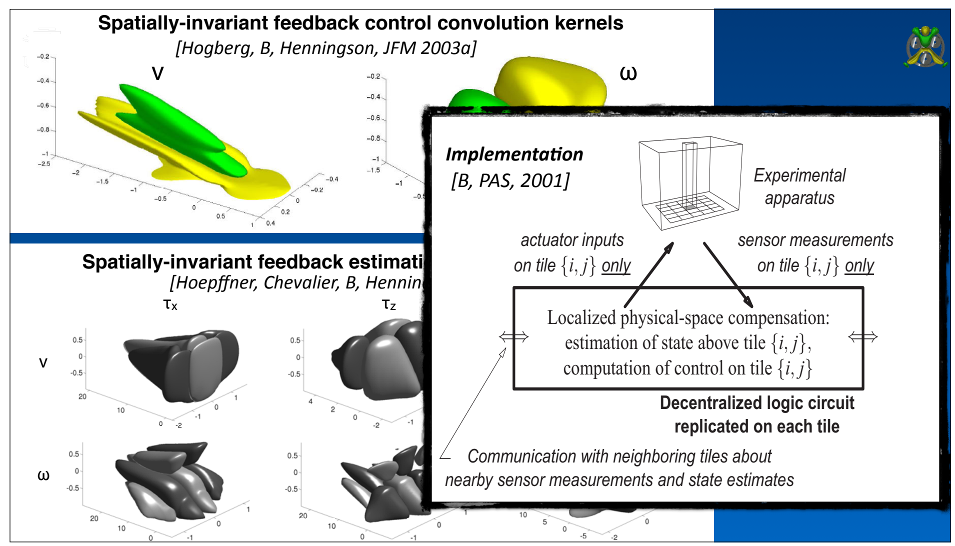





UCSD
$T(y=1) = 1$
Flow ->
$T(y=-1) = -1$
B & Ziane, *TAC* 2007
$\langle \bar{\Psi} \rangle_\infty = k L_x L_z + \frac{k}{2} \langle \| \nabla T' \|^2 \rangle_\infty$
Γ_3^- Γ_2^+ Γ_1^+ Γ_1^- Γ_3^+ Γ_2^- x_1 x_2 x_3
B, *JFM* 2009
$\left\langle \int_{\Gamma_2^\pm} \phi\,(p+\phi^2/2)\,\mathrm{d}x \right\rangle_\infty - [D_L U_B V - \langle D \rangle_\infty U_B V] = \langle \, \|\nabla v\|_2^2 \, \rangle_\infty > 0$

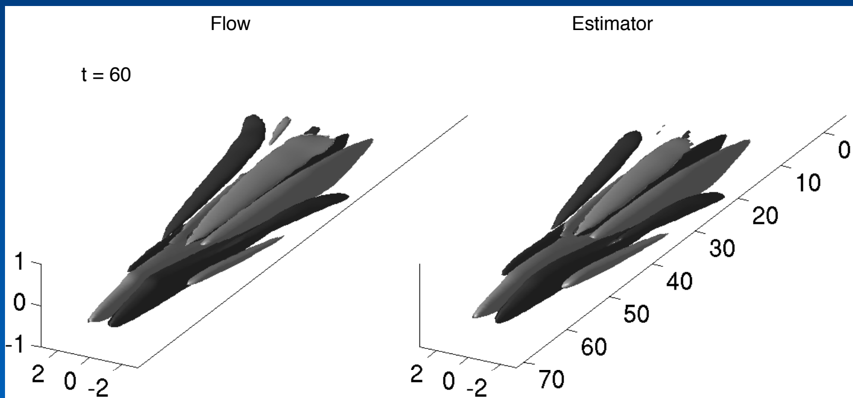

UCSD
← Path A
Path B →
x
y
z
Variational methods revisit past measurements in light of new data.
Ensemble methods provide a low-rank approximation of covariance P.
This work shows how to put these methods together effectively.


Spatially-invariant feedback control convolution kernels
[Hogberg, B, Henningson, JFM 2003a]
v
ω
Spatially-invariant feedback estimation convolution kernels
[Hoepffner, Chevalier, B, Henningson, JFM 2005]
τ_x
τ_z
p
v
ω


Spatially-invariant feedback control convolution kernels
[Hogberg, B, Henningson, JFM 2003a]
v
ω
Spatially-invariant feedback estimati
[Hoepffner, Chevalier, B, Hennin
τ_x
τ_z
v
ω
Implementation
[B, PAS, 2001]
Experimental apparatus
actuator inputs on tile $\{i,j\}$ only
sensor measurements on tile $\{i,j\}$ only
Localized physical-space compensation:
estimation of state above tile $\{i,j\}$,
computation of control on tile $\{i,j\}$
Decentralized logic circuit replicated on each tile
Communication with neighboring tiles about nearby sensor measurements and state estimates

Evolution of small disturbance to state (left) and estimate (right)

[Hoepffner, Chevalier, B, Henningson, JFM 2005]



Positive (light) and negative (dark) iso-surfaces of the streamwise component of velocity. Iso-values at ± 10% of the maximum streamwise velocity of the flow during interval shown.

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x},t) + \mathbf{g}(\mathbf{x},t)$$

$$
\begin{array}{c|cccccc}
0 & 0 & & & & & \\
c_2 & a^{IM}_{2,1} & a^{IM}_{2,2} & & & & \\
c_3 & b_1 & a^{IM}_{3,2} & a^{IM}_{3,3} & & & \\
c_4 & b_1 & b_2 & a^{IM}_{4,3} & a^{IM}_{4,4} & & \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \\
c_s & b_1 & b_2 & \cdots & b_{s-2} & a^{IM}_{s,s-1} & a^{IM}_{s,s} \\
\hline
 & b_1 & b_2 & \cdots & b_{s-2} & b_{s-1} & b_s \\
\hline
 & \hat{b}^{IM}_1 & \hat{b}^{IM}_2 & \cdots & \hat{b}^{IM}_{s-2} & \hat{b}^{IM}_{s-1} & \hat{b}^{IM}_s
\end{array}
\qquad
\begin{array}{c|cccccc}
0 & 0 & & & & & \\
c_2 & a^{EX}_{2,1} & 0 & & & & \\
c_3 & b_1 & a^{EX}_{3,2} & 0 & & & \\
c_4 & b_1 & b_2 & a^{EX}_{4,3} & 0 & & \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \\
c_s & b_1 & b_2 & \cdots & b_{s-2} & a^{EX}_{s,s-1} & 0 \\
\hline
 & b_1 & b_2 & \cdots & b_{s-2} & b_{s-1} & b_s \\
\hline
 & \hat{b}^{EX}_1 & \hat{b}^{EX}_2 & \cdots & \hat{b}^{EX}_{s-2} & \hat{b}^{EX}_{s-1} & \hat{b}^{EX}_s
\end{array}
$$

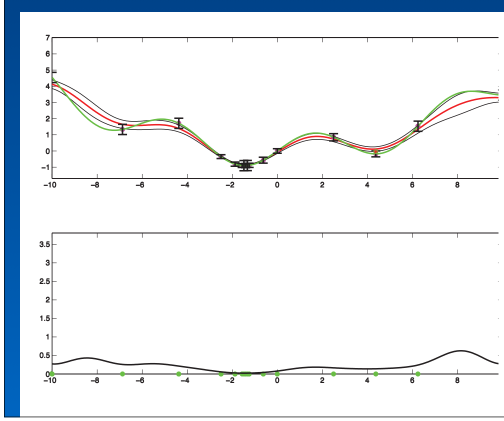
DIRK part **ERK part**

---

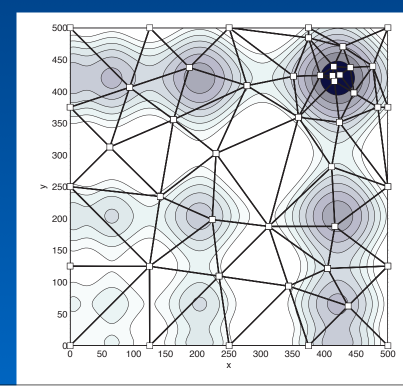
scheme	order	registers	stages (s^{IM}, s^{EX})	stability of DIRK part $[\sigma(z^{IM} \to \infty; z^{EX})]$	Stability of ERK part on negative real axis	truncation error	other properties
IMEXRKCB2	second	[2R]	$(2,3)$	L-stable [0]	$-5.81 \le z^{EX} \le 0$	$A^{(3)} = 0.114$	embedded, SSP ($c = 1.0$)
IMEXRKCB3a	third	[2R]	$(2,3)$	strongly A-stable $[-0.738]$	$-2.51 \le z^{EX} \le 0$	$A^{(4)} = 0.226$	
IMEXRKCB3b			$(3,4)$	strongly A-stable $[-0.732 - 0.366 z^{EX}]$	$-2.21 \le z^{EX} \le 0$	$A^{(4)} = 0.186$	ESDIRK
IMEXRKCB3c				L-stable [0]	$-6.00 \le z^{EX} \le 0$	$A^{(4)} = 0.113$	embedded, SSP ($c = 0.70$)
IMEXRKCB3d					$-2.52 \le z^{EX} \le 0$	$A^{(4)} = 0.207$	embedded, SSP ($c = 0.77$)
IMEXRKCB3e					$-2.79 \le z^{EX} \le 0$	$A^{(4)} = 0.0824$	
IMEXRKCB3f		[3R]	$(4,4)$	L-stable [0]	$-6.00 \le z^{EX} \le 0$	$A^{(4)} = 0.107$	embedded, SO2
IMEXRKCB4	fourth	[3R]	$(6,6)$	L-stable [0]	$-6.04 \le z^{EX} \le 0$	$A^{(5)} = 0.106$	embedded, SO2
CN/RKW3	second	[2R]	$(3,3)$	A-stable $[-1]$	$-2.51 \le z^{EX} \le 0$	$A^{(3)} = 0.0387$	

Cavaglieri & B, JCP 2014

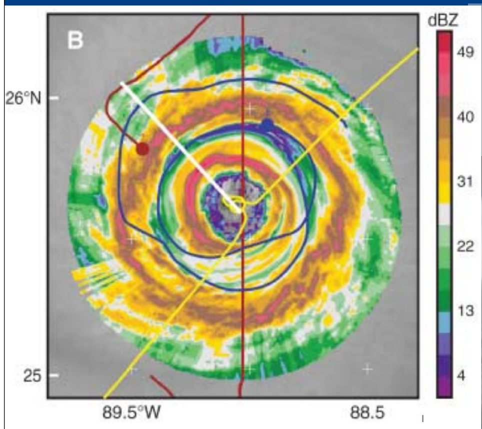
Optimization with “incomplete” function evaluations



Global optimization with surrogate method
(synthetic uncertainty model
built on Delaunay tessellation)



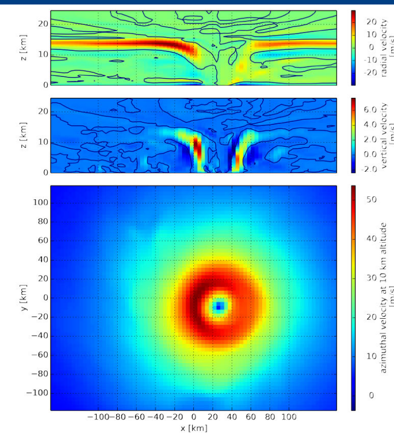
Current: observation of hurricanes via fly-overs by aircraft and dropsondes

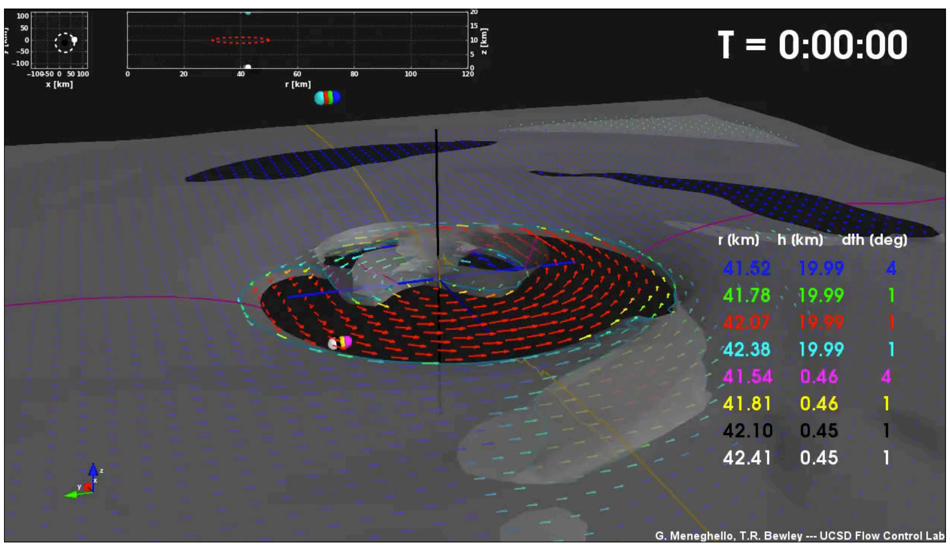


FLIGHT PATHS OVER HURRICANE RITA (HOUZE ET AL., 2007)

New: Underactuated control of sensor balloon distribution in hurricanes

- inflow at lower altitude
- updraft within the eyewall
- outflow at higher altitude (up to 15km)
- azimuthal velocity changes with radius





Methods for solution of large optimal control problems that bypass open-loop model reduction

Thomas Bewley
UCSD Flow Control Lab

joint work with Paolo Luchini, Jan Pralits
Universit`a di Salerno

and Eric Lauga
Formerly UCSD Flow Control Lab, now Cambridge

Adjoint analysis for gradient-based optimization

State equation:

$$\boxed{\begin{array}{ll} E\dot{\mathbf{q}} = N(\mathbf{q},\mathbf{f},\phi,\psi) & \text{on} \quad 0<t<T \\ \mathbf{q}=\mathbf{q}_0 & \text{at} \quad t=0 \end{array}}$$

with: $\mathbf{q}=$ state, $\mathbf{f}=$ external force, $\phi=$ control, $\psi=$ disturbance.

Perturbation equation:

$$\left\{\begin{array}{ll} \mathcal{L}\mathbf{q}' = B_\phi\phi' + B_\psi\psi' & \text{on} \quad 0<t<T \\ \mathbf{q}'=0 & \text{at} \quad t=0 \end{array}\right\} \Rightarrow$$

Small perturbations ϕ' to control ϕ & small perturbations ψ' to disturbance ψ cause small perturbation $\mathbf{q}'$ to state $\mathbf{q}$.

$\mathcal{L}\mathbf{q}' \triangleq \left(E\frac{d}{dt}-A\right)\mathbf{q}'$ is the linearization of the state eqn about the trajectory $\mathbf{q}(\phi,\psi)$.

Cost function (minimize w.r.t. ϕ and maximize w.r.t. ψ):

$$\boxed{\mathcal{J} = \frac{1}{2}\int_0^T (\mathbf{q}^*Q\mathbf{q} + \ell^2\phi^*\phi - \gamma^2\psi^*\psi)\,dt} \quad \Rightarrow \quad \mathcal{J}' = \int_0^T (\mathbf{q}^*Q\mathbf{q}' + \ell^2\phi^*\phi' - \gamma^2\psi^*\psi')\,dt.$$

Statement of adjoint identity. Define inner product $\langle \mathbf{r},\mathbf{q}'\rangle = \int_0^T \mathbf{r}^*\mathbf{q}'\,dt$. Then:

$$\boxed{\langle \mathbf{r}, \mathcal{L}\mathbf{q}'\rangle = \langle \mathcal{L}^*\mathbf{r}, \mathbf{q}'\rangle + \mathbf{b}}$$

with: $\mathbf{r}=$ adjoint, $\mathcal{L}^*\mathbf{r} = \left(-E^*\frac{d}{dt} - A^*\right)\mathbf{r}$, $\mathbf{b} = \mathbf{r}^*E\mathbf{q}'\big|_{t=T} - \mathbf{r}^*E\mathbf{q}'\big|_{t=0}$.

Definition of adjoint equation. Adjoint field easy to compute, though $A=A(\mathbf{q})$.

$$\left\{\begin{array}{ll} \mathcal{L}^*\mathbf{r} = Q\mathbf{q} & \text{on} \quad 0<t<T \\ \mathbf{r}=0 & \text{at} \quad t=T \end{array}\right\} \Leftrightarrow \boxed{\begin{array}{ll} -E^*\dot{\mathbf{r}} = A^*\mathbf{r} + Q\mathbf{q} & \text{on} \quad 0<t<T \\ \mathbf{r}=0 & \text{at} \quad t=T \end{array}}$$

Extraction of gradients. Combining equations, we have:

$$\langle \mathbf{r}, B_\phi\phi' + B_\psi\psi'\rangle = \langle Q\mathbf{q},\mathbf{q}'\rangle \quad \Rightarrow \quad \int_0^T \mathbf{q}^*Q\mathbf{q}'\,dt = \int_0^T \mathbf{r}^*(B_\phi\phi' + B_\psi\psi')\,dt.$$

$$\mathcal{J}' = \int_0^T \left[\left(B_\phi^*\mathbf{r} + \ell^2\phi\right)^*\phi' + \left(B_\psi^*\mathbf{r} - \gamma^2\psi\right)^*\psi'\right]dt \triangleq \int_0^T \left[\left(\frac{\mathscr{D}\mathcal{J}}{\mathscr{D}\phi}\right)^*\phi' + \left(\frac{\mathscr{D}\mathcal{J}}{\mathscr{D}\psi}\right)^*\psi'\right]dt$$

As ϕ' and ψ' are arbitrary, $\boxed{\textbf{the gradient is: } \frac{\mathscr{D}\mathcal{J}}{\mathscr{D}\phi} = B_\phi^*\mathbf{r} + \ell^2\phi, \quad \frac{\mathscr{D}\mathcal{J}}{\mathscr{D}\psi} = B_\psi^*\mathbf{r} - \gamma^2\psi}$.

Riccati analysis for coordinated feedback control

Characterization of saddle point. The control ϕ which minimizes $\mathcal{J}$ and the disturbance ψ which maximizes $\mathcal{J}$ are given by

$$\frac{\mathscr{D}\mathcal{J}}{\mathscr{D}\phi} = 0, \quad \frac{\mathscr{D}\mathcal{J}}{\mathscr{D}\psi} = 0 \quad \Rightarrow \quad \phi = -\frac{1}{\ell^2}B_\phi^*\mathbf{r}, \quad \psi = \frac{1}{\gamma^2}B_\psi^*\mathbf{r}.$$

Combined matrix form. Combining the perturbation and adjoint eqns at the saddle point determined above, assuming $E=I$, gives:

$$\begin{array}{r} \text{Perturbation equation} \rightarrow \\ \text{Adjoint equation} \rightarrow \end{array} \begin{bmatrix} \dot{\mathbf{q}}' \\ \mathbf{r} \end{bmatrix} = \begin{bmatrix} A & \overbrace{-\frac{1}{\ell^2}B_\phi B_\phi^* + \frac{1}{\gamma^2}B_\psi B_\psi^*}^{\text{control and disturbance at saddle point}} \\ -Q & -A^* \end{bmatrix} \begin{bmatrix} \mathbf{q}' \\ \mathbf{r} \end{bmatrix}$$

Solution Ansatz. Relate perturbation $\mathbf{q}'=\mathbf{q}'(t)$ and adjoint $\mathbf{r}=\mathbf{r}(t)$:

$$\boxed{\mathbf{r} = X\mathbf{q}'}, \quad \text{where} \quad X = X(t).$$

Riccati equation. Inserting solution ansatz into the combined matrix form to eliminate $\mathbf{r}$ and combining rows to eliminate $\dot{\mathbf{q}}'$ gives:

$$\left[-\dot{X} = A^*X + XA + X\left(\tfrac{1}{\gamma^2}B_\psi B_\psi^* - \tfrac{1}{\ell^2}B_\phi B_\phi^*\right)X + Q\right]\mathbf{q}'.$$

As this equation is valid for all $\mathbf{q}'$, it follows that:

$$\boxed{-\dot{X} = A^*X + XA + X\left(\tfrac{1}{\gamma^2}B_\psi B_\psi^* - \tfrac{1}{\ell^2}B_\phi B_\phi^*\right)X + Q}.$$

Due to the terminal conditions on $\mathbf{r}$, we must have $\boxed{X=0 \quad \text{at} \quad t=T}$.

Note solutions of this matrix equation satisfy $X^*=X$.

Note also that, by the characterization of the saddle point, we have

$$\psi = \frac{1}{\gamma^2}B_\psi^*\mathbf{r} \quad \text{and} \quad \boxed{\phi = K\mathbf{q}' \quad \text{where} \quad K = -\frac{1}{\ell^2}B_\phi^*X}.$$

This is **the finite-horizon $\mathcal{H}_\infty$ control solution**, and may be solved for linear time-varying (LTV) systems or marched to steady state.

Solving the Riccati equation. Recall $2N \times 2N$ Hamiltonian problem $d\mathbf{v}/dt = Z\mathbf{v}$. The steady-state solution $X > 0$ (of the "CARE") is given by taking the **ordered Schur decomposition** of Z:

$$Z = VTV^{-1}, \quad V = \begin{bmatrix} Q & * \\ R & * \end{bmatrix} = \begin{bmatrix} | & | & & | & \\ \mathbf{v}^1 & \mathbf{v}^2 & \dots & \mathbf{v}^n & * \\ | & | & & | & \end{bmatrix}, \quad \mathbf{v}^i = \begin{bmatrix} \mathbf{q}^i \\ \mathbf{r}^i \end{bmatrix};$$

T is enumerated with its N LHP eigenvalues appearing first. Defining $\mathbf{y} = V^{-1}\mathbf{v}$, we have $d\mathbf{y}/dt = T\mathbf{y}$; stable solutions of $\mathbf{y}$ are spanned by first n columns of T. Since $\mathbf{v} = V\mathbf{y}$, stable solutions of $\mathbf{v}$ are spanned by first n columns of V. To achieve stability of $\mathbf{v}$ via $\mathbf{r} = X\mathbf{q}$ for each of these directions $\mathbf{v}^i$, we have $\mathbf{r}^i = X\mathbf{q}^i$ for $i = 1 \dots n$. Thus:

$$\begin{bmatrix} | & | & & | \\ \mathbf{r}^1 & \mathbf{r}^2 & \dots & \mathbf{r}^n \\ | & | & & | \end{bmatrix} = X \begin{bmatrix} | & | & & | \\ \mathbf{q}^1 & \mathbf{q}^2 & \dots & \mathbf{q}^n \\ | & | & & | \end{bmatrix} \quad \Rightarrow \quad \begin{matrix} R = XQ, \\ X = RQ^{-1}. \end{matrix}$$

A bit of provocative philosophical pondering

Typically, to control high-dimensional discretizations of fluid systems, people first reduce the model in an open loop fashion via an eigenvalue problem (usually a POD, which is just an SVD). Such "open-loop" model reduction methods neglect the control objective, even if they account for B and C in some sort of "balanced" fashion.

Then the control problem is solved, which involves a second eigenvalue problem (i.e., the Schur decomposition of a 2*N* x 2*N* Hamiltonian matrix).

Q: Why TWO eigenvalue problems?? By splitting the problem into two parts, the model is reduced in a manner ignorant of the control problem being considered. That is reckless! We consider here 4 alternatives.

Four approaches to solve large Riccati problems that bypass open-loop model reduction

Chandrasekhar's method

MCE: MINIMUM CONTROL ENERGY STABILIZATION

ADA: THE ADJOINT OF THE DIRECT-ADJOINT

OSSI: OPPOSITELY-SHIFTED SUBSPACE ITERATION

Chandrasekhar's method (Kailath 1973). Consider evolution equations for a low-rank factor $F(t)$ of (dX/dt) and $K(t)$.

$$\frac{dX(t)}{dt} = F_1(t)F_1^H(t) - F_2(t)F_2^H(t) = \begin{pmatrix} F_1 & F_2 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} F_1^H \\ F_2^H \end{pmatrix}$$

Differentiating DRE w.r.t. time and inserting $dX/dt = FHF^H$, we have

$$dK(t)/dt = -R^{-1}B^H F(t) H F^H(t),$$
$$dF(t)/dt = -[A + BK(t)]^H F(t),$$

with terminal conditions

$$K(T) = -R^{-1}B^H Q_T,$$
$$F(T) H F^H(T) = dX(t)/dt\big|_{t=T},$$

where $dX/dt|_{t=T}$ is determined from the original DRE evaluated at $t = T$, and $F(T)$ by its factorization.

UCSD

Four approaches to solve large Riccati problems that bypass open-loop model reduction

Chandrasekhar's method

MCE: MINIMUM CONTROL ENERGY STABILIZATION

ADA: THE ADJOINT OF THE DIRECT-ADJOINT

OSSI: OPPOSITELY-SHIFTED SUBSPACE ITERATION

UCSD

Pole assignment. Consider the ordered Eigen decomposition $Z = V\Lambda V^{-1}$

Define Λ_c with the N desired eigenvalues λ_c of $A+BK$ on the diagonal.

The stable components of the eigen decomposition of Z may be written

$$\begin{bmatrix} A & -BR^{-1}B^H \\ -Q & -A^H \end{bmatrix} V_c = Z V_c = V_c \Lambda_c \quad \text{where} \quad V_c = \begin{bmatrix} X \\ P \end{bmatrix}.$$

We know both Λ_c and Z; we just need to compute V_c. Then, as before,

$$\mathbf{u} = K\mathbf{x} \quad \text{with} \quad K = -R^{-1}B^H W, \quad W = PX^{-1}$$

$$\text{where} \quad AX - BR^{-1}B^H P = X\Lambda_c \quad \text{and} \quad -QX - A^H P = P\Lambda_c.$$

Solving for X and substituting gives

$$AQ^{-1}(A^H P + P\Lambda_c) + BR^{-1}B^H P = Q^{-1}(A^H P + P\Lambda_c)\Lambda_c \qquad (1)$$

$$QX = -(A^H P + P\Lambda_c).$$

Note that (1) is linear in P. Once P is found, calculation of X is trivial.

UCSD

Straightforward simplification of this problem, written in a modal representation of *just the unstable system dynamics*, leads to the following:

Theorem 1 (Lauga & B, *Proc Roy Soc* 2003). *Consider a stabilizable system $d\mathbf{x}/dt = A\mathbf{x} + B\mathbf{u}$ where A has no pure imaginary eigenvalues. Determine the unstable eigenvalues and corresponding left eigenvectors of A such that $Y_u^H A = \Lambda_u Y_u^H$ (i.e., determine the unstable eigenvalues and corresponding right eigenvectors of A^H such that $A^H Y_u = Y_u \Lambda_u^H$). Define $\bar{B}_u = Y_u^H B$ and $C = \bar{B}_u R^{-1} \bar{B}_u^H$, and compute a matrix X with elements $x_{ij} = c_{ij}/(\lambda_i + \lambda_j^*)$. The minimal-energy stabilizing feedback controller is then given by $\mathbf{u} = K\mathbf{x}$, where $K = -R^{-1}\bar{B}_u^H X^{-1} Y_u^H$.*

UCSD

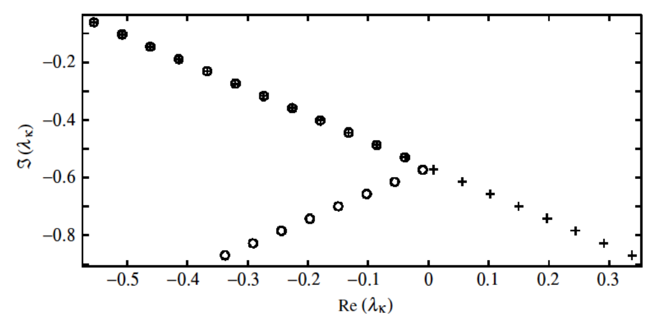


Figure 4. Locus of the first 20 eigenvalues of the CGL operator with supercriticality $\delta = 3$ before (pluses) and after (circles) optimal control is applied (with $\ell = 10^4$ and $x_f = 47$). Note that, in this minimal-energy optimal control setting, the stable eigenmodes of the system matrix are unchanged, and the unstable eigenvalues of the system matrix are reflected across the imaginary axis.

Lauga & B, Proc Roy Soc 2003

Four approaches to solve large Riccati problems that bypass open-loop model reduction

Chandrasekhar's method

MCE: MINIMUM CONTROL ENERGY STABILIZATION

ADA: THE ADJOINT OF THE DIRECT-ADJOINT

OSSI: OPPOSITELY-SHIFTED SUBSPACE ITERATION

For any $\mathbf{x}_0$, adjoint optimization of $\mathbf{u}$ for minimization of $J(\mathbf{u})$ proceeds as before; when converged, $DJ/D\mathbf{u} = 0$ and $\mathbf{u} = -R^{-1}B^H\mathbf{p}$ on $t \in [0,T]$. The input is $\mathbf{x}_0$; focus for now on the output $\mathbf{u}$ at time $t = 0$, denoted $\mathbf{u}_0$. If $\mathbf{x} = \mathbf{x}_{n\times 1}$ and $\mathbf{u} = \mathbf{u}_{m\times 1}$ and we solve this problem n times, then

$$\begin{bmatrix}\mathbf{u}_0^1 & \mathbf{u}_0^2 & \dots & \mathbf{u}_0^n\end{bmatrix} = K_0 \begin{bmatrix}\mathbf{x}_0^1 & \mathbf{x}_0^2 & \dots & \mathbf{x}_0^n\end{bmatrix}. \qquad (1)$$

Solving for K_0 determines the feedback gain matrix K at time $t = 0$,

$$\mathbf{u}(0) = K_0\mathbf{x}(0).$$

This approach requires n optimizations to set up (1), which may be solved to compute the $m \times n$ matrix K_0. If $n \gg m$, it is more efficient to consider the adjoint of this problem, thus leading to an algorithm requiring only m optimizations to compute K_0.

Taking $Q \geq 0$ and $R > 0$, define

$$\mathbf{y} = \begin{bmatrix}\mathbf{p}\\ \mathbf{x}\end{bmatrix} \quad \text{and} \quad \mathbf{L} = \begin{bmatrix} BR^{-1}B^H & d/dt - A \\ -d/dt - A^H & -Q\end{bmatrix}.$$

The converged solution of the "forward TPBVP" may now be written

$$\mathbf{L}\mathbf{y} = 0 \quad \text{with} \quad \begin{cases}\mathbf{x}(0) = \mathbf{x}_0,\\ \mathbf{p}(T) = 0.\end{cases}$$

Defining $\langle \mathbf{a}, \mathbf{b}\rangle = \int_0^T \mathbf{a}^H\mathbf{b}\,dt$, we may express the adjoint identity

$$\langle \tilde{\mathbf{y}}, \mathbf{L}\mathbf{y}\rangle = \langle \mathbf{L}^*\tilde{\mathbf{y}}, \mathbf{y}\rangle + b \quad \text{where} \quad \tilde{\mathbf{y}} = \begin{bmatrix}\tilde{\mathbf{p}}\\ \tilde{\mathbf{x}}\end{bmatrix}.$$

Using integration by parts, it follows that $\mathbf{L}^* = \mathbf{L}$ ($\mathbf{L}$ is "self-adjoint"), and

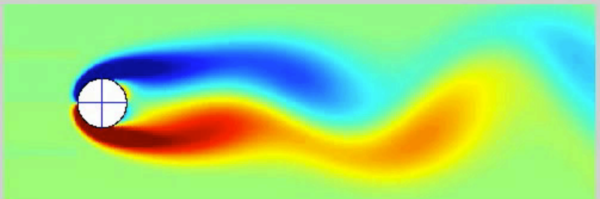
$$b = (\tilde{\mathbf{p}}^H\mathbf{x} - \tilde{\mathbf{x}}^H\mathbf{p})_{t=T} - (\tilde{\mathbf{p}}^H\mathbf{x} - \tilde{\mathbf{x}}^H\mathbf{p})_{t=0}.$$

Now use $\mathbf{L}^*$ to define an appropriate "adjoint TPBVP"

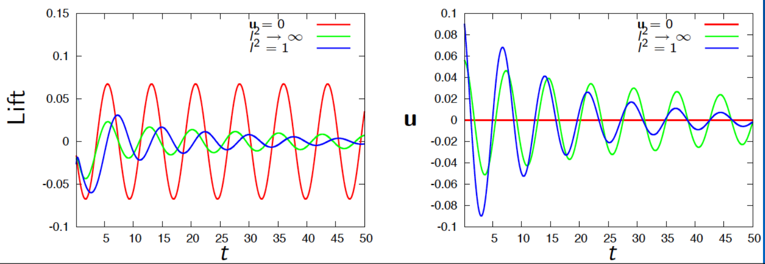
$$\mathbf{L}^*\tilde{\mathbf{y}} = 0 \quad \text{with} \quad \begin{cases}\tilde{\mathbf{x}}(0) = \tilde{\mathbf{x}}_0,\\ \tilde{\mathbf{p}}(T) = 0,\end{cases}.$$

Note that this TPBVP is *exactly* the same as that given previously, it is just written with different variables and has a different interpretation given to the "input" $\tilde{\mathbf{x}}(0)$ and the "output" $\tilde{\mathbf{p}}(0)$. Thus, this "adjoint TPBVP" may also be solved using the same machinery as before.

Substituting into the adjoint identity leads to $[\tilde{\mathbf{p}}(0)]^H\mathbf{x}(0) = [\tilde{\mathbf{x}}(0)]^H\mathbf{p}(0)$. Recall also that $\mathbf{u}(0) = [K_0^H]^H\mathbf{x}(0) = [-BR^{-1}]^H\mathbf{p}(0)$. Thus, setting $\tilde{\mathbf{x}}(0)$ to the first column of $[-BR^{-1}]$ and solving the adjoint TPBVP (iterating all the way to convergence!), the resulting value of $\tilde{\mathbf{p}}(0)$ is simply the first column of K_0^H, etc. That is, after solving the adjoint TPBVP just m times (note: not n times!), the entire K_0 is constructed directly.



Test case: $Re = 55$, control is turned on at $t = 0$



Four approaches to solve large Riccati problems that bypass open-loop model reduction

- Chandrasekhar's method
- MCE: MINIMUM CONTROL ENERGY STABILIZATION
- ADA: THE ADJOINT OF THE DIRECT-ADJOINT
- OSSI: OPPOSITELY-SHIFTED SUBSPACE ITERATION

Prototype subspace iteration algorithm. A basic power method may be visualized as an EE march to convergence

$$d\mathbf{v}/dt = A\mathbf{v} \quad \Rightarrow \quad \mathbf{v}^{k+1} = (I + hA)\mathbf{v}^k = \mathbf{v}^k + hA\,\mathbf{v}^k.$$

A basic block power method may be written $V \leftarrow V + hAV$ with Gram-Schmidt applied at each step. The emerging eigenvalues may be estimated with the Rayleigh quotient ($\sigma = \mathbf{v}^H A\,\mathbf{v}$), which for multiple vectors may be written $\Sigma = V^H AV$. Schur decomposition of Σ then gives

$$\Sigma = \overline{U}\,\overline{T}\,\overline{U}^H = V^H AV \quad \Rightarrow \quad \overline{T} = (V\overline{U})^H A\,(V\overline{U}),$$

and thus

$$V \leftarrow (V\overline{U}) \quad \text{and} \quad \Sigma \leftarrow \overline{T}$$

keeps Σ triangular (and the columns of V orthogonal) as the iterations proceed.

As a further refinement, we may apply a small shift of $-(hV\Sigma)$ to the march for V:

$$V \leftarrow V + h(AV - V\Sigma).$$

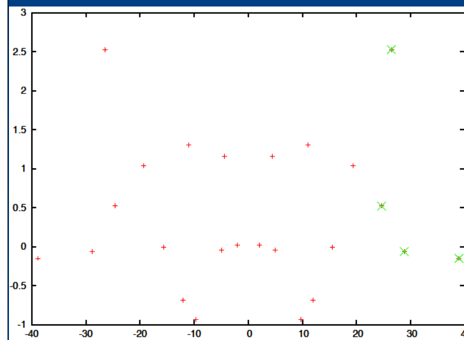
This shifted form has the same essential effect as before (that is, preferential focusing of the columns of V in the direction of the leading Schur vectors of A), with the added benefit of ensuring that the update to V itself approaches zero as V approaches a basis of a set of Schur vectors, and thus $(AV - V\Sigma)$ approaches zero.

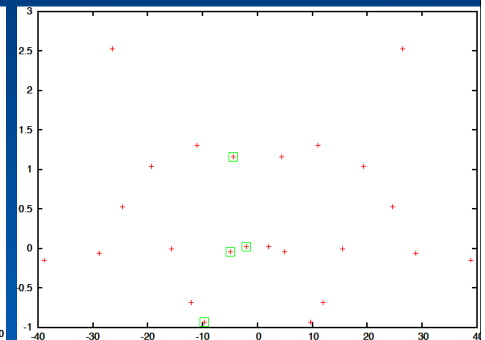
We now extend such methods to find the *central* eigenmodes of a large Hamiltonian matrix Z. We seek the least-stable LHP Schur vectors V, and partition them into their state and adjoint components:

$$Z = \begin{bmatrix} A & -BR^{-1}B^H \\ -Q & -A^H \end{bmatrix}, \qquad V = \begin{bmatrix} X \\ P \end{bmatrix}.$$

We will approximate the resulting feedback gain matrix K using the Moore-Penrose pseudoinverse $K = -R^{-1}B^H(PX^+)$.

The working hypothesis is that the neglected (well-damped) closed-loop Schur vectors of Z likely play a reduced role in the full computation of K [idea motivated by Amodei & Buchot (2010, 2011)].





Standard subspace iteration
-> *leading* (least-damped) eigenvalues

Oppositely-shifted subspace iteration (OSSI)
-> *central* eigenvalues of Hamiltonian matrix

All current subspace iteration methods converge to *extremal* eigenvalues. Convergence to the *central* eigenvalues of Z, using existing algorithms, requires computation of $Z^{-1}\mathbf{x}$, which is expensive. We seek a method to find the central Schur vectors of Z without access to Z^{-1}.

In the MCE case, one of the off-diagonal terms of Z is zero, and the subspace iteration algorithm described above gives (a) the least-stable eigenvalues of A, and (b) the least-stable eigenvalues of $-A^H$. For the control of a system with a few unstable eigenvalues and many stable eigenvalues extending off into the LHP, we need (a), which are the eigenvalues of A near the imaginary axis. Rather than (b), however, we need to the *most*-stable eigenvalues of $-A^H$. These may be found simply by *changing the sign* of the related march!

The eigenvalues of Z vary continuously as its elements are varied. If both Q and $BR^{-1}B^H$ are nonzero but the norm of their product is small (that is, a modest generalization from the MCE limit), application of a slightly modified subspace iteration algorithm, with the adjoint component marching the opposite direction in time, returns those eigenvalues of Z near the least stable eigenvalues of A together with those eigenvalues of Z near the most stable eigenvalues of $-A^H$.

OSSI via an EE discretization. The idea described above is straightforward to implement. The update to X and P is split into two parts, with a *positive* sign in the shift of X, and a *negative* sign in the shift of P:

$$X \leftarrow X + hX_1, \quad P \leftarrow P - hP_1$$

where, leveraging the structure of Z,

$$X_1 = AX - (BR^{-1}B^H)P, \quad P_1 = -QX - A^H P.$$

As before, to accelerate convergence, we may instead apply shifts via

$$X \leftarrow X + h(X_1 - X\Sigma), \quad P \leftarrow P - h(P_1 - P\Sigma).$$

The resulting algorithm is referred to as *opposite shifting*.

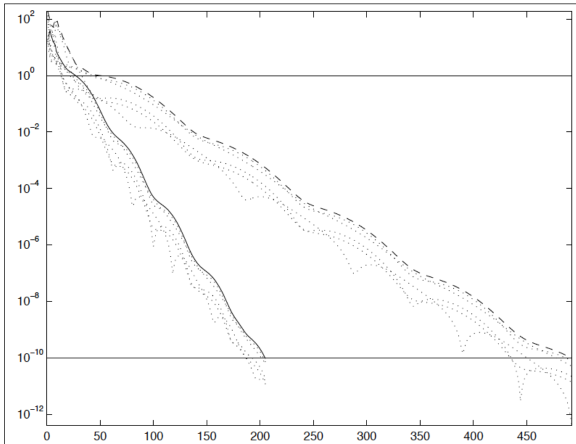


Fig. 2. Convergence of (dashed) the explicit form given in Algorithm 4, and (solid) the implicit form given in Algorithm 5, of the subspace iteration algorithm for the leading Schur vectors of the Hamiltonian matrix Z. The dotted lines indicate the convergence of individual modes.

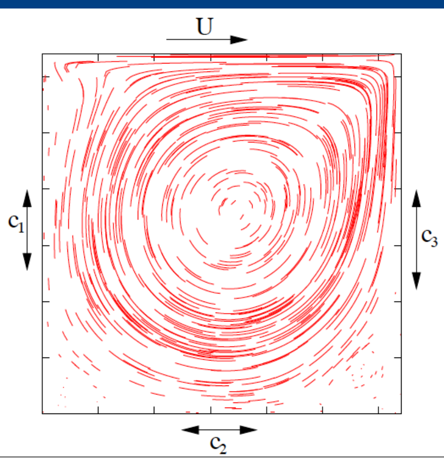
A randomly-generated LQR problem with A, $Q > 0$, $R = I$ and $\bar{B} = BR^{-1}B^H \geq 0$ was defined via

$$A_1 = \text{randn}(n,n); \quad A = -A_1 A_1^H + \text{randn}(n,n);$$
$$Q_1 = \text{randn}(n,n); \quad Q = \alpha Q_1 Q_1^H;$$
$$\bar{B}_1 = \text{randn}(n,m); \quad \bar{B} = \beta \bar{B}_1 \bar{B}_1^H,$$

The system matrix so created has real and complex conjugate pairs of eigenvalues extending off into the LHP, and a few unstable eigenvalues.

Results at left are typical, with parameters: $n = 10$, $m = 4$, $\alpha = 0.1$, and $\beta = 0.01$.

Driven cavity



Leading eigenvalues:
-0.128396-435.659i
-0.128396+435.659i
0.128396+435.659i
0.128396-435.659i
-19.4557
19.4557
-248.216-680.951i
-248.216+680.951i

...

The iteration converges!

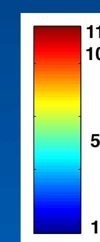
Four approaches to solve large Riccati problems that bypass open-loop model reduction

Chandrasekhar's method

MCE: MINIMUM CONTROL ENERGY STABILIZATION

ADA: THE ADJOINT OF THE DIRECT-ADJOINT

OSSI: OPPOSITELY-SHIFTED SUBSPACE ITERATION



Provocative bottom line message: These alternatives work. There is no need to risk throwing out the baby with the bathwater via open-loop model reduction.